From the DEPARTMENT OF LABORATORY MEDICINE
Karolinska Institutet, Stockholm, Sweden

MOLECULAR INSIGHTS INTO MITOCHONDRIAL DNA REPLICATION

Javier Miralles Fusté

Karolinska Institutet

Stockholm 2012

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by Aidla Trading AB/Kompendiet

ISBN 978-91-7457-752-5

A la meva àvia Mercedes

ABSTRACT

Mitochondria are organelles found in eukaryotic cells. These organelles produce most of the adenosine triphosphate that cells use as a source of energy. Mitochondria contain their own genomic material, a circular DNA genome (mtDNA) that encodes subunits of the respiratory chain complexes and RNA components needed for mitochondrial translation. Many aspects of mtDNA replication are still not understood and in this thesis we address some of the molecular mechanisms of this process in mammalian cells.

DNA synthesis cannot be initiated de novo, but requires a short RNA primer as a starting point. We here demonstrate that the mitochondrial RNA polymerase (POLRMT) is the primase required for initiation of DNA synthesis from the origin of light strand DNA replication (OriL) in human mtDNA. Using purified POLRMT and the core factors of the mitochondrial replisome, we faithfully reconstitute OriL-dependent initiation of replication in vitro. During origin activation, OriL is exposed in its single-stranded conformation and adopts a stem-loop structure. POLRMT initiates primer synthesis from a poly-dT stretch in the single-stranded loop region and after about 25 nt, POLRMT is replaced by the mitochondrial DNA polymerase γ (POLγ) and DNA synthesis is initiated. Our findings also suggest that the mitochondrial single-stranded DNA binding protein directs origin-specific initiation by efficiently blocking unspecific initiation events in other regions of the mtDNA genome.

To analyze the requirements of OriL in vivo, we have used saturation mutagenesis in the mouse combined with in vitro biochemistry and demonstrated that OriL is essential for mtDNA maintenance. OriL requires a stable stem-loop structure and a pyrimidine-rich sequence in the template strand for proper origin function. The OriL mechanism appears to be conserved, since bioinformatics analyses demonstrated the presence of OriL in the mtDNA of most vertebrates including birds. Our findings suggest that mtDNA replication may be performed by a common mechanism in all vertebrates and lend support to the strand-displacement model for mtDNA replication.

A molecular understanding of the mitochondrial DNA replication machinery is also of medical importance. Today, more than 160 mutations in the gene encoding the catalytic subunit of POLγ (POLγA) have been associated with human disease. One example is the Y955C mutation, which causes autosomal dominant progressive external ophthalmoplegia, a disorder characterized by the accumulation of multiple mtDNA deletions. The Y955C mutation decreases POLγ processivity due to a decreased binding affinity for the incoming deoxyribonucleoside triphosphate. However, it is not clear why this biochemical defect leads to a dominant disease. We have used the reconstituted mammalian mtDNA replisome and studied functional consequences of the dominant Y955C mutation. Our study revealed that the POLγA:Y955C enzyme is prone to stalling at dATP insertion sites and instead enters a polymerase/exonuclease idling mode. The mutant POLγA:Y955C competes with wild-type POLγA for access to the primer template. However, once assembled in the replisome, the wild-type enzyme is no longer affected. Our data therefore provide a mechanism for the mtDNA replication phenotypes seen in patients harboring the Y955C mutation.

LIST OF PUBLICATIONS

I. Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis *in vitro*.
Wanrooij S, **Fusté JM**, Farge G, Shi Y, Gustafsson CM, Falkenberg M.
Proc Natl Acad Sci U S A. 2008 Aug 12;105(32): 11122-7.

II. Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication.
Fusté JM, Wanrooij S, Jemt E, Granycome CE, Cluett TJ, Shi Y, Atanassova N, Holt IJ, Gustafsson CM, Falkenberg M.
Mol Cell. 2010 Jan 15;37(1):67-78.

III. *In vivo* mutagenesis reveals that OriL is essential for mitochondrial DNA replication.
Wanrooij S*, **Fusté JM***, Stewart JB, Wanrooij PH, Samuelsson T, Larsson NG, Gustafsson CM, Falkenberg M.
Submitted manuscript 2012.
*Shared first Authorship

IV. Sequence-specific stalling of DNA polymerase γ and the effects of mutations causing progressive ophthalmoplegia.
Atanassova N*, **Fusté JM***, Wanrooij S, Macao B, Goffart S, Bäckström S, Farge G, Khvorostov I, Larsson NG, Spelbrink NJ, Falkenberg M.
Hum Mol Genet. 2011 Mar 15;20(6):1212-23. * Shared first Authorship

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ad	Autosomal dominant
ADP	Adenosine diphosphate
ANT	Adenine nucleotide transporter
ATAD3	AAA Domain containing 3 protein
ATP	Adenosine triphosphate
ar	Autosomal recessive
bp	Base pair
CoQ	Coenzyme Q10 or ubiquinone
COX	Cytochrome c oxidase
CSB	Conserved sequence block
C-terminal	Carboxy terminal
CTP	Cytidine triphosphate
Cyt b	Cytochrome b
D-loop	Displacement loop
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
dsDNA	Double-stranded DNA
FRET	Fluorescence resonance energy transfer
GTP	Guanosine triphosphate
HSP	Heavy strand promoter
Hsp60	Heat Shock protein 60
H-strand	Heavy strand
kb	Kilo base
kDA	Kilodalton
LSP	Light strand promoter
L-strand	Light strand
MDS	Mitochondrial DNA depletion syndrome
mRNA	Messenger ribonucleic acid
mtDNA	Mitochondrial DNA
mtSSB	Mitochondrial single-stranded DNA-binding protein
nt	Nucleotides
N-terminal	Amino terminal
NTP	Nucleotide triphosphate
OriH	Origin of H-strand replication

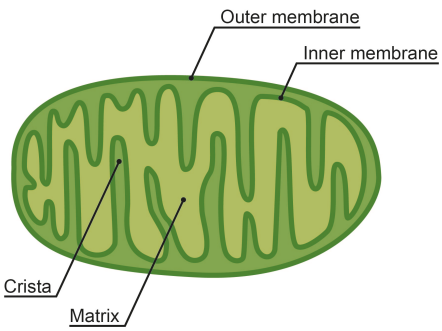
OriL	Origin of L-strand replication
OXPHOS	Oxidative phosphorylation
PEO	Progressive external opthalmaplegia
POLγ	Mitochondrial DNA polymerase gamma (holoenzyme)
POLRMT	Mitochondrial RNA polymerase
RC	Respiratory chain
RITOLS	Ribonucleotide incorporation throughout the lagging strand
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
ssDNA	Single-stranded DNA
TFB1M	Mitochondrial transcription factor B1
TFB2M	Mitochondrial transcription factor B2
TFAM	Mitochondrial transcription factor A
tRNA	Transfer RNA
$tRNA^{Phe}$	Transfer RNA for Phenylalanine
$tRNA^{Val}$	Transfer RNA for Valine
Vero cells	Cell line derived from the kidney of the African gree monkey
WANCY	Tryptophan-Alanine-Asparagine-Cysteine-Tyrosine
2DNAGE	Two dimensional neutral/neutral agarose gel electrophoresis

INTRODUCTION

Mitochondria

Mitochondria are cytoplasmic organelles essential for eukaryotic cells. Most of the ATP required for cellular function is produced within mitochondria, in a process known as Oxidative Phosphorylation (OXPHOS). In this process, the energy derived from the breakdown of carbohydrates and fatty acids is converted into ATP, the major source of cellular energy. Mitochondria are comprised of a double membrane system (Figure 1); the outer mitochondrial membrane separates the organelle from the cytosol. Invaginations in the inner mitochondrial membrane form the cristae where the respiratory chain enzyme complexes are located. The matrix contains the mitochondrial genome (mtDNA) as well as enzymes responsible for many of the central reactions of oxidative metabolism. Although most of the proteins localized in mitochondria are encoded in the nuclear DNA, the proteins encoded in mtDNA are essential for mitochondrial function. MtDNA maintenance is therefore essential for survival of the eukaryotic cell. Mutations and defects in mtDNA are related to mitochondrial disorders, ageing and infertility (Schapira, 2006).

Figure 1. Structure of mitochodria.



Origin

The existence of a separate mitochondrial genome is explained by the broadly accepted endosymbiotic theory, which postulates that mitochondria are the direct descendents of an α–Proteobacterium that once entered into a primitive archaea host cell establishing an endosymbiotic relationship, approximately 2 billion years ago (Gray et al., 1999). Subsequently, the endosymbiont has suffered a massive gene transfer to the nucleus and loss of redundant genes during evolution. Why not all mitochondrial genes have been transferred to the nucleus and why the cell has maintained dedicated systems for mtDNA expression and maintenance through evolution is not completely understood. One possible answer could be that some mtDNA encoded proteins, cannot be imported through the mitochondrial membranes because of their highly hydrophobic nature (von Heijne, 1986)

However, the most accepted hypothesis states that the bi-genomic organization offers substantial compensating benefits, probably in the form of greatly enhanced energy production (Lane and Martin, 2010). The co-location for redox regulation hypothesis suggests that mitochondria have retained those genes whose function in electron transport demands a rapid, direct, and unconditional redox regulatory control for their biosynthesis (Allen, 2003). The fact that mtDNA content correlates with respiratory capacity and ATP availability lends support to this hypothesis (Allen, 1993; Williams, 1986).

Interestingly, in contrast to what one would expect, key components of the mitochondrial replication and transcription machineries are derived from the T-odd lineage of bacteriophages rather than from a α-Proteobacterium as the endosymbiotic theory would predict. Three key factors – the mitochondrial RNA polymerase (POLRMT), the replicative helicase (TWINKLE) and the DNA polymerase γ (POLγ) appear to have shared ancestry with proteins encoded by T-odd bacteriophages. They share high sequence similarity with their respective T-odd homologs. Therefore, the endosymbiosis that originated mitochondria took place with the contribution of genetic information from three different entities: host, eubacterium and phage (Shutt and Gray, 2006a)

Mitochondrial dynamics

Mitochondria constantly fuse and divide to form a dynamic interconnected network. The precise balance between fusion and fission regulates the different mitochondrial morphologies that can be found in specific cell-types, or under different physiological conditions. Enhanced fusion will produce elongated interconnected mitochondria (in metabolically active cells), whereas enhanced fission will lead to small fragmented mitochondria (during the cell cycle or apoptosis). Mitochondrial fusion plays a protective role, facilitating energy transmission (Skulachev, 2001), and genetic complementation between mitochondria (Sato et al., 2006). Disruption of mitochondrial fusion causes a severe defect in respiratory capacity and loss of inner mitochondrial membrane potential (Chen and Chan, 2005; Chen et al., 2003). In addition, fusion is required for stable mtDNA maintenance and it appears to counteract the deleterious effects caused by mtDNA mutations (Chen and Chan, 2010).

Many of the key factors involved in fusion and fission have been extensively studied in yeast and most of them are conserved in mammals. The large GTPases Mitofusin 1 (Mfn1), Mitofusin2 (Mfn2), and Optic atrophy 1 (OPA1), are members of the dynamin family and are the main factors controlling the fusion process in mammals. Mfn1 and Mfn2 are responsible for fusion of the outer mitochondrial membrane. OPA1 is localized in the inner mitochondrial membrane and it is believed to work together with mitofusins to promote fusion. Dynamin 1 (Dnm1) is the major regulator of mitochondrial fission in yeast (Bleazard et al., 1999). Like its human homolog Dynamin-related protein 1 (Drp1), it assembles at division sites around the outer mitochondrial membrane forming multimeric ring complexes that drive the fission

process. Defects of mitochondrial fusion and fission affect especially the central nervous system and mutations in these proteins have been associated with neurodegenerative disorders, developmental problems, neuropathies, obesity, and vascular diseases (Chen and Chan, 2005).

The respiratory chain

During the OXPHOS process, the energy released from the breakdown of nutrients (carbohydrates, lipids and amino acids) is used to produce ATP. The breakdown of glucose, lipids and amino acids generates acetyl-CoA, which enters the Krebs cycle in the mitochondrial matrix. The Krebs cycle enzymes couple oxidation of acetyl-CoA to generation of reduced forms of the electron carriers NADH and FADH2, which feed the respiratory chain (RC). The RC complexes are four large multi-heteromeric protein complexes (complex I-IV) located within the inner mitochondrial membrane (Figure 2). These protein complexes transport the electrons derived from NADH and FADH2 via a series of redox reactions to finally reduce oxygen to water. Two small electron carriers, coenzyme Q and cytochrome *c* transfer the electrons between RC complexes. Electron transfer is coupled to proton pumping from the matrix across the inner membrane to the intermembrane space generating an electrochemical gradient. A fifth complex, the ATP synthase or complex V uses the favourable return of these protons to the matrix as a motor driving force to synthesize ATP, a mechanism known as chemiosmotic coupling. The ATP that is not used inside the mitochondria is then transported to the cytosol across the inner membrane in exchange for ADP by a transmembrane protein called the adenine nucleotide translocase (ANT) (Saraste, 1999).

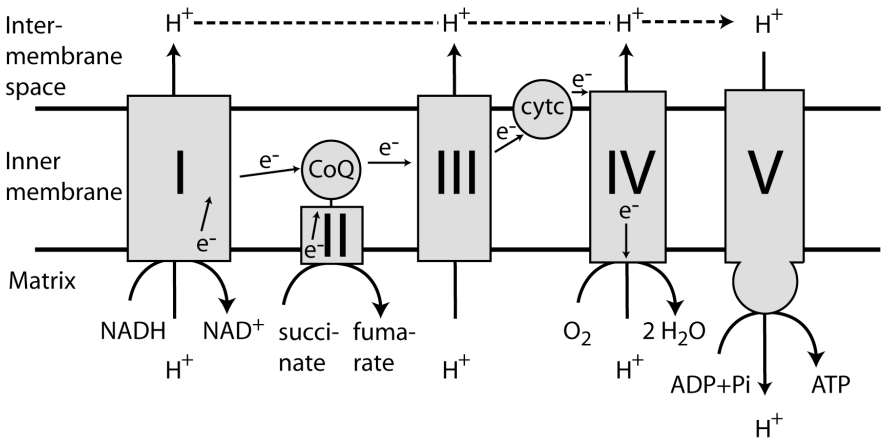


Figure 2. Structure and function of the respiratory chain.
The respiratory chain consists of five different enzyme complexes (complexes I-V), coenzyme Q (CoQ) and cytochrome *c* (cyt c). NADH and succinate are oxidized by complexes I and II, the electrons are transferred to coenzyme Q, complex III, cytochrome *c*, complex IV, and finally to molecular oxygen (1/2 O_2), which is reduced to water (H_2O). Protons (H^+) are pumped out of the mitochondrial matrix by complexes I, III and IV and a proton gradient is formed across the inner membrane of the mitochondrion. The protons reenter the matrix through complex V (ATP synthase) and the energy of the proton gradient is used to synthesize ATP (Adapted from (Larsson and Clayton, 1995)).

Interestingly, the OXPHOS system depends on both the nuclear and mitochondrial genome. Most of the genes (about 70) are encoded in nuclear DNA whereas mtDNA codes for 13 subunits that are part of complexes I, III, IV, V. Complex II is only encoded by nuclear genes. The five RC complexes are organized into higher-order structures called supercomplexes that are thought to be important in channelling electron flow and controlling ROS production. Recently, the first factor controlling supercomplex assembly was identified, Rcf1 in yeast and HIG2A (a homolog of Rcf1) in mammals (Chen et al., 2012; Strogolova et al., 2012). Rcf1 is a subunit of the Cytochrome C oxidase and is required for assembly of the Complex III-Complex IV supercomplex. Disruption of the RCF1 gene caused impaired respiration and mitochondrial oxidative stress (Chen et al., 2012).

The mitochondrial genome

The first evidence for a separate mitochondrial genome appeared in 1960 and was based on electron microscopy pictures of chick embryonic cells (Nass and Nass, 1963). Human mtDNA was the first complete genome to be sequenced (in 1981) and it consists of a small, closed, circular, double stranded molecule of about 16.5 kb in size (Anderson et al., 1981) (Figure 3). The two strands differ in their base composition and it is possible to separate them in a cesium chloride gradient. For this reason, they are named the heavy or H-strand (G-rich) and the light or L-strand (C-rich). In addition to 13 subunits components of the respiratory chain complexes, the mitochondrial genome also encodes for two ribosomal rRNAs (12S and 16S) and 22 tRNAs. The majority of the genes are encoded on the H-strand, whereas the L-strand is poor in gene content encoding just for ND6 mRNA and 8 tRNAs.

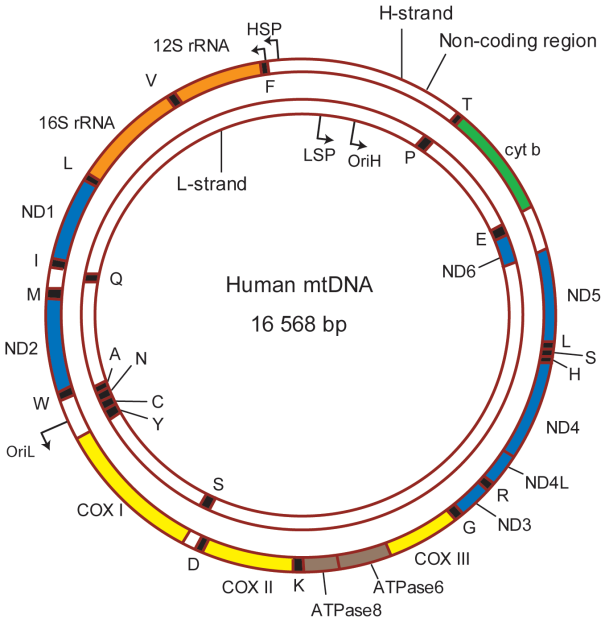


Figure 3. The human mitochondrial genome.
The H-strand and L-strand are represented by circles. The following genes are indicated: NADH dehydrogenase genes (ND1-6 and ND4L) are shown in blue; cytochrome *b* (cyt*b*) gene is shown in green; cytochrome *c* oxidase genes (COXI-III) are shown in yellow; ATP synthase genes (ATP8 and ATP6) are shown in brown; tRNA genes in black and rRNA genes in orange. The arrows represent the heavy (HSP) and light (LSP) promoters and the origins of heavy (OriH) and light (OriL) strand replication. (Adapted from (Wanrooij and Falkenberg, 2010))

The gene organization of mtDNA is very compact, lacking introns. Some of the protein coding genes overlap and the stop codons are created by polyadenylation (Anderson et al., 1981). The genome contains only two non-coding regions, which have regulatory functions. The largest one is the displacement loop (D-loop) or control region, which contains the promoters for light- and heavy-strand transcription (LSP and HSP respectively) and the heavy-strand origin of replication (OriH). This region is characterized by a triple-stranded structure, the D-loop, formed by a nascent H-strand that terminates prematurely and remains annealed to the parental L-strand. This nascent strand known as 7S DNA. It is about 700 nt long and displaces the parental H-strand as a loop (Shadel and Clayton, 1997) (Figure 4).

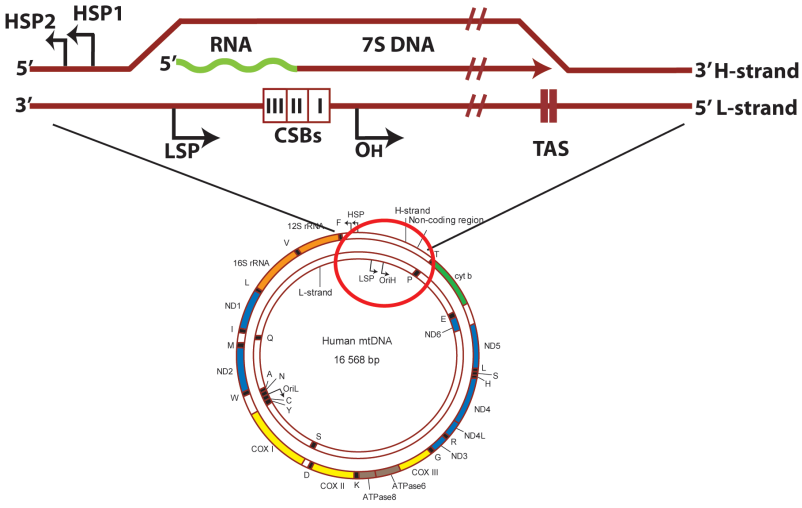


Figure 4. The displacement loop (D-loop) (see text for details).

Although the D-loop region is highly variable between species, some conserved sequences are found and believed to be involved in mtDNA replication. However, the exact function of these sequences is only partially understood. The second non-coding region is the light-strand origin of mtDNA replication (OriL) located within a tRNA cluster. It is 30 bp long and adopts a secondary stem-loop structure in its single-stranded conformation (Tapper and Clayton, 1981).

Nucleoids

Mitochondrial DNA is highly packed and organized in DNA-protein complexes known as nucleoids. Nucleoids are considered the heritable units of mtDNA. They are distributed evenly throughout the mitochondrial network, and they may be associated with the inner mitochondrial membrane. There is no consensus regarding the number of mtDNA molecules present per nucleoid. However, a recent study suggested that mammalian nucleoids contain one mtDNA molecule. Using super-resolution microscopy and quantitative PCR they calculated on average 1.4 mtDNA molecules per nucleoid (Kukat et al., 2011).

More than 30 proteins have been identified to co-localize with nucleoids. However, only one factor is recognized as a core component of the nucleoids. This protein is a histone-like high mobility group (HMG) box protein denoted Abf2 in yeast and mitochondrial transcription factor A (TFAM) in humans (Fisher et al., 1992). TFAM binds, wraps and coats the duplex mtDNA. Structural studies have shown that TFAM organizes mtDNA by introducing repeated U-turns (Rubio-Cosials et al., 2011), and increasing amounts of TFAM compact DNA into nucleoid-like structures *in vitro* (Kaufman et al., 2007). TFAM is also an essential component of the transcription machinery. It binds to the transcription promoters in a sequence-specific manner and promotes transcription (Fisher and Clayton, 1985; Fisher et al., 1989).

Several other proteins have been associated with nucleoids. Some of these have a previously known function in mtDNA transactions i.e. POLγ, TWINKLE, mtSSB, POLRMT (Wang and Bogenhagen, 2006). Other proteins like chaperones, proteases, lipid metabolic enzymes or cytoskeletal components have also been associated to the nucleoids. Some of these proteins are recruited to nucleoids during specific metabolic conditions, suggesting a link between mitochondrial biogenesis and nucleoid activities (Kucej and Butow, 2007).

Transcription of mtDNA

Transcription of mtDNA in mammals is initiated from two specific sites within the control region, the heavy strand (HSP) and the light strand (LSP) promoters (Montoya et al., 1982) giving rise to long polycistronic precursors RNA. The primary transcripts are processed by cleavage of the tRNAs which generates the individual mRNA, tRNA, rRNA molecules (Ojala et al., 1981). Transcription from HSP produces the 12S and 16S rRNAs, 14 tRNAs and all the mitochondrial mRNAs except ND6, whereas transcription from LSP produces 8 tRNAs and ND6 mRNA. A second initiation site has been suggested for H-strand transcription, HSP2, located close to the 5´ end of 12S rRNA. However, *in vitro* studies have questioned whether HSP2 is a functional promoter and its existence is under debate (Litonin et al., 2010). Recently, it has been suggested that HSP2 exists but is regulated in an antagonistic manner compared to LSP and HSP1. HSP2 would be active but in the absence of the mitochondrial transcription factor A (TFAM), which is required for transcription initiation at LSP and HSP1(Lodeiro et al., 2012; Zollo et al., 2012). However, the transcription start site for HSP2 mapped *in vitro* does not correlate exactly with that previously observed *in vivo*. Moreover, mtDNA is always coated with TFAM, raising doubts about the relevance of the *in vitro* findings.

The mitochondrial genome is transcribed by a nuclear-encoded specialized machinery, distinct from that found in the nucleus. Mitochondrial transcription can be reconstituted *in vitro* from purified proteins (Falkenberg et al., 2002). The main factors required for mtDNA transcription are the mitochondrial RNA polymerase (POLRMT), TFAM and TFB2M. The mitochondrial RNA polymerase, which was first identified in yeast

(Greenleaf et al., 1986), is a single subunit enzyme that displays high sequence similarity to the C-terminal part of the RNA polymerases encoded by the T-odd lineage of bacteriophages (Masters et al., 1987; Tiranti et al., 1997). The mitochondrial enzyme also contains a unique N-terminal domain that is absent in its T-odd bacteriophage homologs. In yeast, deletion of the N-terminal region impairs mtDNA genome stability. Interestingly, this deletion does not affect mitochondrial transcription initiation *in vivo*, but may instead link transcription to control of gene expression. The N-terminal extension interacts with Nam1, a protein that affects splicing of introns in budding yeast. The role of the N-terminal extension in higher eukaryotes remains to be established. A recent POLRMT crystal structure showed that the N-terminal domain includes two pentatricopeptide repeat (PPR) motifs (Ringel et al., 2011). The function of the PPR domain in POLRMT is not known, but it could be involved in coupling transcription to post-transcriptional processes since PPR motifs have been found in many plant RNA-binding proteins that are key regulators in plant mitochondrial gene expression (involved in RNA editing, RNA splicing, RNA cleavage and translation) (Schmitz-Linneweber and Small, 2008).

POLRMT can specifically recognize mitochondrial promoters. However, in contrast to the bacteriophage T7 RNA polymerase, it cannot initiate promoter-specific transcription without the assistance of the accessory factors TFAM and TFB2M (Gaspari et al., 2004). The first *in vivo* evidence that TFB2M is essential for mtDNA transcription was provided by studies in *Drosophila melanogaster*, where TFB2M gene silencing resulted in abolishment of mitochondrial transcripts (Adan et al., 2008). The third transcription factor, TFAM binds specifically to a region upstream of the HSP and LSP transcription initiation sites (Fisher and Clayton, 1985). TFAM causes a dramatic bending of the mtDNA at the promoter region resulting in DNA melting and allowing the recruitment of POLRMT and TFB2M to form the initiation complex (Sologub et al., 2009). Recently, some reports have claimed that TFAM is not a core component of the mitochondrial transcription machinery and showed TFAM-independent transcription initiation *in vitro* (Shutt et al., 2010). However, recent results from our laboratory have demonstrated that TFAM is indeed required for transcription initiation, but that low levels of TFAM-independent transcription can be achieved under conditions that favor DNA breathing, e.g. at unphysiological low salt concentrations or when using negatively supercoiled DNA templates (Shi Y et al., 2012, manuscript under review).

DNA replication

Introduction

DNA replication is a fundamental process required for all organisms. The mechanisms of DNA replication are in essence the same in all living organisms. All three domains of life (bacteria, archaea and eukarya) duplicate their DNA in a semiconservative manner, meaning that each strand of the original double-stranded DNA molecule is used as a template to produce a new complementary daughter strand. DNA replication

begins at specific sites in the genome called origins, where DNA is unwound and synthesis of the two new daughter strands forms a DNA replication fork.

The replisome

DNA replication is carried out by a multiprotein specialized enzymatic machinery, a replisome, that is assembled at the DNA replication fork. At the replication fork a helicase unwinds the original duplex DNA by breaking the hydrogen bonds that hold the two parental strands together. As single-stranded DNA (ssDNA) appears at the fork, ssDNA binding proteins rapidly associate with it, preventing the reannealing of the unwound DNA strands. In addition, ssDNA binding proteins are critical for replisome function and often stimulate both the DNA polymerase and helicase activities (Benkovic et al., 2001). The DNA polymerase synthesizes complementary copies of each strand.

Duplex DNA is arranged in an antiparallel configuration, but DNA polymerases can only synthesize DNA in a 5′ to 3′ direction. This apparent directionality problem that the replisome encounters at the replication fork is resolved by the discontinuous synthesis of one strand. The leading-strand is synthesized continuously upon unwinding of the duplex DNA by the helicase. On the other hand, the lagging-strand is synthesized discontinuously in the form of short fragments called Okazaki fragments. Each of these fragments is synthesized in a 5′ to 3′ direction but the overall synthesis of the lagging-strand will be 3′ to 5′. The lagging-strand aligns its polarity with the leading-strand through a loop formation (Trombone model), which allows both leading- and lagging-strand DNA synthesis to advance in parallel.

DNA replication needs to be primed. This is due to the fact that DNA polymerases cannot initiate DNA synthesis *de novo*, but can only elongate an existing RNA or DNA molecule. Specialized RNA polymerases called primases are the enzymes that synthesize short RNA primers needed for replication. The leading-strand needs to be primed only once whereas the lagging-strand needs to be primed continuously at the initiation of each Okazaki fragment.

Most of the current mechanistic knowledge about DNA replication comes from studies made in *Escherichia coli* (*E. coli*) and bacteriophages T7 and T4. Several components of the mtDNA replication and transcription machineries are structurally related to proteins encoded by the T-odd bacteriophage lineage (Shutt and Gray, 2006a). TWINKLE, the mitochondrial helicase, mtDNA polymerase γ (DNA POLγ) and the mitochondrial RNA polymerase, POLRMT share high sequence similarities with their T7 homologs. For this reason, The T7 DNA replication system may be used as a model that can help us understand important aspects of mtDNA replication.

T7 DNA replication

T7 DNA replication can be reconstituted *in vitro* with only four proteins: the DNA polymerase T7 gene 5 protein (gp5), its processivity factor thioredoxin (trx), the hexameric helicase/primase T7 gene 4 protein (gp4), and a single-stranded DNA binding protein, T7 gene 2.5 protein (gp2.5) (Figure 5). The processivity of the DNA polymerase gp5 increases about 100-fold in the presence of thioredoxin (Johnson and Richardson, 2003). Three different interactions with the helicase gp4 increase the processivity of gp5/trx to >17kb during leading-strand synthesis on duplex DNA. Gp5/trx is loaded onto the replication fork by an interaction with the C-terminal tail of gp4. During DNA synthesis, a high affinity complex between the two proteins is formed. A third mode of interaction allows DNA polymerases exchange during DNA synthesis (Zhang et al., 2011). The gp4 protein assembles as a hexamer on the lagging-strand by a ring-opening mechanism without the help of a loading factor (Matson and Richardson, 1983) and uses the energy of dTTP hydrolysis to translocate in a 5′ to 3′ direction during unwinding of duplex DNA (Hamdan and Richardson, 2009). The primase domain of gp4 synthesizes tetraoligoribonucleotides on the lagging-strand in a periodic manner at specific sequences. The formation of a replication loop in the lagging-strand allows for a coordinated synthesis of both strands. The replication loop is formed and released during each cycle of Okazaki fragment synthesis, which implies that the lagging-strand DNA polymerase has to recycle continually.

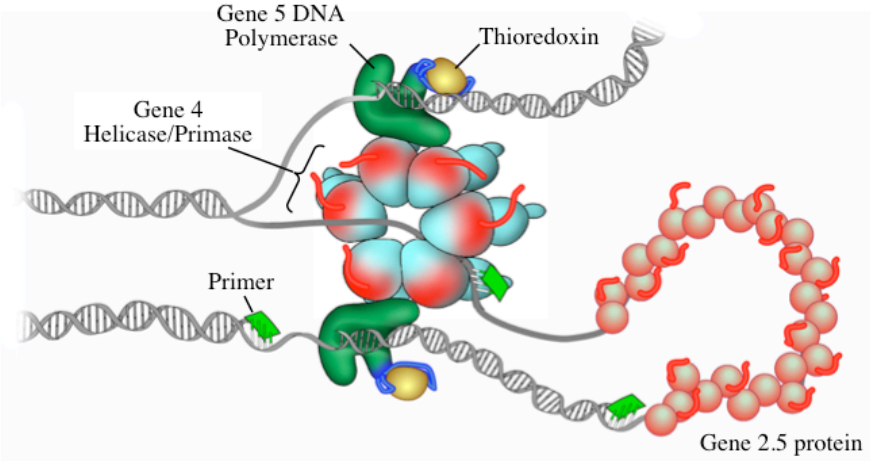


Figure 5. T7 bacteriophage replisome
Model of the T7 replisome. The helicase/primase gp4 encircles the lagging-strand and unwinds the duplex DNA. Gp5/trx bound to the helicase catalyzes the polymerization of nucleotides on the leading-strand. The primase domain of gp4 synthesizes a tetraribonucleotide to serve as a primer for the lagging-strand gp5/trx that is also bound to the helicase. A replication loop is formed in order to align the polarity of both polymerases. Gp2.5 coats the exposed ssDNA (figure courtesy of Prof. Charles C. Richardson).

Besides coating the ssDNA on the lagging-strand and preventing reannealing of the unwound DNA strands, T7 gp2.5 also plays other essential roles in T7 replication. The protein interacts with both gp5/trx and gp4 via its acidic C-terminal tail. These interactions are believed to be essential for coordinated leading-and lagging-strand synthesis. In support of this, 4-fold reduction in lagging-strand DNA synthesis is observed in a reconstituted *in vitro* system using a truncated gp2.5 that lacks its C-teminal tail but conserves the same ssDNA binding activity as the wild type protein (Lee et al., 1998).

Primases

As mentioned before, primases synthesize short RNA molecules that are used by DNA polymerases to initiate replication. Most primases can be divided into two superfamilies. The first group includes prokaryotic (bacterial and bacteriophage) primases whereas the second class includes the small eukaryotic primases that typically form a complex with DNA polymerase alpha. These two groups of primases share no sequence or structure similarities. However, some similarities can be observed in the priming mechanism. Primer synthesis requires a minimum of five steps: template binding, NTP binding, primer initiation, extension and primer "hand off" to DNA polymerase.

Prokaryotic primases are always closely associated with the helicase. Their C-terminal domain can by itself be a DNA helicase like T7 bacteriophage gp4 or directly interact with a DNA helicase like E.coli DnaG. In this way the primase is physically associated with the replisome making priming more efficient. Primases bind to ssDNA generated by the action of helicases in a sequence-specific manner. The recognition sites differ in sequence and length between primases (3´-CTGGG-5´, 3´-CTGTG-5´ or 3´-CTGGT-5´ for T7 gp4 and 5´-CTG-3´ for *E. coli* DnaG) and may play an important role in lagging-strand synthesis coordination. A flexible Zinc-binding domain (ZBD) located within the primase N-terminal domain is responsible for recognition site binding. Eukaryotic primases are less sequence-specific, but, similar to the prokaryotic primases, the sequence-specificity is controlled by a zinc-binding domain (Zenkin and Severinov, 2008).

Primer synthesis is a slow process that determines the rate of lagging-strand DNA synthesis. The rate-limiting step of primer synthesis occurs at or prior to the formation of the first phosphodiester bond. After template binding, primases bind two NTP substrates to form a dinucleotide. The first NTP incorporated at the 5´ end of the primer is always a purine (ATP or GTP). This step requires the formation of a stable complex through sequence-specific interactions between the primase and the template. Primases incorporate NTPs with low fidelity and incorporation of dNTPs has been observed *in vitro*. Fidelity and nucleotide discrimination is strongly dependent on the experimental conditions (Frick and Richardson, 2001). When the RNA product is long enough to form a stable hybrid, the initiation complex is disrupted allowing for transition into

productive synthesis. Failure to escape into productive elongation results in abortive products.

Primer length is controlled by the interactions between the catalytic domain and the DNA binding domain of the primase. As the primer elongates a tension is accumulated in the priming complex until a critical primer length is reached, which causes disruption of the elongation complex and tension release. Phage primases synthesize 4 nt primers whereas bacterial and eukaryotic primases synthesize primers of about 10 nt. The 3´ end of the primer then needs to be transferred to the DNA Polymerase. This occurs in a very direct manner in eukaryotic primases since they form a complex with Pol α, whereas in the case of prokaryotic primases single-strand DNA binding proteins are thought to facilitate primer hand-off to the DNA polymerase.

In addition to these two primase classes it has been shown that RNA polymerases (RNAP) can also prime DNA replication. Already 30 years ago, Arthur Kornberg demonstrated that *E. coli* RNAP was priming origin-dependent M13 bacteriophage DNA replication. RNAP primer synthesis usually involves ssDNA in contrast to regular transcription. The M13 origin of replication (ori) forms two imperfect double-stranded hairpins. Origin-specific primer synthesis depends on *E. coli* SSB, which binds the single-stranded phage genome and thereby excludes RNAP from it. *E. coli* SSB is unable to bind to partially double-stranded regions, which makes the origin accessible for RNAP. Primer synthesis by RNAP occurs by analogous mechanisms described for both eukaryotic and prokaryotic primases (Zenkin and Severinov, 2008).

MtDNA replication

MtDNA replication takes place throughout the cell cycle independently of nuclear DNA replication. MtDNA copy number regulation is thought to depend on the energy requirements of cells, which vary between cell types and different physiological conditions. However, the molecular mechanism behind this essential regulation is not known. All factors required for mtDNA replication are nucleus-encoded and therefore, signaling pathways between mitochondria and the nucleus are required to synchronize mitochondrial biogenesis (Garesse and Vallejo, 2001).

For three decades, the accepted model for mtDNA replication in mammals was the strand-displacement mode. According to this model, DNA synthesis takes place unidirectionally on both strands. DNA replication is asymmetric and initiates from two origins of replication, OriH and OriL. In 2000, this model was questioned and it was suggested that mtDNA replication occurred in a conventional strand-coupled mode. Moreover, a few years later, a third mode for mtDNA replication was reported, which included ribonucleotide incorporation throughout the lagging strand (RITOLS) (Yasukawa et al., 2006) (Figure 7). Currently, there is no consensus about the exact mechanism of mtDNA replication in mammals.

Strand-displacement model

The strand-displacement model for mtDNA replication was proposed based on electron microscopy studies and biochemical characterization (Clayton, 1982; Robberson et al., 1972; Tapper and Clayton, 1981). Replication of both strands is continuous and initiated from two unidirectional origins. H-strand DNA synthesis is initiated at OriH, downstream of LSP, in the D-loop region, displacing the parental H-strand. Replication then continues in one direction until OriL is reached. Upon passage of the leading-strand DNA replication machinery, OriL is exposed in its single-strand conformation, leading to origin activation, and L-strand synthesis is initiated in the opposite direction. Replication of both strands continues until the two daughter molecules are completed (Shadel and Clayton, 1997).

Heavy-strand DNA synthesis

RNA primers required for initiation of H-strand synthesis at OriH are produced by transcription events initiated at LSP. This implies that initiation of mtDNA replication is coupled and depends on mitochondrial transcription (Clayton, 1992). The mechanism that regulates the balance between productive, full-length transcription and primer synthesis remains unknown. Primer formation requires the formation of a stable RNA-DNA hybrid and termination or processing of the LSP transcript at sites near OriH. RNA-to-DNA transitions occur in a region of the D-loop that contains three conserved sequence blocks, CSB I, II and III (Walberg and Clayton, 1981).

Even though RNA-to-DNA transitions have been carefully mapped *in vivo* (Chang and Clayton, 1985), initiation of DNA synthesis at OriH has not yet been reconstituted *in vitro*. It has been suggested that RNA primers are formed by cleavage of the primary transcript by the ribonucleoprotein mitochondrial RNA processing endonuclease (RNase MRP) (Topper et al., 1992). The relevance of these results has however been questioned, since RNase MRP is present mostly in the nucleolus where it is involved in rRNA cleavage. An alternative RNase MRP-independent mechanism has been suggested based on the observation that CSBII promotes LSP-transcription termination *in vitro* (Pham et al., 2006). Recently, it has been demonstrated that formation of a G-quadruplex structure in the nascent RNA transcript is involved in CSBII-specific transcription termination (Wanrooij et al., 2010). As noted above, the newly synthesized primers form a stable and persistent RNA-DNA hybrid near OriH. Thi unusual behaviour of the RNA primer is explained by the formation of a stable G-quadruplex structure, involving the CSBII region in both the nascent RNA and the non-template DNA strand. The stable RNA-DNA hybrid (R-loop) that is formed may be used as a primer for H-strand replication. However, the formation of this complex structure between the primer and the non-template DNA strand, may also present a problem for the DNA replication machinery, since it removes the 3´-end of the RNA primer from the template strand with the consequence that the primer is not accessible (Wanrooij P. H. et al., 2012, manuscript under review). This alternative model for primer formation might therefore require additional factors to prevent the formation of this structure, or alternatively resolve it, so that the RNA primer can be used for initiation of H-strand replication.

Interestingly, when replication has initiated at OriH, only a few initiation events actually undergo complete replication (about 5% of the total initiation events) (Brown et al., 1978). The majority of the nascent H-strands are terminated at the termination-associated sequences (TAS), creating the triplex D-loop structure. TAS sequences are short and conserved in vertebrates, located a few positions upstream of the 3´ends of arrested H-strands (Doda et al., 1981). The mechanism behind what determines whether a nascent H-strand should arrest or proceed all the way around the genome remains unknown, but it is thought to be the key regulatiory step in mtDNA copy number control. A TAS-specific DNA-binding activity has been identified in cows, rat and human using *in vitro* methods and *in organello* foot printing (Madsen et al., 1993; Roberti et al., 2003), which suggests that D-loop formation is mediated through a trans-acting factor binding to TAS. However, the protein/s responsible for this activity have not been identified yet.

Origin of Light-strand DNA replication

The OriL sequence forms a stable stem-loop structure that serves as an initiation site for L-strand synthesis when exposed in its single-stranded conformation (Figure 6). OriL is a short non-coding region of ~ 30 bp flanked by $tRNA^{Asn}$ and $tRNA^{Cys}$ within the WANCY tRNA cluster. In most vertebrates the predicted secondary structure of OriL is conserved whereas the primary sequence differs significantly (Shadel and Clayton, 1997). It is worth mentioning that an OriL-like noncoding sequence is absent in the expected position in the genomes of some vertebrates, e.g. Gallus gallus. This suggests a different location of OriL in some species or that one of the tRNA genes may function as OriL (Desjardins and Morais, 1990).

Alternative initiation sites for L-strand mtDNA synthesis have been suggested by analysis of mtDNA replication intermediates using atomic force microscopy (Brown et al., 2005). However, these other initiation sites may be used at a lower frequency, explaining why only OriL was identified when free 5´ ends of newly synthesized L-strands were mapped *in vivo* (Tapper and Clayton, 1982).

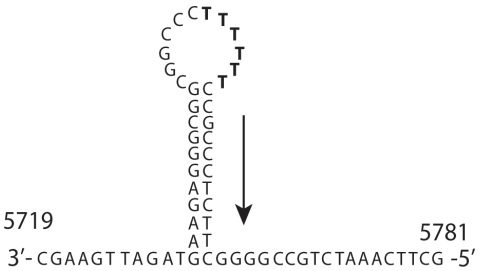


Figure 6. Origin of Light-strand DNA replication.
OriL adopts a stem-loop structure conformation when is exposed as single-strand. The poly-dT stretch where primer synthesis is initiated is highlighted in bold. The arrow inicates the direction of L-strand synthesis (Adapted from (Wong and Clayton, 1985a))

Initiation of DNA replication at OriL requires a primase activity that synthesizes short RNA primers that can be used by the DNA polymerase. Primer synthesis is initiated at the poly-dT stretch of the loop in the predicted OriL stem-loop structure and RNA-to-DNA transition occur in a region immediately 3´ to the base of the stem-loop structure (Wong and Clayton, 1985a). An OriL-specific mitochondrial DNA primase activity

was partially purified from cell extracts (Wong and Clayton, 1985b) but the corresponding protein was not identified and the precise mechanism behind OriL activation has remained obscure for 30 years. The work presented in papers I and II of this thesis identifies the mitochondrial RNA polymerase (POLRMT) as the lagging-strand primase in mammalian mitochondria and attempts to elucidate the molecular mechanism behind OriL-dependent initiation of L-strand synthesis. Furthermore, the work presented in paper III demonstrates that OriL is essential for mtDNA replication *in vivo*.

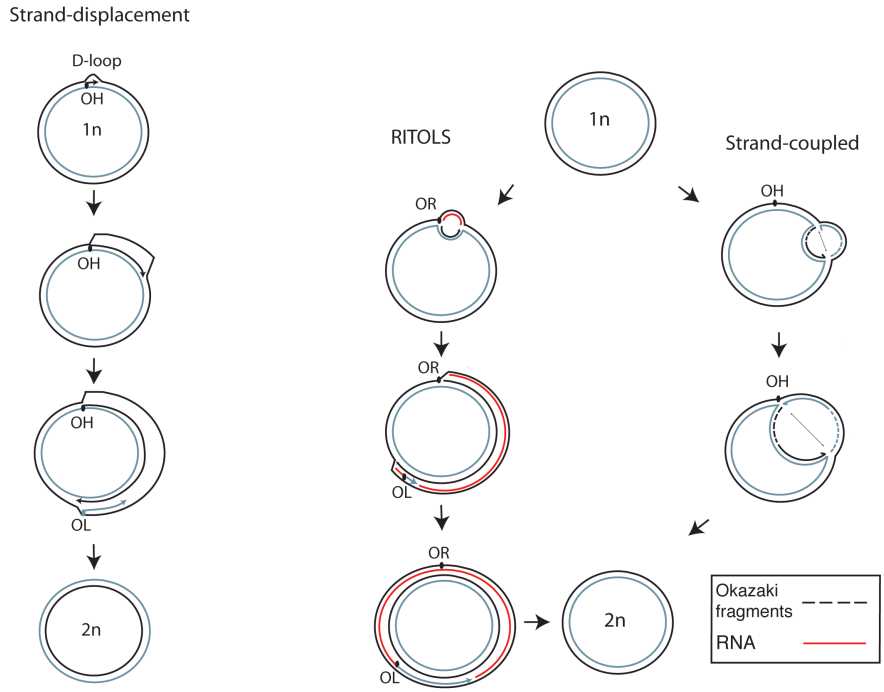


Figure 7. Models for mtDNA replication.
In the strand-displacement model, H-strand DNA replication initiates at OriH, and proceeds unidirectionally displacing the parental H-strand as ssDNA. When OriL is exposed, L-strand DNA replication is initiated in the opposite direction. RITOLS DNA replication involves ribonucleotide incorporation throughout the lagging-strand and subsequently replaced by DNA. During strand-coupled DNA replication, initiation occurs at a broad initiation zone (OriZ) in a bidirectional way. The two replication forks proceed until one of them arrests at OriH (Adapted from a figure provided by Dr.Wanrroij and Prof. H. Spelbrink)

Strand-coupled model

The strand-displacement model of mtDNA replication has been questioned by results obtained from 2D native agarose gel electrophoresis (2DAGE). This method separates DNA molecules by size and shape, and replication intermediates can be recognized according to how they run (Brewer and Fangman, 1987). Thus, restriction fragments containing conventional replication forks, that are essentially duplex DNA, produce a characteristic "Y-arc" on 2D gels. In addition, information about origins of replication and replication pause sites can be obtained with 2D gels. Analysis of mtDNA purified

from human placenta and mouse liver has suggested the existence of fully double-stranded replication intermediates, which is consistent with conventional coupled leading- and lagging-strand DNA synthesis (Holt et al., 2000). Furthermore, 2DAGE analysis of replication intermediates from both mammals and birds indicated that strand-coupled mtDNA replication occurred in a bidirectional manner initiating from a broad initiation zone (OriZ) near OriH. Moreover, OriH was suggested to be a fork arrest point (Bowmaker et al., 2003; Reyes et al., 2005). Another strong argument against the strand-displacement model is the apparent absence of OriL in the expected region of chicken mtDNA (the tRNA WANCY cluster).

The 2DNAGE analysis did also reveal different replication intermediates of single-stranded nature (sensitive to S1 nuclease) and these were assumed to derive from the strand-displacement mode of replication (Holt et al., 2000). However, these intermediates were later shown to be products of partial degradation of RNA/DNA hybrids (RNase sensitive replication intermediates). Extensive RNA tracts were found covering the entire lagging-strand. A novel mtDNA replication mode, entailing ribonucleotide incorporation throughout the lagging-strand, was then suggested, the so-called RITOLS model (Yang et al., 2002; Yasukawa et al., 2006). Like the strand-displacement model, RITOLS DNA replication is unidirectional with H- and L-strands initiating in the control region and at OriL respectively with a significant delay between leading- and lagging-strand DNA synthesis. However, RNA would be first laid down in the lagging-strand and subsequently replaced by DNA in maturation steps. The function of RITOLS was explained as a mechanism to regulate transcription and replication; RNA tracts would represent a roadblock to transcription events while replication is taking place. The precise mechanism by which the provisional lagging-strand RNA is synthesized and how it is replaced by DNA has not been addressed yet.

The mitochondrial replisome

The DNA POLγ holoenzyme, the ring-shaped hexameric helicase TWINKLE, mitochondrial single-strand DNA binding protein (mtSSB), and POLRMT (as demonstrated in this thesis) are the principal factors that constitute the mammalian mitochondrial DNA replisome (Figure 8).

The mitochondrial DNA polymerase- POLγ

DNA polymerase γ is the only DNA polymerase present in animal mitochondria and it is thus responsible for all mtDNA repair and replication (Kaguni, 2004). POLγ is an efficient and processive polymerase and displays high fidelity in nucleotide selectivity and incorporation. In mammals, the holoenzyme forms an asymmetric heterotrimer consisting of one catalytic subunit POLγA and two identical subunits of its accessory subunit POLγB (Carrodeguas et al., 2001; Lee et al., 2009; Yakubovskaya et al., 2006), whereas in *D. melanogaster* (Wernette and Kaguni, 1986) and *X. laevis* (Bogenhagen and Insdorf, 1988) POLγA forms a heterodimer with POLγB. In *S.cerevisiae*, MIP1, the yeast homolog of POLγ is present as a single catalytic subunit (Foury, 1989).

The catalytic subunit POLγA has a "right hand" configuration with palm, fingers and thumb subdomains typical of the family A DNA polymerases like *E. coli* Pol I or T7 DNA polymerases (Kaguni, 2004). In POLγA, a 3′-5′ exonuclease domain is connected to the polymerase domain by a spacer region. A third 5′-dRP lyase activity involved in base excision repair has also been reported (Longley et al., 1998). Sequence alignment of the polymerase and exonuclease domains with other family A members shows highly conserved motifs (A, B and C; I, II and III respectively) common in the family A polymerases, that are essential for polymerase and exonuclease activities. POLγA contains a unique large spacer domain (400 residues), which has been suggested to be involved in DNA binding, subunit interactions and the interplay between the polymerase and exonuclease activities (Luo and Kaguni, 2005). The crystal structure confirmed these results and revealed that the spacer region is organized in two subdomains: a globular subdomain associated with intrinsic processivity (IP) and an extended accessory-interacting determinant (AID) subdomain that contacts one of the POLγB subunits (Lee et al., 2009).

The POLγB accessory factor shares both sequence and structural homology with class IIa aminoacyl-tRNA synthetases. POLγB increases the processivity of POLγ by increasing affinity for DNA, accelerating the polymerization rate, and suppressing exonuclease activity (Carrodeguas et al., 2002; Lim et al., 1999). The two POLγB subunits bind POLγA in an asymmetric manner; one POLγB monomer makes extensive contacts with POLγA and stimulates the holoenzyme DNA-binding affinity by increasing the contact of the holoenzyme to DNA, whereas the distal POLγB monomer presents very few contacts with the catalytic subunit and stimulates the replication rate (Lee et al., 2009; Lee et al., 2010). POLγB has dsDNA-binding activity that is needed for coordination of POLγ holoenzyme and TWINKLE helicase activities at the mtDNA replication fork. It might ensure that the holoenzyme remains bound to the template behind the slow-moving TWINKLE helicase (Farge et al., 2007).

The mitochondrial DNA helicase- TWINKLE

The TWINKLE gene was identified in a search for mutations associated with autosomal dominant progressive external ophthalmoplegia (adPEO), a disorder characterized by mtDNA deletions (Li et al., 1999). The gene product co-localized with mitochondrial nucleoids and homology searches showed high sequence similarity with the T7 gp4 bacteriophage helicase/primase (Spelbrink et al., 2001). These observations immediately pointed to the possibility that TWINKLE could be the long-sought after mitochondrial replicative helicase.

TWINKLE consists of a C-terminal helicase domain and an N-terminal primase-like domain connected by a linker region. The C-terminal helicase domain of TWINKLE displays the same conserved motifs found in T7 gp4. In contrast to the gp4 protein, the N-terminal domain of TWINKLE does not possess primase activity. The human TWINKLE lacks the zinc-binding domain that is essential for primase activity in T7 gp4 as well as a couple of amino acid residues essential for nucleotide polymerization. Bioinformatic analyses strongly suggest that the primase activity of TWINKLE has

been lost in metazoa (Shutt and Gray, 2006b). The exact function of the N-terminal domain of TWINKLE is not known but it has been shown to affect ssDNA binding, helicase activity and replisome processivity (Farge et al., 2008). Similar to T7 gp4, TWINKLE forms hexamers and heptamers in solution (Farge et al., 2008; Spelbrink et al., 2001). The linker region is essential for multimerization of the protein as well as ATP binding and hydrolysis.

TWINKLE unwinds dsDNA with 5′ to 3′ directionality in an ATP-dependent manner. Reconstitution of a minimal mitochondrial replisome *in vitro* demonstrated that TWINKLE is a processive helicase at the replication fork; TWINKLE, in combination with POLγ and mtSSB can utilize a dsDNA template with an artificial replication fork, to synthesize ssDNA molecules of about 16 kb in size, which corresponds to the size of the human mitochondrial genome (Korhonen et al., 2004). *In vivo*, the hexameric helicase TWINKLE has to load onto the closed circular mtDNA genome. In other systems, a specialized helicase loader is needed to facilitate helicase loading by e.g. open up the ring-shape. However, recently it was demonstrated that TWINKLE can load on circular DNA without a specialized loader. Actually, the mitochondrial replication machinery can assemble on a closed circular DNA template and efficiently elongate a DNA primer in a manner that closely resembles initiation of mtDNA synthesis *in vivo* (Jemt et al., 2011).

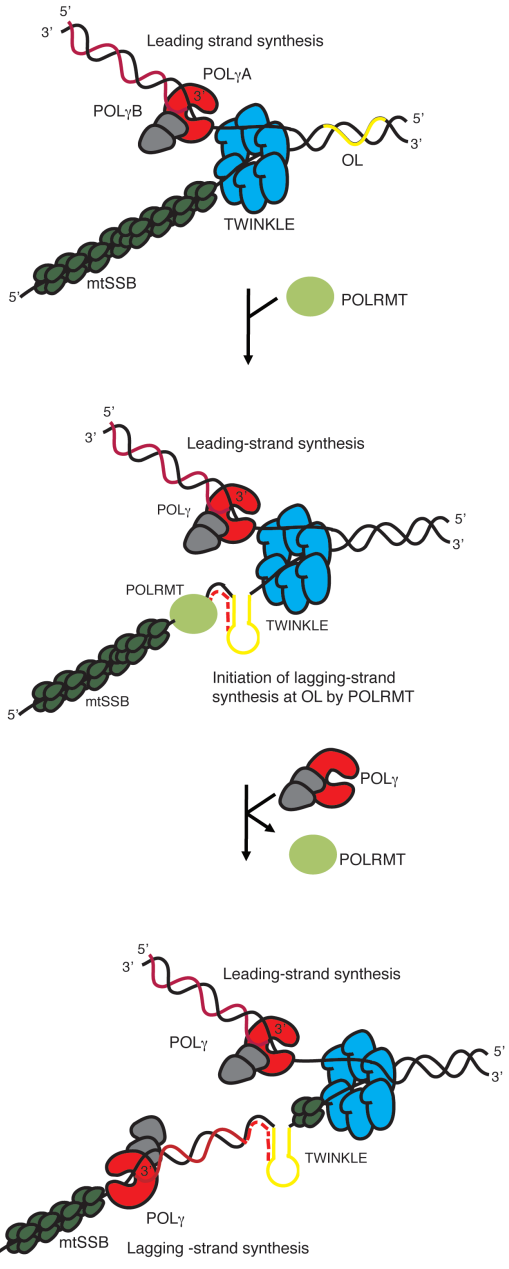
The mitochondrial single-stranded DNA binding protein- mtSSB

Single-stranded DNA binding proteins are not only required to maintain the integrity of DNA, but are also essential components in multiple processes of DNA metabolism such as DNA replication, repair and recombination. mtSSB is evolutionarily conserved throughout eukaryotes and, in contrast to POLγ, TWINKLE and POLRMT, it is not evolutionarily related with the T-odd bacteriophages. mtSSB shares high degree of similarity in aminoacid sequence, structure and biochemical properties with *E. coli* SSB (Tiranti et al., 1993).

Human mtSSB is a small protein with a predicted molecular mass of about 15 kDa. MtSSB homologues have been purified from different species and found to form tetramers in solution (Li and Williams, 1997; Thommes et al., 1995). The crystal structure of human mtSSB has been solved at 2.4 Å resolution and shows that ssDNA wraps around the tetrameric mtSSB through four electropositive channels guided by flexible loops (Yang et al., 1997). Disruption of the mtSSB gene in yeast (RIM1) and Drosophila (lopo) results in loss of mtDNA *in vivo*, a clear demonstration that mtSSB is essential for mtDNA replication (Maier et al., 2001; Van Dyck et al., 1992). In agreement with this, mtSSB knockdown in Drosophila cells results in mtDNA depletion. This phenotype is rescued upon overexpression of wt mtSSB in these cells (Farr et al., 2004).

mtSSB stimulates primer recognition and both exonuclease and polymerase activities of POLγ (Farr et al., 1999). Furthermore, mtSSB specifically enhances TWINKLE DNA unwinding activity of short stretches of dsDNA *in vitro* (Korhonen et al., 2003)

Figure 8. The mitochondrial replisome. Schematic representation of the mitochondrial replisome during origin-dependent initiation of lagging-strand DNA synthesis at OriL. The TWINKLE helicase encircles the lagging-strand and unwinds dsDNA in a 5´ to 3´ direction. The mtSSB protein stabilizes the unwound conformation and stimulates DNA synthesis by the POLγ holoenzyme. The leading-strand (H-strand) replication machinery approaches the OriL sequence and, upon passage, OriL is exposed as single-strand and adopts a stem-loop structure. MtSSB is unable to bind to the stem-loop and the poly-dT stretch in the single-stranded loop region is presented to POLRMT for initiation of primer synthesis. After about 25 nt, POLRMT is replaced by POLγ, and lagging-strand (L-strand) DNA synthesis is initiated.The newly synthesized DNA daugther strands are shown as a red solid line. (courtesy of Dr. Wanrooij and Prof. Falkenberg)



and has a strong stimulatory effect on DNA synthesis catalyzed by POLγ and TWINKLE on dsDNA templates (Korhonen et al., 2004). These data demonstrate functional interactions between mtSSB and the replisome, although little is known about how they are established. A recent study has shown that both N-terminal and C-terminal parts of mtSSB modulate in a negative manner functional interactions with POLγ. N- and C-terminal deleted mtSSB variants stimulated DNA synthesis at 1.4 to 2-fold higher level than the wild-type protein (Oliveira and Kaguni, 2010).

The mitochondrial primase- POLRMT

As noted above, POLRMT generates the primers required for initiation of DNA replication at OriH. In addition, as shown in papers I, II and II of this thesis, POLRMT also acts as a primase involved in initiation of DNA synthesis at OriL.

mtDNA genetics

Today, it is well recognized that mtDNA is maternally inherited. Several observations support this: mtDNA copy number in sperm cells is very low (50-70) compared to at least 100,000 copies in mature oocytes, therefore maternal inheritance could easily be due to a dilution effect. One single case of paternal mtDNA transmission has been described (Schwartz and Vissing, 2002). However, this case is a rare exception and it is known today that sperm mitochondria are actively degraded. After fertilization of the oocyte an autophagic process is initiated that selectively degrades the paternal mitochondria. The exact mechanisms of this process are not understood, but the paternal organelles targeted for destruction are ubiquitinated (Al Rawi et al., 2011).

Mitochondrial DNA (mtDNA) copy number is generally in the range of 100-10000 per somatic cell. This variability depends on each cell type and its energy demand. In most cases, all the mtDNA molecules within the cells of an individual are identical. We refer to this situation as homoplasmy. However, spontaneous mtDNA mutations may occur leading to a situation known as heteroplasmy, where a mixture of wild type and mutant genotypes coexist within a cell. MtDNA molecules segregate randomly to daughter cells during cell division and such distribution may generate different levels of heteroplasmy in different cells and tissues. Relatively high percentage levels of mutated molecules are tolerated before a given cell develops respiratory chain deficiency. The precise threshold for a pathologic phenotype to manifest itself varies for different mutations and for different tissues, but usually ranges from 60-90%. It has been shown that the percentage of mutated mtDNA in affected tissues correlates with the severity of the disease (Chinnery and Turnbull, 1997).

The levels of heteroplasmy often vary between the mother and each of her children. This variation is explained by a genetic “bottleneck” that occurs during early oogenesis. According to the bottleneck hypothesis, there is a random segregation or unequal amplification of different mitochondrial genomes. The exact mechanisms are still debated and this remains an outstanding question in mitochondrial biology. The

mitochondrial bottleneck has important implications for human disease, since the fraction of mutated mtDNA molecules may decide if a child is healthy or sick.

Mitochondrial diseases

The first report of a patient with a mitochondrial disorder was described in 1962 (Luft 1962). However, it was not until the late 80s when the first pathogenic mtDNA mutations were found (Holt et al., 1988; Wallace et al., 1988). Since then, more than 300 disease-causing mtDNA mutations have been reported (Mitomap). The prevalence of mitochondrial disorders has been estimated at 1 in 5000 individuals (Schaefer et al., 2004), demonstrating that these diseases are among the most common genetic disorders, in contrast to what was initially thought.

The onset of mitochondrial disorders ranges from infancy to adulthood and can affect one or multiple organs. The clinical manifestations of mitochondrial disorders are extremely heterogeneous. However, tissues and organs with a high energy demand such as the central nervous system, skeletal muscle and heart are generally affected. The molecular defect in the cells of the affected tissues is respiratory chain deficiency. Diagnosis is usually confirmed by histological and biochemical analysis of skeletal muscle biopsies where cytochrome c oxidase (COX) negative fibers can be detected. The RC deficiency can be caused either by primary mtDNA mutations or mutations in nuclear genes encoding RC subunits or proteins necessary for complex assembly, mitochondrial translation or mtDNA maintenance. As a result of the dual genetic control, mitochondrial disorders can be maternally inherited (for primary mtDNA mutations) or follow a Mendelian pattern of inheritance (nuclear encoded genes) (Ylikallio and Suomalainen, 2012).

Mitochondrial DNA mutations

Primary mtDNA mutations include mtDNA point mutations and rearrangements (large-scale deletions or duplications). Large-scale deletions have typically been associated with progressive external ophthalmoplegia (PEO), Kearns-Sayre syndrome (KSS) and Pearson's marrow-pancreas syndrome but have also been described in patients affected with mitochondrial encephalomyopathy syndromes (Schapira, 2006). Most deletions span the region between the two origins of mtDNA replication (OriH and OriL; Mitomap) and are flanked by short direct repeats. Deletions exist in heteroplasmic form and usually >60 % of mtDNA molecules present deletions before a biochemical defect is observed (Sciacco et al., 1994). The mechanism behind deletion formation is not understood but they are believed to be generated during mtDNA replication, by a strand-slippage mechanism (Shoffner et al., 1989). A new model has been proposed in which deletions are formed during repair of double-strand breaks (DSBs) (Krishnan et al., 2008). This new model is supported by the observation that induction of DSBs in mtDNA by targeting the restriction endonuclease *Pst*1 to mouse muscle mitochondria leads to deletion formation (Srivastava and Moraes, 2005).

MtDNA point mutations have been characterized in protein coding genes, tRNAs and rRNAs (Mitomap). Mutations in protein coding genes affect one specific respiratory chain complex whereas mutations in tRNAs and rRNAs cause a defect of overall mitochondrial protein synthesis. More than half of the reported mutations are located within tRNA genes even though these genes represent less than 5 % of the mitochondrial genome (Greaves et al., 2012). The clinical symptoms vary and include phenotypes like MELAS, MERRF (myoclonic epilepsy and ragged-red fibers), NARP (neuropathy ataxia and retinitis pigmentosa), MILS (maternally inherited Leigh syndrome) and LHON (Leber hereditay optic neuropathy) (Schapira, 2006).

MtDNA maintenance disorders

A large group of mtDNA disorders are caused by mutations in nuclear genes that code for proteins involved in mtDNA maintenance. Therefore, mutations in these genes result in secondary mtDNA defects i.e. mtDNA deletions, point mutations or depletion of mtDNA but are inherited in a Mendelian manner. These disorders include mitochondrial depletion syndrome (MDS), autosomal dominant (ad) or recessive (ar) progressive external ophthalmoplegia (PEO), Alpers syndrome and mitochondrial neurogastrointestinal encephalomyopathy (MNGIE). Most of the mentioned mtDNA maintenance diseases are caused by mutations in genes with a direct role in mtDNA replication (POLγA, POLγB, TWINKLE) or in mitochondrial deoxynucleoside triphosphate (dNTPs) pool maintenance. Interestingly, mutations in one single mtDNA maintenance gene can cause several different types of disorders.

Correct replication and repair of both nuclear and mitochondrial DNA requires sufficiently large dNTP pools. It is also crucial that the relative proportion of the four dNTPs is balanced to avoid mutagenesis since the replication machinery is very sensitive to dNTP pool imbalances. Two pathways regulate dNTP pools in cells, *de-novo* and the salvage pathways. In the *de-novo* pathway, ribonucleotide reductase (RNR) converts ribonucleotides derived from the purine and pyrimidine biosynthetic pathways into dNTPs. In the salvage pathway, deoxynucleosides undergo sequential phosphorylations until dNTPs are formed. Mutations in the genes that control both de-novo and salvage pathways for deoxynucleotide synthesis cause MDS, a severe disorder characterized by mtDNA depletion in the affected tissues and early onset of the disease (Moraes et al., 1991). The symptoms manifest in a tissue specific manner and can vary depending on the gene causing the disease. Mutations in thymidine kinase 2 (TK2) (Saada et al., 2001), thymidine phosphorylase (TP) (Nishino et al., 1999) and deoxyguanosine kinase (dGK) (Mandel et al., 2001) genes have been associated with MDS. Mutations in the gene encoding the p53 inducible small subunit of RNR (p53R2) also cause MDS, which suggests a crucial role for p53R2 in dNTP supply for mtDNA synthesis (Bourdon et al., 2007).

More than 150 disease-causing mutations have been identified in the POLG1 gene, which encodes for the catalytic subunit of POLγA. The spectrum of diseases caused by POLγA mutations is highly heterogeneous. The three main clinical manifestations are

PEO, Alpers syndrome and ataxia-neuropathy disorders such as MIRAS (mitochondrial recessive atxia syndrome). In adittion, POLγA mutations have also been linked to premature menopause, parkinsonism and male infertility (Chan and Copeland, 2009). PEO is a late onset disorder characterized by a gradual accumulation of multiple mtDNA deletions without a decrease in mtDNA content in postmitotic tissues. The main clinical manifestations of the disease are ptosis, external ophthalmoplegia, progressive skeletal muscle weakness and exercise intolerance, as well as cardiomyopathy and ataxia (Van Goethem et al., 2001). Mutations in the polymerase domain of POLγA are usually dominant since the DNA binding activity is unaffected and the mutant can compete with the wild type protein whereas recessive mutations usually are located in the spacer or the exonuclease domain and affect DNA binding and accessory subunit interactions (Chan and Copeland, 2009). Mutations in POLG2 gene (POLγB) causing PEO have been identified in two patients so far (Longley et al., 2006). Mutations in TWINKLE have been described in patients with adPEO (Spelbrink et al., 2001).

It has been suggested that replication stalling could be the first step towards multiple deletion formation in adPEO. In support of this, mice expressing TWINKLE with a linker duplication observed in a Finnish adPEO family, show enhanced replication stalling and accumulate mtDNA deletions (Tyynismaa et al., 2005). In addition, overexpression in human cells of adPEO-related TWINKLE and POLγ mutants induced severe replisome stalling (Wanrooij et al., 2007). Stalling could lead to double-strand breaks and, as mentioned above, repair of double-strand breaks would induce deletion formation.

THESIS AIMS

Although the main components of the mammalian mtDNA replisome are known, the exact mechanisms of mtDNA replication have not been explained in molecular detail. The aim of this thesis was to provide new insights into the molecular mechanisms of mammalian mtDNA replication, mainly focusing on the initiation of lagging-strand DNA synthesis:

PAPER I: To identify the long-sought after lagging-strand primase of mammalian mtDNA replication.

PAPER II: To characterize the molecular basis for the activation of OriL.

PAPER III: To investigate whether OriL is essential for mtDNA replication *in vivo* and to identify and characterize the sequence elements required for a functional OriL.

PAPER IV: To study the functional consequences of the dominant Y955C and recessive G848S PEO-causing POLγA mutations in the context of the replisome.

RESULTS AND DISCUSSION

Paper I. Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis *in vitro*

MtDNA replication requires an enzyme that synthesizes RNA primers on the lagging-strand. An OriL-specific primase activity had been reported in mitochondrial extracts but the responsible enzyme was never identified (Wong and Clayton, 1985b). TWINKLE, the mitochondrial replicative helicase, which shares high sequence similarity with T7 bacteriophage gp4 primase/helicase protein, is not an active primase in mammals (Shutt and Gray, 2006b). An earlier report showed that Vero cell extracts infected with herpes simplex virus type I displayed a DNA primase activity with the enzymatic characteristics of mitochondrial RNA polymerase (POLRMT) but the enzyme was never directly identified (Tsurumi and Lehman, 1990).

In our study, we demonstrate that POLRMT can act as a primase *in vitro*. POLRMT has two distinct modes of action. The enzyme is highly processive on a double-stranded DNA template, but on a single-stranded DNA template it synthesizes short RNA products of 25-75 nucleotides in size. In contrast, the structurally related T7 RNA polymerase is highly processive on both ssDNA and dsDNA templates suggesting that the low processivity on ssDNA is a specific characteristic of POLRMT. We also investigated whether POLγ could use the short RNA molecules formed by POLRMT to initiate DNA synthesis. Using radiolabeled NTPs and single-stranded DNA as a template we observed that the primers synthesized could indeed be extended in the presence of POLγ. The length of the newly synthesized DNA products decreased at high POLRMT concentrations. This was expected since POLγ will synthesize DNA until it encounters the next primer. Thus, the more primers on the template, the shorter stretches of ssDNA will be available for POLγ synthesis. We observed a severe decrease of POLRMT primer synthesis in the presence of saturating levels of mtSSB. However, even if the overall primer synthesis levels were decreased, POLγ could efficiently extend the few primers synthesized by POLRMT. In addition, mtSSB stimulated the action of POLγ and significantly longer DNA replication products were obtained. In contrast, the heterologous T7 RNA polymerase failed to support DNA synthesis in the presence of mtSSB.

Finally, we were able to reconstitute simultaneous leading- and lagging-strand DNA synthesis on a double-stranded template in the presence of the minimal mtDNA replisome (POLγ, TWINKLE and mtSSB) and POLRMT. Using the minimal replisome on a dsDNA minicircle template (Korhonen 2004) resulted in robust leading-strand DNA synthesis. After addition of POLRMT, lagging-strand DNA synthesis was observed as well, resulting in the synthesis of dsDNA. This result was confirmed by southern blotting using strand-specific probes for either the leading- or the lagging-strand. Our findings provided the first biochemical evidence supporting that POLRMT is the lagging-strand primase in mammalian mitochondria. However, according to the strand-displacement mode of mtDNA replication, lagging-strand synthesis is mainly

initiated at OriL, which presumably forms a stem-loop structure in its single-strand conformation. Therefore, the next question we had to address was whether OriL is a preferred site for POLRMT-dependent primer synthesis or not.

Paper II. Mitochondrial RNA polymerase is needed for activation of the origin of Light-strand DNA replication

We used a combination of *in vitro* biochemistry and *in vivo* approaches to demonstrate that POLRMT is the OriL-dependent primase in human cells. First we used RNA interference to knock down POLRMT in human cell lines and analyzed the effects on lagging-strand DNA synthesis using 2D native gel electrophoresis. These experiments demonstrated a significant delay of lagging-strand initiation in POLRMT-depleted cells showing that POLRMT is involved in the timing and initiation of lagging-strand DNA synthesis *in vivo.*

Next, we were able to reconstitute OriL-dependent initiation of lagging-strand replication *in vitro*, using a double-stranded template with a preformed replication fork that contained the OriL sequence. Incubating the template with the components of the mitochondrial replisome (POLγ, TWINKLE and mtSSB) and POLRMT, we obtained OriL-dependent lagging-strand DNA products that were confirmed with southern blotting using strand-specific probes. The reaction was POLRMT-dependent and, interestingly, TFAM and TFB2M, which are essential for transcription initiation at the promoters, were not required for efficient primer synthesis. Primer initiation sites and RNA-to-DNA transitions were mapped by S1 nuclease protection assays and primer extension experiments respectively and corresponded exactly with those sites mapped *in vivo* in the past (WongClayton1985, Kang1997). Primer initiation takes place at the poly-dT stretch in the loop region and RNA-to-DNA transitions take place immediately 3´ to the base of the stem-loop.

Deletion of the OriL stem-loop structure from the DNA template resulted in loss of site-specific lagging-strand synthesis. OriL mutations that destabilized the base pairing of the stem led to a severe decrease in OriL-specific lagging-strand synthesis. Moreover, replacement of the poly-dT stretch in the OriL loop with a poly-dA sequence completely abolished primer synthesis. Addition of mtSSB abolished non-specific primer synthesis on a ssDNA template without affecting OriL-specific initiation events, suggesting a role for mtSSB in OriL-dependent initiation of mtDNA replication similar to the role of *E. coli* ssDNA binding protein in M13 DNA replication. Our results clearly demonstrate that POLRMT is the OriL-specific primase and also that the overall structure of OriL and the poly-dT stretch in the loop are essential elements for primer synthesis.

Paper III. *In vivo* mutagenesis reveals that OriL is essential for mitochondrial DNA replication

The exact mechanism for mtDNA replication is still under intense debate. According to the strand-displacement model, L-strand is initiated at OriL whereas in the strand-coupled model lagging-strand initiates at multiple sites, which implies that OriL is dispensable. To elucidate the *in vivo* functional role and the sequence requirements at OriL we used an *in vivo* saturation mutagenesis approach combined with *in vitro* biochemistry and bioinformatics analyses to study the structure and sequence requirements of the OriL sequence.

We used transgenic mice homozygous for a proofreading deficient form of the mitochondrial DNA polymerase γ (POLγA^{D257A}/POLγA^{D257A}). These mtDNA mutator mice accumulate high levels of somatic mtDNA point mutations. MtDNA point mutations accumulate randomly in all coding positions, whereas the mutation load in the mitochondrial control region is significantly lower, most likely due to a reduction in the replicative capacity of these molecules (Ameur et al., 2011; Trifunovic et al., 2004). We decided to investigate whether mutations at OriL are also selected against and used PCR-sequencing to analyze this region. The transgenic mutator mouse displayed a 5-fold lower mutation load at OriL (0.2 mutations/1000bp) compared with coding genes (~1 mutation/1000bp), suggesting that mutations at OriL are rapidly selected against. Interestingly, the few point mutations found at OriL were all located at the non-template purine-rich strand side of the stem. We also found that the length of the poly-dT stretch in the loop region was highly variable and that these changes were tolerated *in vivo*. Our *in vitro* analyses confirmed that the OriL mutations found *in vivo* did not affect primer synthesis and hence only mtDNA molecules containing a functional OriL were maintained in the mouse. Our data therefore shows a distinct pattern of tolerated variability at OriL.

Additional *in vitro* analyses demonstrated that a functional OriL requires a stable stem-loop structure. A shorter, and therefore less stable double-stranded stem caused a decrease in OriL activity. This supports the idea that the functional role of the stem-loop structure is to prevent mtSSB binding into the origin region, and present the single-stranded loop region to POLRMT. In accordance with this, POLRMT uses ATP as the initiating nucleotide and could initiate primer synthesis from any poly-dT stretch on linear ssDNA templates that lacked any potential secondary structure. However, this activity was abolished in the presence of mtSSB. Our results showed that mtSSB governs OriL specificity. Rolling-circle DNA replication in the absence of mtSSB resulted in initiation of lagging-strand DNA synthesis at many different sites. Addition of mtSSB silenced all these atypical priming sites and only OriL-specific lagging-strand DNA synthesis remained. It is thus possible that, *in vivo*, mtSSB levels regulate lagging-strand initiation events at OriL and other sites in the genome and, furthermore, that low mtSSB concentrations *in vivo* creates a complicated pattern of replication intermediates that could resemble strand-coupled replication.

In our study, we used different bioinformatic approaches and demonstrated that OriL is strongly conserved throughout vertebrates and is an early invention in vertebrate evolution. In addition, the bioinformatic analysis of conserved sequence elements at OriL showed a highly variable loop region and a remarkably conserved stem. Also, we found that the pyrimidine-rich sequence at the stem presents a conserved CCCGCC motif, which was previously shown to be essential for primer synthesis. Therefore, the bioinformatic analysis supported our *in vitro* and *in vivo* characterization of OriL.

A strong argument used against the strand-displacement model for mtDNA replication is the absence of OriL in some species, especially in birds. There are studies suggesting a bidirectional mtDNA replication mechanism in chicken (Reyes). We therefore decided to use our methods for OriL prediction and analyzed the mtDNA genomes of birds. We identified an OriL-like stem-loop structure, but it was located in a different region in the genome. This hairpin structure was found in the D-loop region, about 800 bp downstream of OriH, in Galliformes (Gamefowl) and Anseriformes (Waterfowl). To test this prediction, we performed primer extension experiments from isolated mtDNA from chicken liver and observed that free 5´-ends mapped within the putative chicken OriL loop region. RNA-to-DNA transitions mapped 100-150 bp downstream of the stem-loop structure. Moreover, rapid amplification of cDNA ends (RACE) mapped the 5´-ends of the RNA primers to the loop region of the chicken OriL. The loop region contained a poly-dG that is conserved in birds, which suggests that the poly-dT stretch found in mammals has been replaced by Gs in birds. The conservation of OriL in vertebrates and the fact that a functional OriL structure is essential for mtDNA replication *in vivo* strongly supports the strand-displacement model for mtDNA replication.

Paper IV. Sequence-specific stalling of DNA polymerase γ and the effects of mutations causing progressive ophthalmoplegia

In this study, we investigated the functional consequences of the dominant Y955C and recessive G848S POLγA PEO-causing mutations *in vitro* in the context of the mtDNA replisome. Overexpression of the Y955C mutant in human cell lines resulted in severe decrease levels of mtDNA copy number and a replication-stalling phenotype verifying the dominant nature of the Y955C mutant *in vivo.*

We were able to characterize the dominant and recessive effects of Y955C and G848S *in vitro*. Addition of PolγA:G848S mutant did not affect the wild-type POLγA neither on ssDNA templates, nor on dsDNA templates in the context of the mitochondrial replisome in the presence of TWINKLE and mtSSB. Our findings were thus in agreement with the recessive nature of this mutant and reflects the *in vivo* situation found in heterozygous individuals. In contrast, the dominant POLγA:Y955C could compete with wild-type POLγA for binding to a primed template, therefore repressing wild-type POLγA-dependent DNA synthesis. However, wild-type POLγA was processive in the replisome context and POLγA:Y955C failed to compete with the wild-type protein for access to the primed template once the wild-type replisome had

initiated DNA synthesis. Two populations of long and short DNA replication products were observed in rolling-circle reactions in the presence of wild-type POLγA and POLγA:Y955C.
Biochemical analysis revealed that POLγA:Y955C exhibited low polymerase activity and a replication stalling phenotype. In addition, POLγA:Y955C displayed an imbalance between the exonuclease and polymerase activities and favored the exonuclease mode. POLγA:Y955C backtracked on DNA destroying DNA synthesis products, which could lead to stretches of single-stranded gaps on DNA. This effect could explain the deletion phenotype observed in patients since generation of single-stranded regions may facilitate strand breaks and, in turn, cause mtDNA deletions in postmitotic tissues (Krishnan et al., 2008).

As suggested in earlier studies in yeast (Baruffini et al., 2007), we demonstrated that the non-processive nature of POLγA:Y955C could be overcome by increasing dNTP concentrations, resulting in near wild-type levels of DNA synthesis. Furthermore, we observed that the requirement for higher dNTP concentrations was restricted to dATP. At low dATP concentrations POLγA:Y955C stalled just before dATP insertion sites and entered into a polymerase/exonuclease idling mode. Interestingly, we observed that the wild-type POLγA was also extra sensitive to dATP, however at much lower concentrations compared to POLγA:Y955C. Tyr955, together with Glu895 and Tyr951 forms the hydrophobic pocket that accommodates and stabilizes the incoming dNTP during DNA synthesis. The phenyl ring of Tyr955 stabilizes the template base by stacking interactions, preventing it from wobbling and thus allowing base pairing with the incoming nucleotide. The balance between exonuclease and polymerase activities is established by the stability of the template base and the newly synthesized base. Therefore, the exonuclease activity will be more pronounced upon dATP or dTTP insertions since dA-dT base pairing is less stable than dG-dC base pairing.

In conclusion, this study explains the dominant nature of POLγA:Y955C and demonstrates that mutations in the polymerase domain of human POLγA may shift the balance between exonuclease and polymerase activities leading to mitochondrial disease in affected individuals.

CONCLUDING REMARKS

The work presented in this thesis has contributed to a better understanding of the basic molecular mechanisms of mtDNA replication.

In our work, we have demonstrated that POLRMT is the mitochondrial primase required for initiation of DNA synthesis at OriL. Depletion of POLRMT *in vivo* results in delayed lagging-strand synthesis. Furthermore, we have faithfully reconstituted OriL-dependent initiation of replication *in vitro* using only POLRMT and DNA replication factors, and elucidated in molecular detail how origin specific initiation of mtDNA synthesis is achieved. We have demonstrated that mtSSB may govern origin specificity. POLRMT only needs a short poly-dT stretch to commence primer synthesis and in the absence of mtSSB, primer synthesis can be initiated at multiple sites on the mitochondrial genome. However, mtSSB efficiently blocks unspecific initiation of DNA replication and directs POLRMT to the poly-dT stretch in the OriL loop region. The stable stem-loop structure prevents mtSSB binding to OriL and presents the poly-dT region in its single-stranded conformation.

Using a saturation mutagenesis approach in the mouse in combination with *in vitro* biochemistry, has allowed us to characterize the OriL sequence *in vivo*. We have found that OriL is an essential element of the mitochondrial genome since only mtDNA molecules containing a functional OriL are maintained in the mouse. Mitochondria cannot be transfected and it is therefore not possible to investigate the *in vivo* relevance of individual DNA elements that regulate transcription, translation, replication etc. Our saturation mutagenesis approach in the mouse presents an alternative solution to this problem and allows us to analyze the importance of mtDNA sequences in the mouse. We may use this approach in the future to identify conserved sequence elements in order to address their functional significance for mtDNA replication, transcription, etc.

The fact that OriL is conserved throughout vertebrates strongly suggests that the strand-displacement model is the preferred mechanism for mtDNA replication in all vertebrates. However, even if the strand-displacement model has been demonstrated here, we cannot exclude that alternative mechanisms for mtDNA replication are used under certain conditions and further investigations are needed in order to confirm this.

A challenge we have in our lab is to reconstitute transcription coupled replication *in vitro* by combining the basal replication and transcription machinery. This seems to be achieved by a much more complicated mechanism and all our attempts to reconstitute this reaction *in vitro* have, so far, failed. While initiation of lagging-strand replication from OriL seems to occur by a rather simple, firing mechanism, initiation at OriH is likely to be a tightly regulated step.

Our study with the POLγ PEO-related mutants allowed us to reconstitute the dominant and recessive nature of these mutations *in vitro*. The biochemical characterization of

disease-related mutations appears necessary in order to understand the molecular basis of PEO. Elucidating the mechanism behind large-scale mtDNA deletions will be crucial to find an effective therapeutic treatment for PEO patients.

ACKNOWLEDGEMENTS

My PhD studies have been split between Stockholm and Göteborg and have been a great adventure in which I have learnt a lot about lab work and science. However, the best part of it has been the people I have met during all these years. This thesis wouldn't have been possible without the contribution of many people and I would specially like to thank:

My supervisor **Maria Falkenberg**, not only for giving me the chance to work in your lab but mostly for showing me that with curiosity, motivation and positive thinking (and also with a lot of hard work!) one can achieve anything in life. You have always been there when I needed to discuss my results. I really admire you as a person and as a scientist. Thanks for always believing in me.

My co-supervisor **Anna Karlsson** for being so helpful with anything I needed.

Claes Gustafsson for your support, help and positive feedback about my work, but mainly for teasing me all the time, you are almost as cool as me! I am sure you will miss me a little bit when I leave the lab. Now seriously, thank you for always having time for me to discuss things and chat about everything.

Special thanks to **Sjoerd Wanrooij**, my project colleague, but most important my friend. It has been a lot of fun working together and I have learnt a lot from you all these years about the mitochondrial field and the science world, in general. It has been a lot of struggling but I think we are the best team ever, almost better than FCBarcelona! Thank you so much for your constant support and for reviewing my thesis. For me you are number one (not five)!

Moving to Sweden was a big change in my life and I have to say it was tough in the beginning. I moved from sunny and warm Barcelona to Stockholm in November 2006. It was dark and cold, plus I didn't know anyone, so I was homesick all the time. However, time passed by and I started making friends and getting to know my colleagues at the lab, and in a few months Stockholm became the best place on Earth! I would like to thank all people at the division of Metabolic Diseases, for creating such a nice environment, for super nice mitochondrial retreats and pubs, those were nice days. Special thanks to:

Mina Pellegrini for taking care of me when I first arrived to the lab in Stockholm. I remember I was constantly coming with questions and you always helped me. My italian sister, I really miss all the good times we spent together. The lab is definitely not the same without you.
Jordi Asin Cayuela for your help and advice. Poder parlar català amb algú al laboratori em va fer sentir una mica com si estigués a casa.

Géraldine Farge for teaching me the basics of lab work and how to be ironic (in a good way).
Neli Atanassova for your always nice words, it's been fun working with you.

Chan Bae Park, **Mügen Terzioglu, Metodi Metodiev, Benedetta Ruzzenente, Yolanda Cámara** for showing me what hard work means and for so much fun moments! Roooooll the circleeeeeeeeee!!!!!!!(Thank you all the body!)
Emmanouella Chatzidakis for being so positive and always care about people.
Jim Stewart for the nice times spent at the mitochondrial retreats (beers included) and your help and contribution in our last manuscript.
Camilla Koolmeister for arranging everything for the dissertation.

My spaniard-swedish friends as well: **Patxi, Laura, Juanma, Monica, Maite, Pipi, Jonatan, Fernando, Juan Carlos, Angel Cedazo** and **Angel Valero.** Déu n'hi do!!!!muy heavy!

Moving to Göteborg was not that bad in the end, and now it is impossible for me to say which city is better, thanks to:
Stefan Bäckström my first office mate and one of my first friends in Göteborg, for showing me that Göteborg is also a nice city, for your support and help, du är min favoriten gamla smurff! **Jessica Tångefjord** for you endless kindness and the best 80's hits.
All the lab members at medkem (you ar so many!) specially:
Paulina Wanrooij for all these years working together (in Stockholm as well), we have had parallel PhDs and shared not only frustrations but many great moments. Thanks for showing me how to keep fighting for it!
Ka-Wei Tang for your encouragement, humor, narcissism and our conversations about future and more (and for organizing parties at my professor office). So many great lunches together with Sjoerd at Mario's!
Emily Hoberg for bringing order to the lab, teaching me some skånska and being always up to help people. Sharing office with you it's been herrrrregud awesome!
Yonghong Shi for all the discussions about our experiments and your help. We are so close to being famous!
Koviljca du är bästa på labbet!tack så jätte mycket för allt!
Jonas Carlsten for contributing to a nice environment and bringing so good ideas to improve the functioning of the lab; **Jennifer Uhler**, without your help there would be many spelling mistakes here! **Elisabeth Jemt** for very interesting chats. I will admit it…you are NOT boring! **Bertil Macao** for your encouragement when writing my thesis; **Majda Mehmedovic** for having so much energy and for always letting people talk, it's not your fault you support the worse team in the world; The yeast gang **Xuefeng Zhu, Gabor Banyai** and **Zsolt Szilagyi.**
Per Elias for nice discussions about football and for sharing your wisdom and your phylosophy about the scientific world with me, and your help these last days.
Isabella Muylaert for being always so nice, but let's not talk about Messi.

Tore Samuelsson for your bioinformatic analyses of OriL, the chicken hunting and your always calm temper.
Carlos Cabrera por los ánimos, consejos, charlas y sobretodo por motivarme con la escalada. Visca Catalunya!
Mario, has hecho que nuestra hora de comer sea puro disfrute y alegría. Esta tesis también ha sido posible gracias a los manjares que preparas en Lyktan.

Huge thank to family and friends:
All my friends in Stockholm and Göteborg (you are too many to mention you all!), specially **Ola** and **Luis**, bästa kompisar! my life here in Sweden deffinitely wouldn't have been the same without you. "En gångs grill" och "de la verga (o era chingón?)"!!!!
Climbing has been an important part of my life in Sweden: **Arvid, Danny, Akil, Fredrik** and **Mirko**, kom igen!
Thanks **Sara** for repeating over and over that everything would go fine, I really needed it!
Thanks also to all my friends back in Barcelona (you are too many to mention you all as well!), with a special mention to **Aitor** and **Marc**, for always giving a call once in a while and being there.

Thanks to my parents **Jaime** and **M. Eugenia**, my twin brother **Luis** and my sisters **Clara** and **Ana**, and my brother-in-law **Xavi**, vuestro apoyo ha sido inmenso. El més dur de tot això ha estat ser lluny de vosaltres tot aquest temps. Us estimo molt!
The **Miralles** and **Fusté** family members, specially my cousins **Nacho**, **Borja**, and **Jaime**, hopefully we will go for a skål soon!

Finally I would like to thank **Sylvia**, my fairytale. You have given me the strength continue until the end. I am so glad I have met you. You are Zo Zekzi!!!

REFERENCES

Adan, C., Matsushima, Y., Hernandez-Sierra, R., Marco-Ferreres, R., Fernandez-Moreno, M.A., Gonzalez-Vioque, E., Calleja, M., Aragon, J.J., Kaguni, L.S., and Garesse, R. (2008). Mitochondrial transcription factor B2 is essential for metabolic function in Drosophila melanogaster development. J Biol Chem *283*, 12333-12342.

Al Rawi, S., Louvet-Vallee, S., Djeddi, A., Sachse, M., Culetto, E., Hajjar, C., Boyd, L., Legouis, R., and Galy, V. (2011). Postfertilization autophagy of sperm organelles prevents paternal mitochondrial DNA transmission. Science *334*, 1144-1147.

Allen, J.F. (1993). Control of gene expression by redox potential and the requirement for chloroplast and mitochondrial genomes. J Theor Biol *165*, 609-631.

Allen, J.F. (2003). Why chloroplasts and mitochondria contain genomes. Comp Funct Genomics *4*, 31-36.

Ameur, A., Stewart, J.B., Freyer, C., Hagstrom, E., Ingman, M., Larsson, N.G., and Gyllensten, U. (2011). Ultra-deep sequencing of mouse mitochondrial DNA: mutational patterns and their origins. PLoS Genet *7*, e1002028.

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

Baruffini, E., Ferrero, I., and Foury, F. (2007). Mitochondrial DNA defects in Saccharomyces cerevisiae caused by functional interactions between DNA polymerase gamma mutations associated with disease in human. Biochim Biophys Acta *1772*, 1225-1235.

Benkovic, S.J., Valentine, A.M., and Salinas, F. (2001). Replisome-mediated DNA replication. Annu Rev Biochem *70*, 181-208.

Bleazard, W., McCaffery, J.M., King, E.J., Bale, S., Mozdy, A., Tieu, Q., Nunnari, J., and Shaw, J.M. (1999). The dynamin-related GTPase Dnm1 regulates mitochondrial fission in yeast. Nat Cell Biol *1*, 298-304.

Bogenhagen, D.F., and Insdorf, N.F. (1988). Purification of Xenopus laevis mitochondrial RNA polymerase and identification of a dissociable factor required for specific transcription. Mol Cell Biol *8*, 2910-2916.

Bourdon, A., Minai, L., Serre, V., Jais, J.P., Sarzi, E., Aubert, S., Chretien, D., de Lonlay, P., Paquis-Flucklinger, V., Arakawa, H., *et al.* (2007). Mutation of RRM2B, encoding p53-controlled ribonucleotide reductase (p53R2), causes severe mitochondrial DNA depletion. Nat Genet *39*, 776-780.

Bowmaker, M., Yang, M.Y., Yasukawa, T., Reyes, A., Jacobs, H.T., Huberman, J.A., and Holt, I.J. (2003). Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone. J Biol Chem *278*, 50961-50969.

Brewer, B.J., and Fangman, W.L. (1987). The localization of replication origins on ARS plasmids in S. cerevisiae. Cell *51*, 463-471.

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

Brown, W.M., Shine, J., and Goodman, H.M. (1978). Human mitochondrial DNA: analysis of 7S DNA from the origin of replication. Proc Natl Acad Sci U S A *75*, 735-739.

Carrodeguas, J.A., Pinz, K.G., and Bogenhagen, D.F. (2002). DNA binding properties of human pol gammaB. J Biol Chem *277*, 50008-50014.

Carrodeguas, J.A., Theis, K., Bogenhagen, D.F., and Kisker, C. (2001). Crystal structure and deletion analysis show that the accessory subunit of mammalian DNA polymerase gamma, Pol gamma B, functions as a homodimer. Mol Cell *7*, 43-54.

Chan, S.S., and Copeland, W.C. (2009). DNA polymerase gamma and mitochondrial disease: understanding the consequence of POLG mutations. Biochim Biophys Acta *1787*, 312-319.

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

Chen, H., and Chan, D.C. (2005). Emerging functions of mammalian mitochondrial fusion and fission. Hum Mol Genet *14 Spec No. 2*, R283-289.

Chen, H., and Chan, D.C. (2010). Physiological functions of mitochondrial fusion. Ann N Y Acad Sci *1201*, 21-25.

Chen, H., Detmer, S.A., Ewald, A.J., Griffin, E.E., Fraser, S.E., and Chan, D.C. (2003). Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. J Cell Biol *160*, 189-200.

Chen, Y.C., Taylor, E.B., Dephoure, N., Heo, J.M., Tonhato, A., Papandreou, I., Nath, N., Denko, N.C., Gygi, S.P., and Rutter, J. (2012). Identification of a protein mediating respiratory supercomplex stability. Cell Metab *15*, 348-360.

Chinnery, P.F., and Turnbull, D.M. (1997). Clinical features, investigation, and management of patients with defects of mitochondrial DNA. J Neurol Neurosurg Psychiatry *63*, 559-563.

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

Clayton, D.A. (1992). Transcription and replication of animal mitochondrial DNAs. Int Rev Cytol *141*, 217-232.

Desjardins, P., and Morais, R. (1990). Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. J Mol Biol *212*, 599-634.

Doda, J.N., Wright, C.T., and Clayton, D.A. (1981). Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc Natl Acad Sci U S A *78*, 6116-6120.

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

Farge, G., Holmlund, T., Khvorostova, J., Rofougaran, R., Hofer, A., and Falkenberg, M. (2008). The N-terminal domain of TWINKLE contributes to single-stranded DNA binding and DNA helicase activities. Nucleic Acids Res *36*, 393-403.

Farge, G., Pham, X.H., Holmlund, T., Khorostov, I., and Falkenberg, M. (2007). The accessory subunit B of DNA polymerase gamma is required for mitochondrial replisome function. Nucleic Acids Res *35*, 902-911.

Farr, C.L., Matsushima, Y., Lagina, A.T., 3rd, Luo, N., and Kaguni, L.S. (2004). Physiological and biochemical defects in functional interactions of mitochondrial DNA polymerase and DNA-binding mutants of single-stranded DNA-binding protein. J Biol Chem *279*, 17047-17053.

Farr, C.L., Wang, Y., and Kaguni, L.S. (1999). Functional interactions of mitochondrial DNA polymerase and single-stranded DNA-binding protein. Template-primer DNA binding and initiation and elongation of DNA strand synthesis. J Biol Chem *274*, 14779-14785.

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

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

Fisher, R.P., Parisi, M.A., and Clayton, D.A. (1989). Flexible recognition of rapidly evolving promoter sequences by mitochondrial transcription factor 1. Genes Dev *3*, 2202-2217.

Foury, F. (1989). Cloning and sequencing of the nuclear gene MIP1 encoding the catalytic subunit of the yeast mitochondrial DNA polymerase. J Biol Chem *264*, 20552-20560.

Frick, D.N., and Richardson, C.C. (2001). DNA primases. Annu Rev Biochem *70*, 39-80.

Garesse, R., and Vallejo, C.G. (2001). Animal mitochondrial biogenesis and function: a regulatory cross-talk between two genomes. Gene *263*, 1-16.

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

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

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

Greenleaf, A.L., Kelly, J.L., and Lehman, I.R. (1986). Yeast RPO41 gene product is required for transcription and maintenance of the mitochondrial genome. Proc Natl Acad Sci U S A *83*, 3391-3394.

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

Holt, I.J., Harding, A.E., and Morgan-Hughes, J.A. (1988). Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. Nature *331*, 717-719.

Holt, I.J., Lorimer, H.E., and Jacobs, H.T. (2000). Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. Cell *100*, 515-524.

Jemt, E., Farge, G., Backstrom, S., Holmlund, T., Gustafsson, C.M., and Falkenberg, M. (2011). The mitochondrial DNA helicase TWINKLE can assemble on a closed circular template and support initiation of DNA synthesis. Nucleic Acids Res *39*, 9238-9249.

Johnson, D.E., and Richardson, C.C. (2003). A covalent linkage between the gene 5 DNA polymerase of bacteriophage T7 and Escherichia coli thioredoxin, the processivity factor: fate of thioredoxin during DNA synthesis. J Biol Chem *278*, 23762-23772.

Kaguni, L.S. (2004). DNA polymerase gamma, the mitochondrial replicase. Annu Rev Biochem *73*, 293-320.

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

Korhonen, J.A., Gaspari, M., and 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*, 48627-48632.

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

Krishnan, K.J., Reeve, A.K., Samuels, D.C., Chinnery, P.F., Blackwood, J.K., Taylor, R.W., Wanrooij, S., Spelbrink, J.N., Lightowlers, R.N., and Turnbull, D.M. (2008). What causes mitochondrial DNA deletions in human cells? Nat Genet *40*, 275-279.

Kucej, M., and Butow, R.A. (2007). Evolutionary tinkering with mitochondrial nucleoids. Trends Cell Biol *17*, 586-592.

Kukat, C., Wurm, C.A., Spahr, H., Falkenberg, M., Larsson, N.G., and 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*, 13534-13539.

Lane, N., and Martin, W. (2010). The energetics of genome complexity. Nature *467*, 929-934.

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

Lee, J., Chastain, P.D., 2nd, Kusakabe, T., Griffith, J.D., and Richardson, C.C. (1998). Coordinated leading and lagging strand DNA synthesis on a minicircular template. Mol Cell *1*, 1001-1010.

Lee, Y.S., Kennedy, W.D., and Yin, Y.W. (2009). Structural insight into processive human mitochondrial DNA synthesis and disease-related polymerase mutations. Cell *139*, 312-324.

Lee, Y.S., Lee, S., Demeler, B., Molineux, I.J., Johnson, K.A., and 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*, 1490-1499.

Li, F.Y., Tariq, M., Croxen, R., Morten, K., Squier, W., Newsom-Davis, J., Beeson, D., and Larsson, C. (1999). Mapping of autosomal dominant progressive external ophthalmoplegia to a 7-cM critical region on 10q24. Neurology *53*, 1265-1271.

Li, K., and 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*, 8686-8694.

Lim, S.E., Longley, M.J., and Copeland, W.C. (1999). The mitochondrial p55 accessory subunit of human DNA polymerase gamma enhances DNA binding, promotes processive DNA synthesis, and confers N-ethylmaleimide resistance. J Biol Chem *274*, 38197-38203.

Litonin, D., Sologub, M., Shi, Y., Savkina, M., Anikin, M., Falkenberg, M., Gustafsson, C.M., and Temiakov, D. (2010). Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro. J Biol Chem *285*, 18129-18133.

Lodeiro, M.F., Uchida, A., Bestwick, M., Moustafa, I.M., Arnold, J.J., Shadel, G.S., and Cameron, C.E. (2012). Transcription from the second heavy-strand promoter of human mtDNA is repressed by transcription factor A in vitro. Proc Natl Acad Sci U S A.

Longley, M.J., Clark, S., Yu Wai Man, C., Hudson, G., Durham, S.E., Taylor, R.W., Nightingale, S., Turnbull, D.M., Copeland, W.C., and Chinnery, P.F. (2006). Mutant POLG2 disrupts DNA polymerase gamma subunits and causes progressive external ophthalmoplegia. Am J Hum Genet *78*, 1026-1034.

Longley, M.J., Ropp, P.A., Lim, S.E., and Copeland, W.C. (1998). Characterization of the native and recombinant catalytic subunit of human DNA polymerase gamma: identification of residues critical for exonuclease activity and dideoxynucleotide sensitivity. Biochemistry *37*, 10529-10539.

Luo, N., and Kaguni, L.S. (2005). Mutations in the spacer region of Drosophila mitochondrial DNA polymerase affect DNA binding, processivity, and the balance between Pol and Exo function. J Biol Chem *280*, 2491-2497.

Madsen, C.S., Ghivizzani, S.C., and Hauswirth, W.W. (1993). Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. Mol Cell Biol *13*, 2162-2171.

Maier, D., Farr, C.L., Poeck, B., Alahari, A., Vogel, M., Fischer, S., Kaguni, L.S., and Schneuwly, S. (2001). Mitochondrial single-stranded DNA-binding protein is required for mitochondrial DNA replication and development in Drosophila melanogaster. Mol Biol Cell *12*, 821-830.

Mandel, H., Szargel, R., Labay, V., Elpeleg, O., Saada, A., Shalata, A., Anbinder, Y., Berkowitz, D., Hartman, C., Barak, M., *et al.* (2001). The deoxyguanosine kinase gene is mutated in individuals with depleted hepatocerebral mitochondrial DNA. Nat Genet *29*, 337-341.

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

Matson, S.W., and Richardson, C.C. (1983). DNA-dependent nucleoside 5'-triphosphatase activity of the gene 4 protein of bacteriophage T7. J Biol Chem *258*, 14009-14016.

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

Moraes, C.T., Shanske, S., Tritschler, H.J., Aprille, J.R., Andreetta, F., Bonilla, E., Schon, E.A., and DiMauro, S. (1991). mtDNA depletion with variable tissue expression: a novel genetic abnormality in mitochondrial diseases. Am J Hum Genet *48*, 492-501.

Nass, M.M., and Nass, S. (1963). Intramitochondrial Fibers with DNA Characteristics. I. Fixation and Electron Staining Reactions. J Cell Biol *19*, 593-611.

Nishino, I., Spinazzola, A., and Hirano, M. (1999). Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. Science *283*, 689-692.

Ojala, D., Montoya, J., and Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. Nature *290*, 470-474.

Oliveira, M.T., and Kaguni, L.S. (2010). Functional roles of the N- and C-terminal regions of the human mitochondrial single-stranded DNA-binding protein. PLoS One *5*, e15379.

Pham, X.H., Farge, G., Shi, Y., Gaspari, M., Gustafsson, C.M., and Falkenberg, M. (2006). Conserved sequence box II directs transcription termination and primer formation in mitochondria. J Biol Chem *281*, 24647-24652.

Reyes, A., Yang, M.Y., Bowmaker, M., and Holt, I.J. (2005). Bidirectional replication initiates at sites throughout the mitochondrial genome of birds. J Biol Chem *280*, 3242-3250.

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

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

Roberti, M., Polosa, P.L., Bruni, F., Musicco, C., Gadaleta, M.N., and Cantatore, P. (2003). DmTTF, a novel mitochondrial transcription termination factor that recognises two sequences of Drosophila melanogaster mitochondrial DNA. Nucleic Acids Res *31*, 1597-1604.

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

Saada, A., Shaag, A., Mandel, H., Nevo, Y., Eriksson, S., and Elpeleg, O. (2001). Mutant mitochondrial thymidine kinase in mitochondrial DNA depletion myopathy. Nat Genet *29*, 342-344.

Saraste, M. (1999). Oxidative phosphorylation at the fin de siecle. Science *283*, 1488-1493.

Sato, A., Nakada, K., and Hayashi, J. (2006). Mitochondrial dynamics and aging: Mitochondrial interaction preventing individuals from expression of respiratory deficiency caused by mutant mtDNA. Biochim Biophys Acta *1763*, 473-481.

Schaefer, A.M., Taylor, R.W., Turnbull, D.M., and Chinnery, P.F. (2004). The epidemiology of mitochondrial disorders--past, present and future. Biochim Biophys Acta *1659*, 115-120.

Schapira, A.H. (2006). Mitochondrial disease. Lancet *368*, 70-82.

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

Schwartz, M., and Vissing, J. (2002). Paternal inheritance of mitochondrial DNA. N Engl J Med *347*, 576-580.

Sciacco, M., Bonilla, E., Schon, E.A., DiMauro, S., and Moraes, C.T. (1994). Distribution of wild-type and common deletion forms of mtDNA in normal and respiration-deficient muscle fibers from patients with mitochondrial myopathy. Hum Mol Genet *3*, 13-19.

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

Shoffner, J.M., Lott, M.T., Voljavec, A.S., Soueidan, S.A., Costigan, D.A., and Wallace, D.C. (1989). Spontaneous Kearns-Sayre/chronic external ophthalmoplegia plus syndrome associated with a mitochondrial DNA deletion: a slip-replication model and metabolic therapy. Proc Natl Acad Sci U S A *86*, 7952-7956.

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

Shutt, T.E., and Gray, M.W. (2006b). Twinkle, the mitochondrial replicative DNA helicase, is widespread in the eukaryotic radiation and may also be the mitochondrial DNA primase in most eukaryotes. J Mol Evol *62*, 588-599.

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

Skulachev, V.P. (2001). Mitochondrial filaments and clusters as intracellular power-transmitting cables. Trends Biochem Sci *26*, 23-29.

Sologub, M., Litonin, D., Anikin, M., Mustaev, A., and Temiakov, D. (2009). TFB2 is a transient component of the catalytic site of the human mitochondrial RNA polymerase. Cell *139*, 934-944.

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

Srivastava, S., and Moraes, C.T. (2005). Double-strand breaks of mouse muscle mtDNA promote large deletions similar to multiple mtDNA deletions in humans. Hum Mol Genet *14*, 893-902.

Strogolova, V., Furness, A., Robb-McGrath, M., Garlich, J., and Stuart, R.A. (2012). Rcf1 and Rcf2, Members of the Hypoxia-Induced Gene 1 Protein Family, Are Critical Components of the Mitochondrial Cytochrome bc1-Cytochrome c Oxidase Supercomplex. Mol Cell Biol *32*, 1363-1373.

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

Tapper, D.P., and Clayton, D.A. (1982). Precise nucleotide location of the 5' ends of RNA-primed nascent light strands of mouse mitochondrial DNA. J Mol Biol *162*, 1-16.

Thommes, P., Farr, C.L., Marton, R.F., Kaguni, L.S., and Cotterill, S. (1995). Mitochondrial single-stranded DNA-binding protein from Drosophila embryos. Physical and biochemical characterization. J Biol Chem *270*, 21137-21143.

Tiranti, V., Rocchi, M., DiDonato, S., and Zeviani, M. (1993). Cloning of human and rat cDNAs encoding the mitochondrial single-stranded DNA-binding protein (SSB). Gene *126*, 219-225.

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

Topper, J.N., Bennett, J.L., and Clayton, D.A. (1992). A role for RNAase MRP in mitochondrial RNA processing. Cell *70*, 16-20.

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

Tsurumi, T., and Lehman, I.R. (1990). Release of RNA polymerase from vero cell mitochondria after herpes simplex virus type 1 infection. J Virol *64*, 450-452.

Tyynismaa, H., Mjosund, K.P., Wanrooij, S., Lappalainen, I., Ylikallio, E., Jalanko, A., Spelbrink, J.N., Paetau, A., and Suomalainen, A. (2005). Mutant mitochondrial helicase Twinkle causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. Proc Natl Acad Sci U S A *102*, 17687-17692.

Van Dyck, E., Foury, F., Stillman, B., and Brill, S.J. (1992). A single-stranded DNA binding protein required for mitochondrial DNA replication in S. cerevisiae is homologous to E. coli SSB. EMBO J *11*, 3421-3430.

Van Goethem, G., Dermaut, B., Lofgren, A., Martin, J.J., and Van Broeckhoven, C. (2001). Mutation of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. Nat Genet *28*, 211-212.

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

Walberg, M.W., and Clayton, D.A. (1981). Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Res *9*, 5411-5421.

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

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

Wanrooij, P.H., Uhler, J.P., Simonsson, T., Falkenberg, M., and Gustafsson, C.M. (2010). G-quadruplex structures in RNA stimulate mitochondrial transcription termination and primer formation. Proc Natl Acad Sci U S A *107*, 16072-16077.

Wanrooij, S., and Falkenberg, M. (2010). The human mitochondrial replication fork in health and disease. Biochim Biophys Acta *1797*, 1378-1388.

Wanrooij, S., Goffart, S., Pohjoismaki, J.L., Yasukawa, T., and Spelbrink, J.N. (2007). Expression of catalytic mutants of the mtDNA helicase Twinkle and polymerase POLG causes distinct replication stalling phenotypes. Nucleic Acids Res *35*, 3238-3251.

Wernette, C.M., and Kaguni, L.S. (1986). A mitochondrial DNA polymerase from embryos of Drosophila melanogaster. Purification, subunit structure, and partial characterization. J Biol Chem *261*, 14764-14770.

Williams, R.S. (1986). Mitochondrial gene expression in mammalian striated muscle. Evidence that variation in gene dosage is the major regulatory event. J Biol Chem *261*, 12390-12394.

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

Wong, T.W., and Clayton, D.A. (1985b). Isolation and characterization of a DNA primase from human mitochondria. J Biol Chem *260*, 11530-11535.

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

Yang, C., Curth, U., Urbanke, C., and Kang, C. (1997). Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 A resolution. Nat Struct Biol *4*, 153-157.

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

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

Ylikallio, E., and Suomalainen, A. (2012). Mechanisms of mitochondrial diseases. Ann Med *44*, 41-59.

Zenkin, N., and Severinov, K. (2008). RNA polymerase -the third class of primases. Cell Mol Life Sci *65*, 2280-2288.

Zhang, H., Lee, S.J., and Richardson, C.C. (2011). Essential protein interactions within the replisome regulate DNA replication. Cell Cycle *10*, 3413-3414.

Zollo, O., Tiranti, V., and Sondheimer, N. (2012). Transcriptional requirements of the distal heavy-strand promoter of mtDNA. Proc Natl Acad Sci U S A.