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

# MOLECULAR MECHANISMS OF MITOCHONDRIAL DNA REPLICATION

Teresa Holmlund


Karolinska Institutet

Stockholm 2009

Published by Karolinska Institutet. Printed by Larserics Digital Print AB.

ISBN 978-91-7409-541-8

# ABSTRACT

Mitochondria are the energy producing organelles of eukaryotic cells. The organelle has its own genome, the mitochondrial DNA (mtDNA) that encodes 13 subunits of the respiratory chain (RC) complexes, two rRNAs and 22 tRNAs. Nuclear genes encode the majority of the RC subunits and all the factors required for transcription and replication of the mtDNA. Mutations in mtDNA replication factors are associated with human diseases affecting mitochondrial genome stability and maintenance. The human mtDNA replication system has been reconstituted *in vitro* and involves the combined actions of the DNA polymerase γ holoenzyme (POLγ), the TWINKLE helicase and the single-stranded DNA binding protein mtSSB. The general aim of this thesis has been to further investigate the molecular mechanisms of mtDNA replication, with a major focus on the mitochondrial hexameric helicase TWINKLE, as well as the accessory B subunit of POLγ.

A biochemical characterization of POLγB demonstrated that the protein blocks the exonuclease activity of the catalytic POLγA subunit. In addition, the dsDNA-binding activity of POLγB was required for the TWINKLE-dependant stimulation of the POLγ holoenzyme.

TWINKLE displays sequence similarity to the bacteriophage T7 gene 4 protein (gp4) which contains the DNA helicase and primase activities needed at the bacteriophage replication fork. The C-terminal domain of TWINKLE is indeed an active helicase, but there have been no reports of primase activity. The functional role of the TWINKLE N-terminus was therefore investigated in this work. The N-terminal domain was found to contribute to ssDNA-binding and helicase activities of TWINKLE, and was ultimately required for full replisome activity. A structural model of TWINKLE was constructed based on homology modeling with T7 gp4. This model displayed a conserved region with significant electropositive potential, which in structurally related primases has been suggested to interact with ssDNA.

Mutations in both POLγ and TWINKLE can cause autosomal dominant progressive external ophtalmoplegia (adPEO). To investigate the molecular mechanisms behind this disorder, we performed a detailed biochemical analysis on eleven different adPEO-causing TWINKLE mutations, seven in the linker-region and four in the N-terminal domain. Distinct molecular phenotypes were observed, with individual consequences for multimerization, ATPase activity, helicase activity and ability to support mtDNA synthesis *in vitro*. The different molecular phenotypes could be interpreted using our structural model of TWINKLE. Two of the mutations in the linker region affected multimerization, whereas the N-terminal mutations showed a striking reduction in ATPase activity and were thus proposed to impair the interplay between ssDNA-binding and ATP hydrolysis, an essential element of the catalytic cycle of related hexameric helicases.



**Keywords:** Mitochondria, mtDNA, replication, helicase, TWINKLE, POLγ, POLγB, mtSSB, adPEO.

# LIST OF PUBLICATIONS

This thesis is based on the following papers, which will be referred to in the text by their roman numerals:

**I. The accessory subunit B of DNA polymerase gamma is required for mitochondrial replisome function.**

Farge G, Pham XH, Holmlund T, Khorostov I, Falkenberg M.

*Nucleic Acids Res*. 2007;35(3):902-11.

**II. Structure-function defects of the TWINKLE linker region in progressive external ophthalmoplegia.**

Korhonen JA, Pande V, Holmlund T, Farge G, Pham XH, Nilsson L, Falkenberg M.

*J Mol Biol.* 2008 Mar 28;377(3):691-705.

**III. The N-terminal domain of TWINKLE contributes to single-stranded DNA binding and DNA helicase activities.**

Farge G, Holmlund T, Khvorostova J, Rofougaran R, Hofer A, Falkenberg M.

*Nucleic Acids Res*. 2008 Feb;36(2):393-403.

**IV. Structure-function defects of the TWINKLE amino-terminal region in progressive external ophthalmoplegia.**

Holmlund T, Farge G, Pande V, Korhonen J, Nilsson L, Falkenberg M.

*Biochim Biophys Acta*. 2009 Feb;1792(2):132-9.

# TABLE OF CONTENTS

# LIST OF ABBREVIATIONS

| | |
|---|---|
| adPEO | autosomal dominant Progressive External Ophtalmoplegia |
| ANT | Adenine nucleotide translocator |
| ADP | Adenosine diphosphate |
| ATAD3 | AAA Domain containing 3 protein |
| ATP | Adenosine triphosphate |
| bp | Base pair |
| COX | Cytochrome C oxidase |
| CSB | Conserved sequence block |
| C-terminal | Carboxy-terminal |
| D-loop | Displacement loop |
| dsDNA | Double-stranded DNA |
| Hsp60 | Heat-shock protein 60 |
| kb | Kilo base |
| kDa | Kilo Dalton |
| mRNA | Messenger RNA |
| mtDNA | Mitochondrial DNA |
| mtSSB | Mitochondrial single-stranded DNA binding protein |
| N-terminal | Amino-terminal |
| NTP | Nucleotide triphosphate |
| NTPase | NTP hydrolysis |
| $O_H$ | Origin of heavy strand |
| $O_L$ | Origin of light strand |
| OXPHOS | Oxidative phosphorylation |
| POLγ | Mitochondrial DNA polymerase gamma (holoenzyme) |
| POLRMT | Mitochondrial RNA polymerase |

| | |
|---|---|
| RC | Respiratory chain |
| rRNA | Ribosomal RNA |
| ssDNA | Single-stranded DNA |
| TFAM | Mitochondrial transcription factor A |
| TFB1M | Mitochondrial transcription factor B1 |
| TFB2M | Mitochondrial transcription factor B2 |
| tRNA | Transfer RNA |
| $tRNA^{Phe}$ | Transfer RNA for Phenylalanine |
| $tRNA^{Val}$ | Transfer RNA for Valine |

# Introduction

## Mitochondria

Mitochondria are eukaryotic subcellular organelles which are commonly referred to as the powerhouse of the cell, since they produce most of the cell's energy supply. This process, called oxidative phosphorylation, converts the energy derived from breakdown of carbohydrates, fatty acids and amino acids into ATP, which is the main energy source in the cell. Mitochondria also play an important role in many other cellular processes, such as calcium buffering and apoptosis.

Mitochondria are unique among the cytoplasmic organelles in that they contain multiple copies of their own DNA, which encodes a small number of proteins, highly essential for the organelle. However, nuclear genes encode the majority of the proteins that are essential for mitochondrial function, including those responsible for expression and maintenance of the mtDNA. Proper mitochondrial function thus relies on the coordinated expression of two separate genetic systems, the mitochondrial and the nuclear genomes.

The fundamental importance of this organelle is highlighted by the escalating number of human disorders that in recent years have been shown to be due to mitochondrial dysfunction. Many rare genetic disorders are caused by mutations either in the mitochondrial DNA (mtDNA), or in nuclear genes encoding mitochondrial factors. In addition, mitochondrial dysfunction has also been implicated in more common human diseases, such as Parkinson's and Alzheimer's disease, diabetes, and even in the natural process of ageing (Trifunovic et al, 2004).

### Mitochondrial dynamics

Mitochondria adopt different morphologies depending on the cell type and the metabolic requirements of the cell. The previous view of mitochondria as kidney bean-shaped isolated organelles has been challenged with the discovery that mitochondria function in extended filaments, networks, or clusters that undergo constant migration, fission and fusion. Mitochondrial movement and distribution appear to be mediated through association with the cytoskeleton, especially microtubules. In differentiated cells, mitochondria are often localized to specific cytoplasmic regions rather than randomly distributed. Mitochondria have been found to undergo developmentally programmed changes in shape and distribution at key stages in cellular differentiation (Sato et al, 2006; Skulachev, 2001; Yaffe, 1999).

The balance between mitochondrial fission and fusion is of great importance in several aspects of cell life and death. Extended mitochondrial networks have been demonstrated to efficiently transmit energy, in the form of membrane potential, between remote parts of the cell (Skulachev, 2001). Mitochondrial fusion is also thought to counteract the manifestation of respiratory chain deficiencies caused by mutant mtDNA, since the fusion process allows complementation of mtDNA gene products between mitochondria (Sato et al, 2006). The mitochondrial fission machinery is important for mitochondrial inheritance during cytokinesis, and is also thought to actively participate in apoptosis by inducing fragmentation of the mitochondrial network before the release of cyt c and caspase

activation (Youle & Karbowski, 2005). Recent studies indicate that a disruption in the fission and fusion balance of mitochondria may underlie neurodegenerative disorders (Knott & Bossy-Wetzel, 2008).

## Mitochondrial energy production

Mitochondria are surrounded by a double membrane system. The outer membrane separates the organelles from the cytosol, while the inner membrane forms numerous folds (cristae) which extend into the interior (matrix) of the organelle, substantially increasing its surface area. The matrix contains the enzymes of the citric acid cycle, and the inner membrane contains the protein complexes involved in electron transport and oxidative phosphorylation. The inner mitochondrial membrane is, in contrast to the outer membrane, impermeable to most ions and small molecules, a property critical to maintain the proton gradient that drives oxidative phosphorylation.

The initial stage of glucose metabolism (glycolysis), with the conversion of glucose to pyruvate, takes place in the cytosol. Pyruvate is then transported into mitochondria where it is subsequently oxidized to $CO_2$ to yield most of the cellular energy in the form of ATP. Pyruvate is first oxidized to acetyl CoA, which is then metabolized to $CO_2$ in the citric acid cycle. The degradation of fatty acids and amino acids also yields acetyl CoA, which is broken down in the citric acid cycle.

The enzymes of the citric acid cycle are located in the mitochondrial matrix. The oxidation of acetyl CoA to $CO_2$ is coupled to the reduction of $NAD^+$ and FAD to NADH and $FADH_2$, respectively. The high-energy electrons from NADH and $FADH_2$ are then transferred through the electron transport chain to the final electron acceptor; $O_2$, which is reduced to $H_2O$. The electron transport chain, or the respiratory chain (RC), consists of a series of five multimeric protein complexes located in the inner mitochondrial membrane (Figure 1). The RC also requires two small electron carriers, coenzyme Q and Cytochrome c. The electron transport is coupled to the generation of an electrochemical proton gradient across the inner membrane. The potential energy stored in this gradient is harvested by complex V (ATP synthase) which couples ATP production to the energetically favorable return of protons to the mitochondrial matrix. The transport of ATP and ADP across the inner mitochondrial membrane is mediated by an integral membrane protein called the adenine nucleotide translocator (ANT), which transports one molecule of ADP into the matrix in exchange for one ATP that is transferred to the cytosol, an energetically favorable reaction driven by the voltage gradient across the membrane (Saraste, 1999).

The RC consists of more than 90 different protein subunits, with the majority being nuclear encoded and only 13 being encoded by the mtDNA. The mtDNA in humans encode for subunits in four out of five complexes; complex II is entirely nuclear encoded. Recent studies suggest that the RC complexes are organized into supercomplexes, so called “respirasomes”, that have been found in mitochondria of yeast, plants and mammals (Krause et al, 2004; Schafer et al, 2006; Schagger & Pfeiffer, 2000).

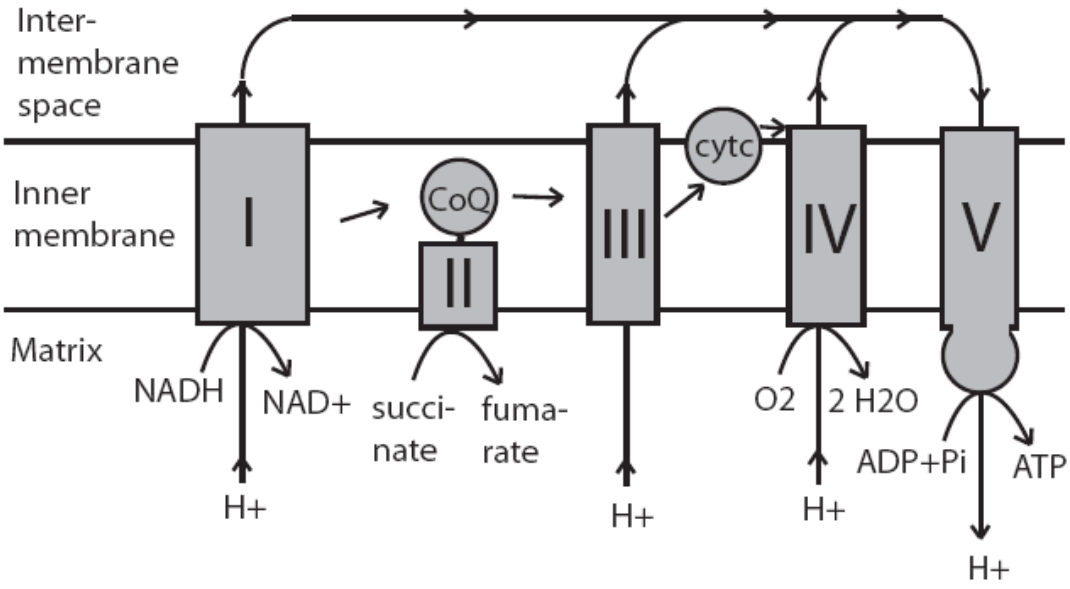


**Figure 1.** Schematic illustration of the respiratory chain. The respiratory chain consists of five different enzyme complexes (complexes I-V), coenzyme Q (CoQ) and cytochrome c (cyt c). NADH and succinate are oxidized by complexes I and II, the electrons are transferred to coenzyme Q, complex III, cytochrome c, complex IV, and finally to molecular oxygen ($0_2$), which is reduced to water ($H_20$). 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 & Clayton, 1995)).

## Origin of mitochondria

The presence of a separate mitochondrial genome can be explained by the widely accepted endosymbiotic theory, according to which mitochondria derive from an α-proteobacterium that entered an Archea-type host around 2 billion years ago and the two organisms subsequently developed a symbiotic relationship (Gray et al, 1999). The ability of these bacteria to perform respiration in host cells that had relied on glycolysis and fermentation would have provided a considerable evolutionary advantage. The closest known eubacterial relative of mitochondria is *Rickettsia prowazekii*, an obligate intracellular parasite that causes epidemic typhus in humans. During the course of evolution, ancient genes from the endosymbiont were transferred to the nuclear genome, as is evident by the presence of orthologous genes in the mtDNA in some species and in the nuclear genome in other species (Andersson et al, 2003).

Because of the bacterial origin of mitochondria, one might expect the enzymes responsible for mtDNA replication and transcription to resemble those of eubacteria. Instead three of the key components (the mitochondrial RNA polymerase POLRMT, the catalytic subunit of mtDNA polymerase POLγA, and the replicative mitochondrial helicase TWINKLE) are similar to proteins encoded by the T-odd lineage of bacteriophages, including bacteriophages T7 and T3. Probably a free-living phage was acquired in addition to, and at the same time as

the endosymbiont genome, and became the source of the phage-like replication proteins (Shutt & Gray, 2006a).

## Organization of the mitochondrial genome

Mitochondrial DNA (mtDNA) was discovered in the 1960s (Nass & Nass, 1963) and the human mtDNA molecule was the first to be fully sequenced (Anderson et al, 1981). Most mammalian cells contain hundreds of mitochondria, and each mitochondrion in turn contains several (2-10) copies of mtDNA.

The mtDNA is a double stranded circular molecule of about 16 kb. In addition to 13 respiratory chain complex components (Figure 2), the mtDNA also encodes 2 ribosomal RNAs and 22 transfer RNAs that are required for the translation of the mtDNA encoded proteins in the mitochondrial matrix. The two strands are designated heavy (H) and light (L) strand because they can be separated on a cesium chloride gradient due to differences in nucleotide content; the H strand is guanine rich and the L strand is guanine poor (Falkenberg et al, 2007).

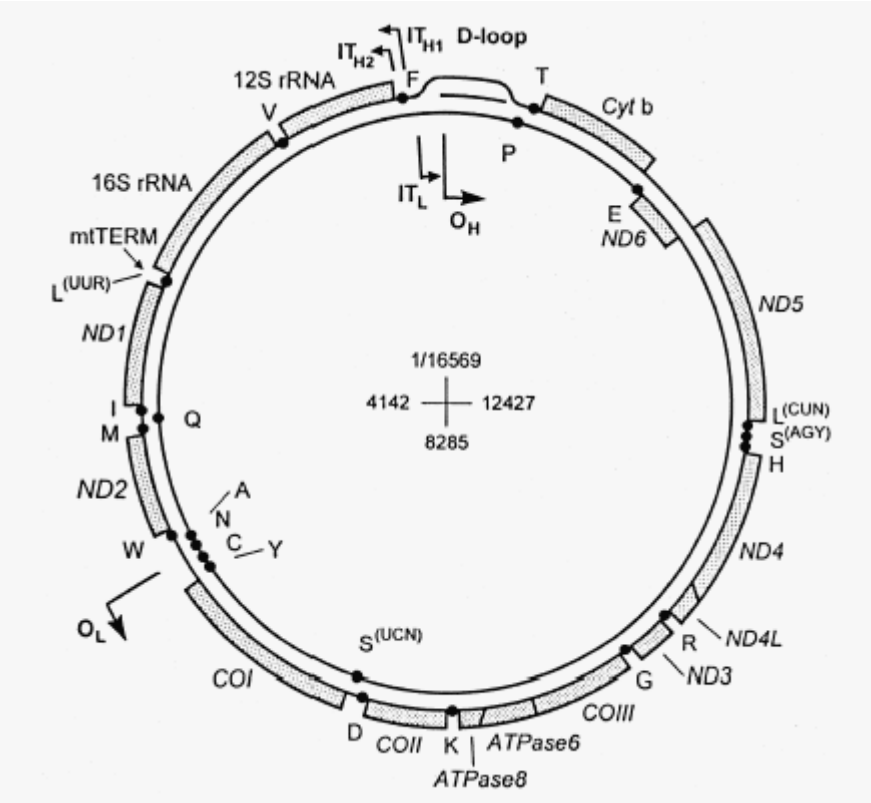


**Figure 2.** Gene content and organization of the human mitochondrial DNA. The outer circle represents the heavy-strand (H-strand) and the inner circle represents the light-strand (L-strand). The initiation of transcription sites ($IT_L$, $IT_{H1}$, $IT_{H2}$) and the direction of RNA synthesis are denoted by short bent arrows. The D-loop is shown as a triple-stranded structure. The origins of H-strand ($O_H$) and L-strand ($O_L$) replication and the direction of DNA synthesis are indicated by long bent arrows. The 22 tRNA genes are depicted by dots and the genes are depicted by shaded boxes (Taanman, 1999).

The mitochondrial genome has a very compact organization, lacking introns and often stop codons are not encoded in the DNA, but are created by mRNA- polyadenylation (Anderson et al, 1981). mtDNA has only two noncoding regions, which are control elements for mtDNA expression. The largest noncoding region is the displacement loop (D-loop) which is a triple stranded DNA region of about 1 kb in humans (Figure 4). This region contains both the origin of heavy strand replication ($O_H$) and the promoters for H- and L-strand transcription. The D-loop is formed by premature termination of H-strand DNA synthesis soon after initiation. The arrested nascent heavy strand, of about 500-700 nucleotides (called the 7S DNA), remains hybridized to the parental molecule, creating a triple stranded structure characterized by displacement of the parental H-strand (Shadel & Clayton, 1997). The second noncoding region, comprising the origin of light strand replication ($O_L$), is a 30 nt long sequence located within a tRNA cluster at about 2/3 of the mtDNA length from $O_H$. (Fernandez-Silva et al, 2003).

### Crosstalk between the nuclear and mitochondrial genomes

Mitochondria are not self-sufficient organelles but also rely on the nuclear genome for function (Figure 3). The majority of the components of the respiratory chain complexes, as well as the enzymes required for mtDNA expression and maintenance, are encoded by the nuclear DNA. Only about 1-2% of the mitochondrial proteins are encoded by the mtDNA and synthesized within the organelle. Therefore import of nuclear encoded proteins into mitochondria, and sorting of the proteins to the correct submitochondrial compartment is of crucial importance for mitochondrial function.

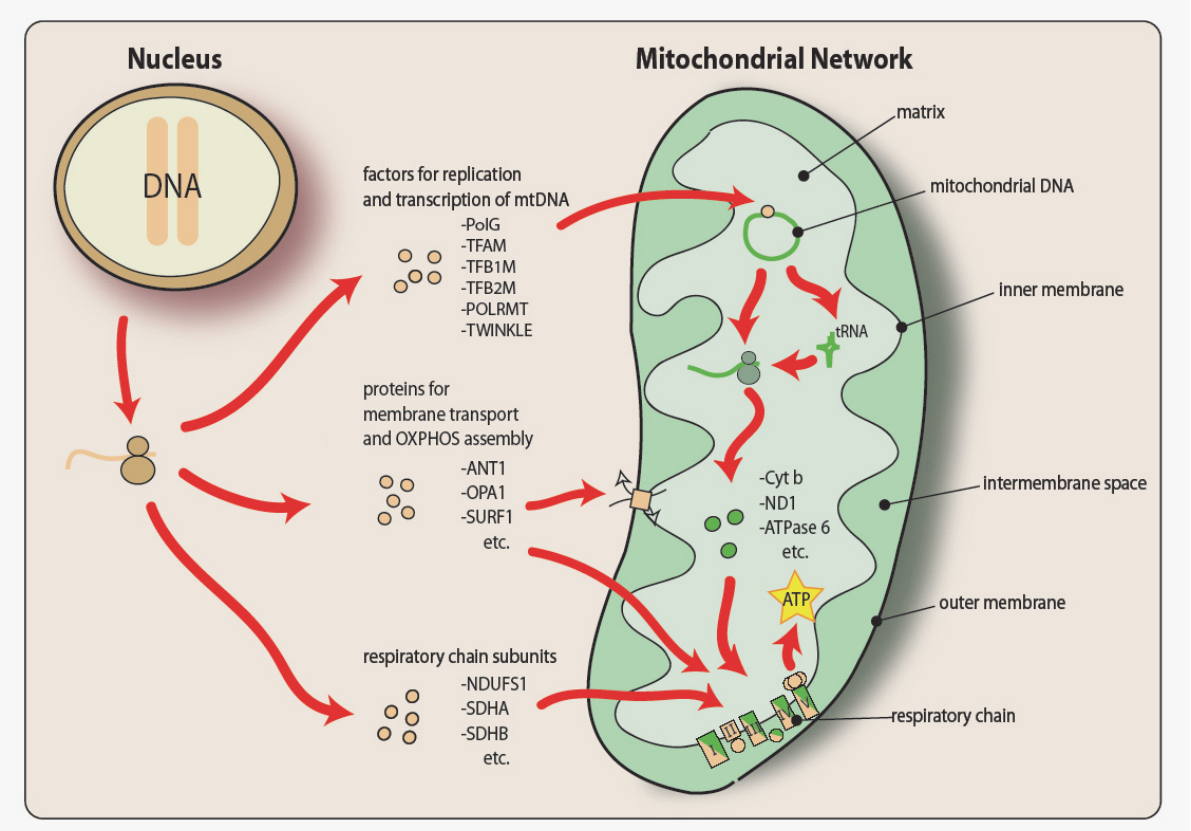


**Figure 3.** Mitochondrial function relies on a coordinated expression of mitochondrial proteins from two genomes. All factors required for replication and transcription of the mtDNA are encoded in the nucleus. In courtesy of Lene Sörensen.

Precursor proteins destined for the mitochondrial matrix generally carry an N-terminal pre-sequence that targets the protein to the organelle. The pre-sequence is cleaved during the mitochondrial import which takes place through specialized carriers located in the mitochondrial membranes. Membrane proteins on the other hand carry internal targeting sequences that are not cleaved upon import (Habib et al, 2007).

## Nucleoids

Although more than four decades have gone since the discovery of mtDNA, its organization and inheritance is still not fully understood. Mitochondrial DNA is organized in aggregates known as mitochondrial nucleoids (mt-nucleoids). Each nucleoid, which is considered the heritable unit of mtDNA, contains several mtDNA molecules, packed with a large number of different proteins in a space with a diameter of 70 nm. About 30 different proteins in different species have been identified as potential components of mt-nucleoids. The nucleoids are believed to be associated with the inner mitochondrial membrane, and are dynamic structures able to divide and redistribute in the mitochondrial network (Kucej & Butow, 2007). The functional roles for the nucleoids are not yet fully established, but they have been proposed to be of importance for protection, segregation and inheritance of the mtDNA, as well as being the biosynthetic centers for mtDNA transactions such as replication, transcription and repair (Bogenhagen et al, 2008; Chen & Butow, 2005).

Much of the current knowledge about nucleoids is based on studies of the yeast *S. Cervisiae*. Cytological, biochemical and genetic studies have helped to shed light on their organization and composition. The main packaging proteins of mt-nucleoids is a non-histone, high mobility group (HMG) protein, denoted Abf2 in budding yeast and transcription factor A (TFAM) in human cells (Chen & Butow, 2005).

Human nucleoids have been isolated by immunoaffinity purification and analyses of the protein content identified 21 proteins, including many proteins with previously known function in mtDNA replication and transcription, e.g. TFAM, mtSSB and TWINKLE (Wang & Bogenhagen, 2006). The major nucleoid protein TFAM, the mammalian homologue of yeast Abf2, has the ability to bind, unwind and bend DNA *in vitro*, and is thought to be a key player in mtDNA packaging in mammals (Fisher et al, 1992). Other examples of nucleoid proteins identified are some heat shock proteins, proteases, cytoskeletal components and a number of metabolic enzymes. The recruitment of these proteins to mt-nucleoids may allow a coupling of mitochondrial metabolism and biogenesis to mtDNA and nucleoid activities (Chen & Butow, 2005; Kucej & Butow, 2007).

Recently, a layered structure of nucleoids has been proposed, with proteins involved in mtDNA transactions such as replication, transcription and repair, associated with mtDNA in a central core. Meanwhile, translation and RC complex assembly instead occur in the peripheral region of the nucleoid, upon which nascent proteins are inserted into the inner mitochondrial membrane. In support of this model, many proteins known to be involved in mtDNA replication and transcription were able to cross-link to mtDNA, e.g. TFAM, mtSSB, TWINKLE and POLγ. Several other metabolic proteins and chaperones identified in native nucleoids were not able to cross-link to mtDNA and were thus proposed to be located in the nucleoid periphery, for example ATAD3, Hsp60 and prohibitin (Bogenhagen et al, 2008).

## Transcription of mtDNA

Transcription of human mitochondrial DNA is initiated from two distinct promoters, both located in the D-loop; the light strand promoter (LSP) and the heavy strand promoter (HSP). Transcription from these promoters produces large polycistronic precursor RNAs, which are then further processed to generate the individual mRNA, rRNA and tRNA molecules. In addition, transcription from LSP provides the RNA primers necessary for initiation of leading heavy (H) strand DNA replication (Shadel & Clayton, 1997).

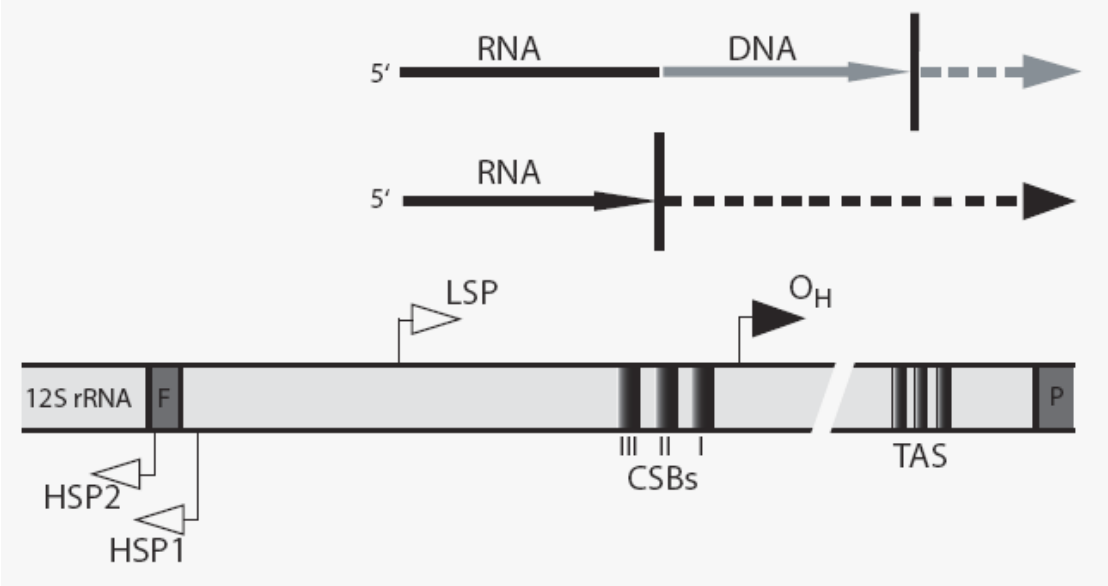


**Figure 4.** mtDNA regulatory elements. See text for details. (Adapted from (Falkenberg et al, 2007)).

### Heavy and light strand promoters

Transcription of the heavy strand is initiated from two different sites; $HSP_1$ and $HSP_2$ (Montoya et al, 1982). The $HSP_1$ site is located upstream of the $tRNA^{Phe}$ gene and produces a transcript which terminates at the 3′ end of the 16S rRNA gene. The $HSP_1$ site operates about 20 times more frequently than $HSP_2$ and generates two rRNAs (12S and 16S) and two tRNAs ($tRNA^{Phe}$ and $tRNA^{Val}$). The second initiation site, $HSP_2$, is located close to the 12S rRNA 5′ end and produces a polycistronic transcript covering almost the entire H-strand. The light strand promoter (LSP), located about 150 bp away from $HSP_1$, initiates a single polycistronic transcript, encoding 8 tRNAs and one mRNA (Fernandez-Silva et al, 2003).

Nearly all the protein and rRNA genes are flanked by at least one tRNA gene. Excision of tRNA molecules is required to produce mature mRNA and rRNA molecules from the polycistronic transcripts. The secondary structures of the tRNA sequences have been proposed as punctuation marks in the reading of the mitochondrial information, by providing the substrates for RNA processing enzymes (Ojala et al, 1981).

### Factors required for mtDNA transcription

The human mitochondrial RNA polymerase (POLRMT) is a single subunit enzyme of 120 kDa which shows a high sequence similarity with both the T3 and T7 bacteriophage polymerases (Tiranti et al, 1997). Although POLRMT can specifically recognize mitochondrial promoters, it

cannot initiate transcription without assistance of two additional transcription factors, TFAM and either TFB1M or TFB2M (Falkenberg et al, 2002; Gaspari et al, 2004).

Like other HMG proteins, TFAM can bind, unwind and bend DNA without sequence specificity (Fisher et al, 1992). TFAM is one of the major protein components of mitochondrial nucleoids and has been proposed to coordinate the compaction of several mtDNA molecules in nucleoid formation (Kaufman et al, 2007).

TFAM has also been proposed as an important regulator of mtDNA copy number (Ekstrand et al, 2004). In addition to the sequence-unspecific DNA binding activity, TFAM also binds sequence-specifically to promoter elements immediately upstream of HSP and LSP. This sequence-specific binding has been suggested to cause structural alterations in the promoter region, possibly unwinding the start site for transcription to facilitate transcription initiation (Clayton, 1991). The vital importance of TFAM for mtDNA maintenance was demonstrated in studies of homozygous TFAM knockout mice, which showed embryonic lethality and depletion of mtDNA (Larsson et al, 1998).

With the discovery of the human transcription factors B1 and B2, the basal human transcription machinery could be reconstituted in a pure *in vitro* system with recombinant proteins (Falkenberg et al, 2002). Both TFB1M and TFB2M can form a heterodimeric complex with POLRMT, and each complex could independently together with TFAM support promoter-specific initiation of transcription. TFB2M is at least 10 times more active than TFB1M in promoting transcription. Furthermore, both factors were shown to be highly homologous to bacterial rRNA dimethyltransferases, suggesting that an RNA modifying enzyme was recruited during evolution to function as a mitochondrial transcription factor. Both TFB1M and TFB2M can act as rRNA methyltransferases (Cotney & Shadel, 2006; Seidel-Rogol et al, 2003); however TFB2M appears to be a less efficient enzyme than TFB1M. Furthermore, a mouse model showed that TFB1M is essential for the stability of the small subunit of the mitochondrial ribosome. A total loss of TFB1M was embryonic lethal and disruption of the gene in heart lead to an impaired assembly of the mitochondrial ribosome and abolished mitochondrial translation (Metodiev et al, 2009). Biochemical data also demonstrated that TFB1M does not activate or repress transcription in the presence of TFB2M (Metodiev et al, 2009), inconsistent with previous reports suggesting TFB1M to be a strong transcriptional activator (McCulloch et al, 2002; McCulloch & Shadel, 2003).

### Transcriptional termination and regulation

Transcription of the H-strand starting at $HSP_1$ terminates immediately downstream of the 16S rRNA gene. The termination event is mediated by specific binding of the mitochondrial transcription termination factor, mTERF (Montoya et al, 1983). The molecular mechanism of mTERF-dependent termination remains obscure and the structure of mTERF has not been elucidated.

Transcription starting at LSP and $HSP_2$ produces long polycistronic transcripts containing all the genetic information encoded on the respective strands. A termination site for the $HSP_2$ transcript has been found immediately upstream of the $tRNA^{Phe}$ gene and two proteins were described to bind this region, although their identities are still not known (Camasamudram et al, 2003). The precise location and mechanism for L-strand termination is still not known,

although involvement of mTERF has been implicated also in this termination event (Shang & Clayton, 1994).

Recently, three novel genes encoding proteins homologous to mTERF were identified and predicted to have mitochondrial localization (Linder et al, 2005). These were denoted MTERF2-4, with MTERF1 corresponding to the previously characterized mTERF. MTERF3 was subsequently demonstrated to be a negative regulator of transcription initiation, rather than operating in transcription termination (Park et al, 2007). MTERF2 was demonstrated to be a nucleoid component of mammalian mitochondria (Pellegrini et al, 2009) and MTERF2 knockout mice displayed mild myopathy and memory deficits associated with decreased mitochondrial transcription (Wenz et al, 2009). *In vitro* studies were indicative of MTERF2 interacting with HSP, suggesting a role in transcription initiation (Wenz et al, 2009). The function of mTERF4 is still unclear.

## DNA replication – An introduction

DNA replication is a fundamental process occurring in all living organisms to copy their DNA. This process requires a large number of proteins properly assembled at a DNA replication fork. Much of today's knowledge about DNA replication is derived from studies of *Escherichia coli* (*E. coli*) and its bacteriophages T7 and T4. These different prokaryotic systems use the same basic components for DNA synthesis, although the number and nature of individual proteins can vary. The main components at the DNA replication fork include: a DNA polymerase, a polymerase accessory factor, a single-stranded DNA binding protein and a primosome consisting of DNA helicase and DNA primase activities (Benkovic et al, 2001). Before mtDNA replication is considered, a brief introduction to the field of DNA replication is presented.

### Initiation of DNA replication

The process of DNA replication is semiconservative, meaning that each strand of the original DNA duplex serves as a template for a new complementary strand. The replication process initiates at specific sites in the genome, called replication origins. These origins are targeted by proteins that separate the two strands and initiate DNA synthesis. A bacterial chromosome has one origin, whereas each eukaryotic chromosome contains multiple origins. DNA synthesis usually proceeds bidirectionally away from an origin, producing a replication bubble with two growing forks moving in opposite directions. The process is terminated when two opposing replication forks meet.

### The replisome

The multiple actions at the DNA replication fork are carried out by a multi-protein complex commonly referred to as the replisome. The replication fork structure is created by a helicase which breaks the hydrogen bonds holding the two DNA strands together; thereby unwinding the original duplex DNA. The DNA polymerase then uses each strand as a template to synthesize a new complementary strand. Single-stranded DNA binding proteins (SSBs) have critical functions in DNA replication, as well as other cellular reactions that require the presentation of ssDNA. They prevent the reannealing of separated DNA strands

and have also been shown to stimulate the activity of DNA polymerases (Fernandez-Silva et al, 2003).

The anti-parallel nature of DNA presents a challenge for the replication machinery since DNA polymerases can only synthesize DNA in a 5′ to 3′direction. The cellular replisomes overcome this problem by synthesizing leading and lagging strands in a continuous and discontinuous manner, respectively. On the leading strand, the DNA polymerase operates continuously, behind the helicase that unwinds the duplex DNA. On the lagging strand however, a second DNA polymerase synthesizes DNA in form of short so called Okazaki fragments. Each Okazaki fragment is synthesized in the 5′ to 3′ direction, but the overall movement of the lagging strand is 3′ to 5′ relative to the replication fork (Pomerantz & O'Donnell, 2007).

DNA polymerases cannot initiate DNA synthesis de novo. Instead they begin by extending the 3´end of an RNA primer. Short RNA primers are synthesized by a specialized RNA polymerase called primase. A primase is usually only required once at the leading strand, whereas at the lagging strand, this activity is required to initiate the synthesis of each Okazaki fragment.

In order to coordinate the leading and lagging strand movements, the lagging strand is thought to form a loop back through the replisome. The leading and lagging strand DNA polymerases may thus be spatially coupled within a single replisome complex, and the direction of lagging strand movement is aligned with the overall movement of the replication fork (Sinha et al, 1980) (Figure 5).

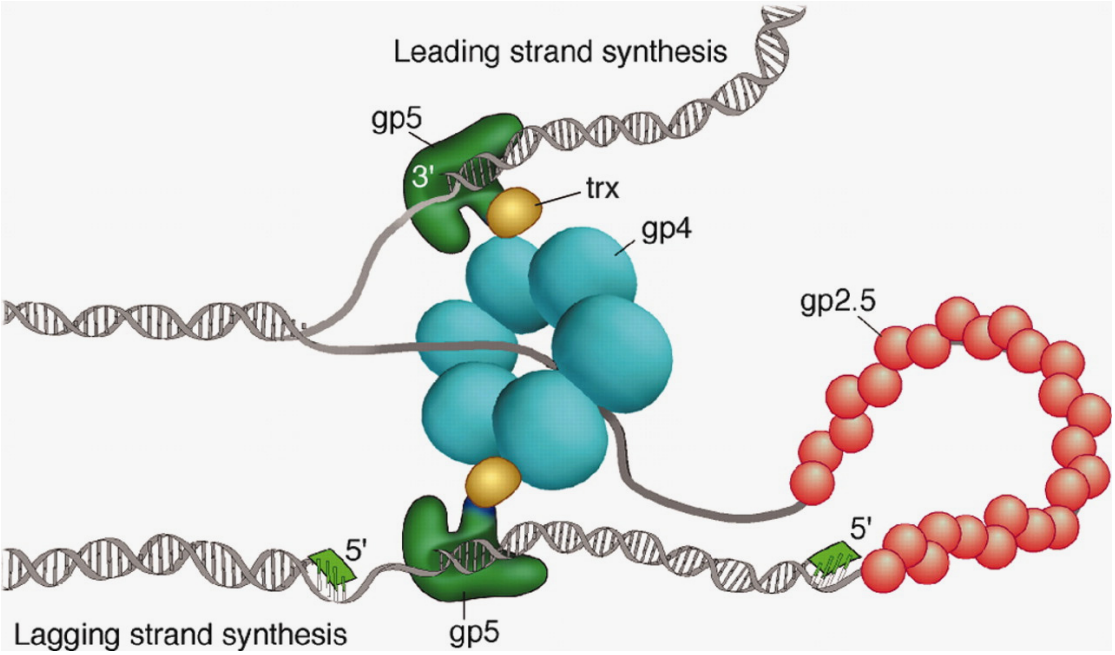


**Figure 5.** Model of the T7 replisome. The replisome consists of four proteins; DNA polymerase (gp5), the processivity factor thioredoxin (trx), the hexameric helicase/primase (gp4), and the ssDNA-binding protein (gp2.5). The polarity of lagging-strand synthesis is aligned with the synthesis of the leading-strand through loop formation (Johnson et al, 2007).

In addition, a topoisomerase is required to regulate the supercoiling of the duplex DNA that occurs during replication, and a DNA ligase activity is required to seal the gaps between neighboring Okazaki fragments.

### *The T7 replisome*

The bacteriophage T7 replication system has several similarities with the mitochondrial DNA replication system and can therefore be used as a model when studying mtDNA replication. At least three different proteins involved in mtDNA replication or transcription share homology with their counterparts in T7; the mitochondrial helicase TWINKLE, the mtDNA polymerase POLγ and the mtRNA polymerase POLRMT (Lecrenier et al, 1997; Spelbrink et al, 2001; Tiranti et al, 1997). The four proteins that are responsible for the basic reactions at the T7 replication fork are the gene 4 helicase-primase, the gene 5 DNA polymerase, the host *E. coli* encoded processivity factor thioredoxin, and the gene 2.5 single-stranded DNA binding protein (Figure 5). The polymerase forms a 1:1 complex with its processivity factor thioredoxin, and direct interactions have also been observed between all three phage encoded factors. The C-terminal tail of gp4 is responsible for interaction with the polymerase.(Richardson, 1983).

### Hexameric helicases

Helicases are motor proteins that catalyze the unwinding of duplex nucleic acid by coupling the chemical energy of NTP hydrolysis to the mechanical energy of translocation along DNA or RNA. Helicases play important roles in various cellular activities that require presentation of single-stranded nucleic acids, including replication, recombination, repair and transcription. Mutations in several genes encoding helicases are known to cause human diseases, for example xeroderma pigmentosum, Cockayne's syndrome and Werner's syndrome (Mackintosh & Raney, 2006). Some helicases are encoded by viruses, e.g. herpes simplex virus and hepatitis C virus, and are therefore targets of antiviral drug discovery. The central role of helicases, both in disease states and basic cellular processes, makes this group of enzymes an important target of research. In human cells over twenty different helicases have been identified.

Helicases can be divided into six superfamilies based on primary sequence similarities (Singleton et al, 2007). However, they can also be classified based on nucleic acid specificity (DNA, RNA or DNA-RNA hybrid), directionality (3′ to 5′, or 5′ to 3′), oligomeric state (usually monomeric, dimeric or hexameric) or whether they translocate on single-stranded or double-stranded molecules.

The hexameric helicases is a diverse group of enzymes comprising superfamilies 3-6. Many replicative helicases are hexameric and encircle the template DNA in the central channel whilst unwinding. This may be an advantage in terms of processivity, since a closed ring structure is less likely to disengage from the template strand than a monomeric or dimeric protein. However, the loading of hexameric helicases onto the template strand usually requires accessory proteins that aid in this process. An exception is bacteriophage T7 gene 4 protein (gp4) that is able to encircle a closed circular DNA molecule without additional cofactors (Ahnert et al, 2000).

Most hexameric helicases require $Mg^{2+}$ or NTP binding for hexamer formation and DNA (or RNA) greatly stabilizes the hexameric structure. Hexamer formation is required for DNA binding, and DNA binding in turn stimulates NTP hydrolysis. Conformational changes give rise to the energy transduction that couples NTP hydrolysis to movement and hence unwinding. Hexamer formation is thus of key importance for the enzymatic activity. Mutant helicases that are unable to hexamerize also show defects in DNA binding, NTPase and helicase activities (Patel & Picha, 2000).

### *T7 gene 4 protein primase-helicase*

The replicative helicase of bacteriophage T7, gene 4 protein (gp4), is probably the most extensively characterized of the hexameric helicases. Because of the homology between TWINKLE and the phage helicase, gp4 is a good model for understanding the role of TWINKLE.

Gp4 belongs to the smaller RecA/DnaB superfamily or superfamily 4. Members of this family show a conserved organization with five motifs (H1, H1a, H2, H3 and H4) important for helicase function. Only motifs H1 and H2 are similar to the conserved motifs of the two largest superfamilies 1 and 2.

In bacteria, the helicase associates with a primase (e.g. DnaB and DnaG proteins of *E. coli*), whereas in T7 the helicase and primase functions are located in one single polypeptide, with two distinct domains separated by a flexible linker. The bacteriophage T7 gene 4 encodes two proteins, a 63 kDa helicase-primase (gp4a), and a 56 kDa helicase (gp4b), translated from separated in frame start sites. The helicase activity resides in the C-terminal part of the protein, and the primase activity in the N-terminal part. The intervening linker region is believed to be critical for hexamer formation (Guo et al, 1999). Nucleotides bind at the interface between helicase monomers and thereby stimulate oligomerization (Sawaya et al, 1999). The gene 4 protein encircles the lagging strand and translocates in a 5′ to 3′ direction, thereby unwinding duplex DNA by acting as a wedge (Patel & Picha, 2000). The crystal structure of gp4 primase-helicase has been determined as a heptamer (Toth et al, 2003), however the functional unit for unwinding is believed to be a hexamer. The heptameric form is able to accommodate dsDNA, and gp4 has been shown to translocate on dsDNA and drive branch migration of a synthetic Holliday junction, suggesting that it may also be involved in recombination or repair activities (Kaplan & O'Donnell, 2002).

The gp4 has been proposed to load onto the closed circular DNA molecule by a ring-opening mechanism, where the initial contact with DNA is made by the N-terminal primase domain (Ahnert et al, 2000). The primase domain contains a conserved region with significant electropositive potential suggested both in gp4 and the structurally related DnaG primase of *E. coli* to interact with the backbone of the ssDNA template (Kato et al, 2003; Keck et al, 2000; Podobnik et al, 2000) This region represents the catalytic center of the primase, which has earlier been thought to lie on the outside of the hexameric ring (VanLoock et al, 2001). The previously accepted model was that the unwound DNA was folded back in order for the enzyme to put primers on the lagging strand (Lee & Richardson, 2002). The crystal structure of the primase-helicase instead opens up for an alternative model. The primase and helicase domains are shown to be very flexibly attached to each other and one primase

monomer can likely dip into the center of the ring and prime ssDNA in the central channel (Toth et al, 2003). This flexible structure may also allow the primase to transiently bind DNA whilst loading the hexamer onto the DNA molecule.

## Mitochondrial DNA replication

Mitochondrial biogenesis is a fundamental process in animal cell proliferation, and replication of the mitochondrial DNA is a crucial element of that process. mtDNA replication takes place in the mitochondrial matrix, independently from cell cycle phase and nuclear DNA replication (Bogenhagen & Clayton, 1977). Mitochondria replicate and divide mainly in response to the cell's energy requirements. The energetic demands can vary substantially between different cell types and under different physiological conditions requiring an adaptation in mitochondrial biogenesis. Proper mitochondrial function depends on a coordination of the nuclear and mitochondrial genomes, but the molecular mechanisms involved in this process are still poorly understood. The link between transcription and mtDNA replication suggests that mtDNA copy number in part may be controlled by the frequency of transcription initiation at LSP (Shadel & Clayton, 1997).

The mechanism by which mtDNA is replicated is not completely understood. The since long established "strand-displacement mode" (Clayton, 1982) has been challenged by a "strand-coupled mode" (Holt et al, 2000) (Figure 6). Since the strand-coupled mode was presented there has been an intense debate concerning the mode of mtDNA replication, and it is evident that further experimental work is required to clarify the situation (Bogenhagen & Clayton, 2003a; Bogenhagen & Clayton, 2003b; Holt & Jacobs, 2003).

### The strand-displacement mode (Clayton mode)

The strand-displacement mode or the Clayton mode was defined by Clayton and coworkers based on numerous biochemical and electron microscopy studies of replicative intermediates in mammalian cell lines, reviewed in Clayton (1982) and Clayton (1991). According to this mode, replication of mammalian mtDNA is thought to occur asymmetrically from two distinct strand-specific origins of replication. Transcription from LSP provides primers for initiation of leading heavy (H) strand synthesis at the origin of heavy strand ($O_H$). Replication proceeds unidirectionally, displacing the parental H strand as single-stranded DNA (ssDNA). When leading-strand synthesis has proceeded two-thirds of the genome, it exposes another major origin, the origin of light-strand synthesis ($O_L$), which is activated, and lagging-strand synthesis initiates in the opposite direction. The new mtDNA molecules are then ligated to form closed circles prior to the introduction of superhelical turns.

Recently, analysis of mtDNA replication intermediates by atomic force microscopy revealed that multiple additional initiation sites for lagging-strand mtDNA synthesis are present on the L-strand (Brown et al, 2005). These alternative initiation sites are probably less abundant, explaining why only $O_L$ was identified when 5′ end mapping was used to identify initiation sites for lagging-strand mtDNA synthesis (Tapper & Clayton, 1981).

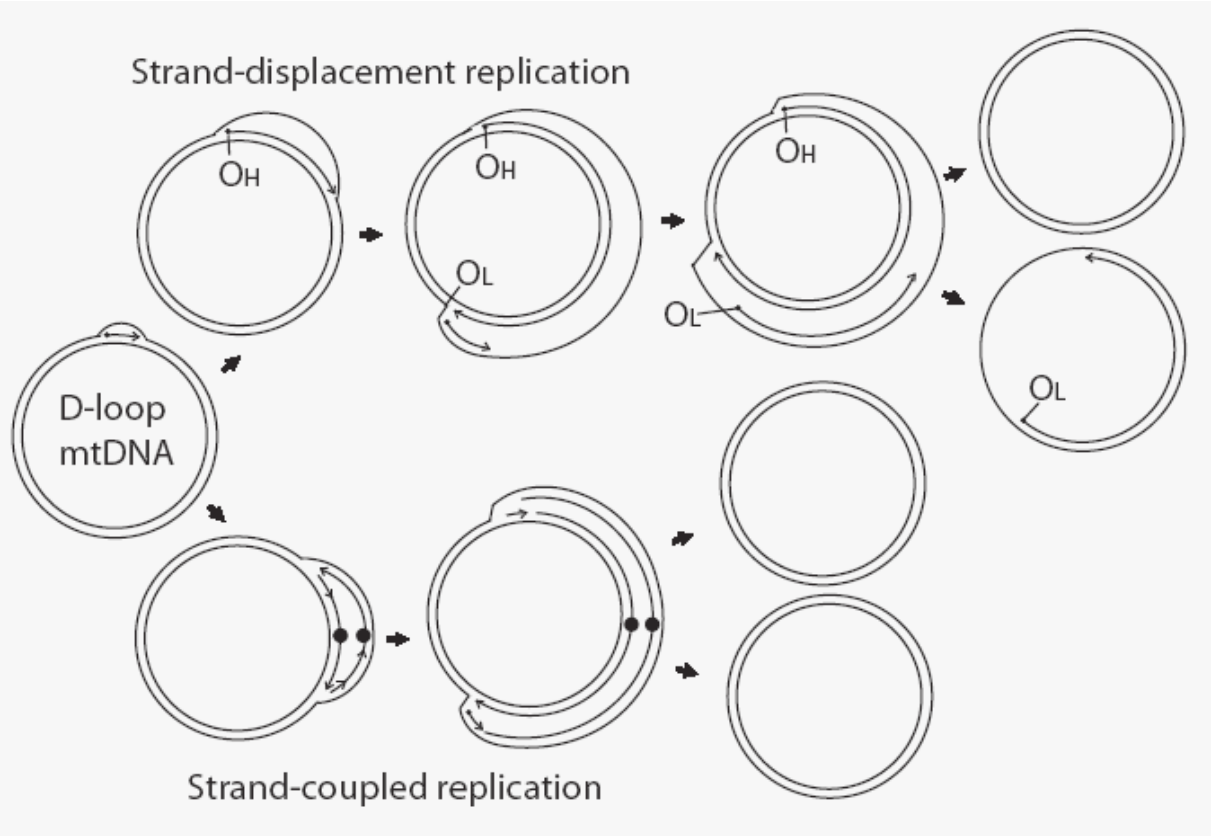


**Figure 6.** Strand-displacement (left) and strand-coupled (right) modes of mtDNA replication. According to the displacement-mode, replication initiates at $O_H$ and proceeds unidirectionally, displacing the parental H strand as ssDNA. When $O_L$ is exposed, lagging-strand synthesis initiates in the opposite direction. The asymmetry of strand synthesis leaves one segregated daughter molecule with an incompletely synthesized L-strand. In the strand-coupled mode, mtDNA replication is thought to occur symmetrically, with leading- and lagging-strand synthesis progressing bidirectionally from multiple origins across a broad zone. (Adapted from (Brown et al, 2005)).

### *Initiation of H-strand synthesis*

Transcription from LSP is not only required for gene expression, but also to generate the RNA primers used by the DNA polymerase to initiate leading (H) strand DNA replication at $O_H$. The newly synthesized RNA remains hybridized to the DNA template upstream of $O_H$ to form a stable so-called R-loop structure. To provide a properly positioned primer, the LSP transcript must either terminate or be processed at sites near $O_H$ (Shadel & Clayton, 1997). The RNA to DNA transition is thought to involve three conserved sequence blocks (termed CSBI, CSBII and CSBIII) that have been found downstream of LSP in the D-loop region of vertebrates (Walberg & Clayton, 1981) (Figure 4).

The primary LSP transcripts may be cleaved by an endoribonuclease activity at sites near $O_H$ to yield primers for DNA synthesis. A site-specific mitochondrial RNA processing endoribonuclease (RNase MRP) with this activity was identified in mouse and human cells (Clayton, 1991). The relevance of these findings has however been questioned because RNase MRP is mainly localized to the nucleolus, where it plays an important role in rRNA

processing. An alternative model was recently suggested in which CSBII can act as an LSP-transcription termination element *in vitro*, suggesting that a sequence-specific DNA element in the D-loop could also mediate the RNA to DNA transition (Pham et al, 2006). This does not exclude the RNase MRP-dependant model for primer formation, but suggests an alternative model that may exist independent of RNase MRP action.

Only a few initiation events at $O_H$ actually result in the synthesis of a full-length genome, whereas nearly 95% arrest at the termination associated sequences (TAS), creating the triplex D-loop structure. TAS are short (15 bp) DNA elements conserved in vertebrates and located at the 3' end of nascent D-loop H strands (Figure 4). The TAS sequence has been proposed to function as a major regulatory point of mtDNA replication (Madsen et al, 1993).

The triple-stranded D-loop is a conserved feature of mtDNA in vertebrates, suggesting that it has a crucial role in mtDNA metabolism. It has however been difficult to determine its precise role(s) since the frequency of D-loop structures in a population of mtDNA molecules depends on the cell type and on cell growth phase. Suggested roles for the D-loop include keeping a number of mtDNA molecules in a state primed for subsequent DNA replication, and providing an mtDNA configuration that is required either for transcriptional regulation or for mtDNA segregation (Shadel & Clayton, 1997).

### *Initiation of L-strand synthesis*

Replication of the light (L) strand is initiated when leading H strand synthesis has reached about two-thirds of its total length, revealing the origin of light-strand synthesis ($O_L$). $O_L$ consists of a small (30 bp) noncoding region and is surrounded by several tRNA genes (Figure 2). The DNA in this sequence has the potential to create a stable stem-loop structure that is thought to form once the leading strand has displaced this region as ssDNA. The stem-loop structure is conserved in evolution; however the exact nature of the primary sequence may differ significantly. It is noteworthy that $O_L$ is located within a cluster of tRNA genes that also have potential DNA secondary structures (Clayton, 1991). In chicken mtDNA $O_L$ is missing, although the five tRNA genes in this region are conserved, and it has been suggested that the tRNA genes may function as a substitute to $O_L$ in this species (Desjardins & Morais, 1990).

The initiation of DNA replication at $O_L$ requires a DNA primase activity to produce short RNA primers complementary to a T-rich portion of the stem-loop. A transition from RNA to DNA synthesis occurs near the base of the stem at a site that is actually part of a tRNA gene (Clayton, 1991). A mitochondrial DNA primase activity involved in RNA priming at $O_L$ has been partially purified from cell extract (Wong & Clayton, 1985) but the corresponding protein was never identified.

Recent work suggests that L-strand mitochondrial DNA replication is primed by the mitochondrial RNA polymerase POLRMT (Wanrooij et al, 2008). This report demonstrates that POLRMT has two distinct modes of action. The enzyme efficiently transcribes long regions of dsDNA, but becomes much less processive on ssDNA, producing short RNA stretches of only 25-75 nt. The short RNA primers can be used by the mitochondrial DNA polymerase γ to initiate DNA synthesis *in vitro* and this reaction is stimulated by the mitochondrial ssDNA binding protein (mtSSB). It was further demonstrated that, when

combined, POLRMT, DNA polymerase γ, TWINKLE and mtSSB are capable of simultaneous leading-and lagging-strand DNA synthesis *in vitro* (Wanrooij et al, 2008).

In a subsequent study $O_L$-dependant initiation was reconstituted *in vitro*, demonstrating that POLRMT indeed can function as an $O_L$-specific primase in mammalian mitochondria. POLRMT was shown to initiate primer synthesis from a poly-dT stretch in the single-stranded stem-loop structure. POLRMT synthesized about 25 nt of RNA primer, then POLRMT was replaced by POLγ to start DNA synthesis (Fuste et al, 2009).

Following the initiation at $O_L$, lagging strand synthesis proceeds over the whole length of the strand and ends after the synthesis of the nascent H-strand has terminated (Clayton, 1982). Upon completion of DNA synthesis, the new mtDNA molecules need to be separated, RNA primers removed, gaps ligated and superhelical turns introduced, but the exact mechanisms of these processes are still not well understood.

### The strand-coupled mode (Holt mode)

In recent years, several papers have appeared that challenge the strand-displacement mode of mtDNA replication (Bowmaker et al, 2003; Holt et al, 2000; Yang et al, 2002; Yasukawa et al, 2005). Holt and coworkers have instead proposed an alternative mode of mammalian mtDNA replication, based on neutral/neutral two-dimensional agarose gel electrophoresis (2D-NAGE). DNA restriction fragments that contain candidate replication origins and replication forks can be detected using this technique. DNA fragments are separated in the first dimension based on size and in the second dimension based on strand configuration. Particular types of replication intermediates can thus be resolved in predictable ways. DNA fragments with conventional synchronous leading- and lagging-strand replication forks display so-called y-arcs. Such y-arcs were revealed within an area between the two origins $O_H$ and $O_L$ (Holt et al, 2000). This and related observations lead to the conclusion that mtDNA actually replicates symmetrically, with leading- and lagging-strand synthesis progressing bidirectionally from multiple origins across a region including the cyt b, NAD5 and NAD6 genes. When the replication process reaches $O_H$, the replication fork is arrested and replication is restricted to one direction only (Bowmaker et al, 2003).

The strand-coupled mode was later modified to include an additional mode of replication observed in birds and mammals; the RITOLS mode (ribonucleotide incorporation throughout the lagging strand) (Yasukawa et al, 2006). During RITOLS replication, mtDNA synthesis starts unidirectionally from a region in the D-loop. In contrast to the strand-coupled mode, the L-strand is initially laid down as RNA before being converted into DNA by an unknown mechanism. The authors speculate that the RNA may help to protect and stabilize the displaced ssDNA, alternatively act to block transcription machineries that could interfere with the replication process.

The strand-coupled mode of mtDNA replication has been questioned by proponents of the strand-displacement model who claim that the distribution of y-arcs and other 2D-gel patterns might be explained by both alternative L-strand origins and branch migration of nascent strands in the strand-displacement mode of replication (Brown et al, 2005). Atomic force microscopy analyses of mtDNA from mouse liver tissue provided further evidence of the strand-displacement mode of mtDNA replication. These mtDNA molecules also

displayed y-arcs, thereby reconciling the original displacement mode of replication with the data obtained from 2D-gel analyses (Brown et al, 2005).

The observation of stable, nonreplicative and partially hybridized RNA (so called R-loops) on the mtDNA template casts doubt on the RITOLS mode of replication. These RNAs are claimed to have confounding interpretation of 2D-gels by mimicking patterns commonly associated with replication intermediates (Brown et al, 2008). The presence of these R-loops appeared to restrict the conformation of supercoiled mtDNA, implying a larger structural role within the mitochondrial nucleoid and possibly multiple consequences for the maintenance and expression of mtDNA.

The RITOLS mode of replication recognizes the $O_L$ region as a transition point between RNA and DNA (Yasukawa et al, 2006), meaning that the major difference between RITOLS and the strand-displacement mode is that ssDNA is supposed to be protected by RNA in the first case, and by SSB in the latter. With the discovery of additional initiation sites of lagging strand replication (Brown et al, 2005), the boundaries between the different proposed modes of mtDNA replication becomes somewhat blurred. Future studies may help to give a final picture of the mechanism of this fundamental process.

## The mitochondrial replisome

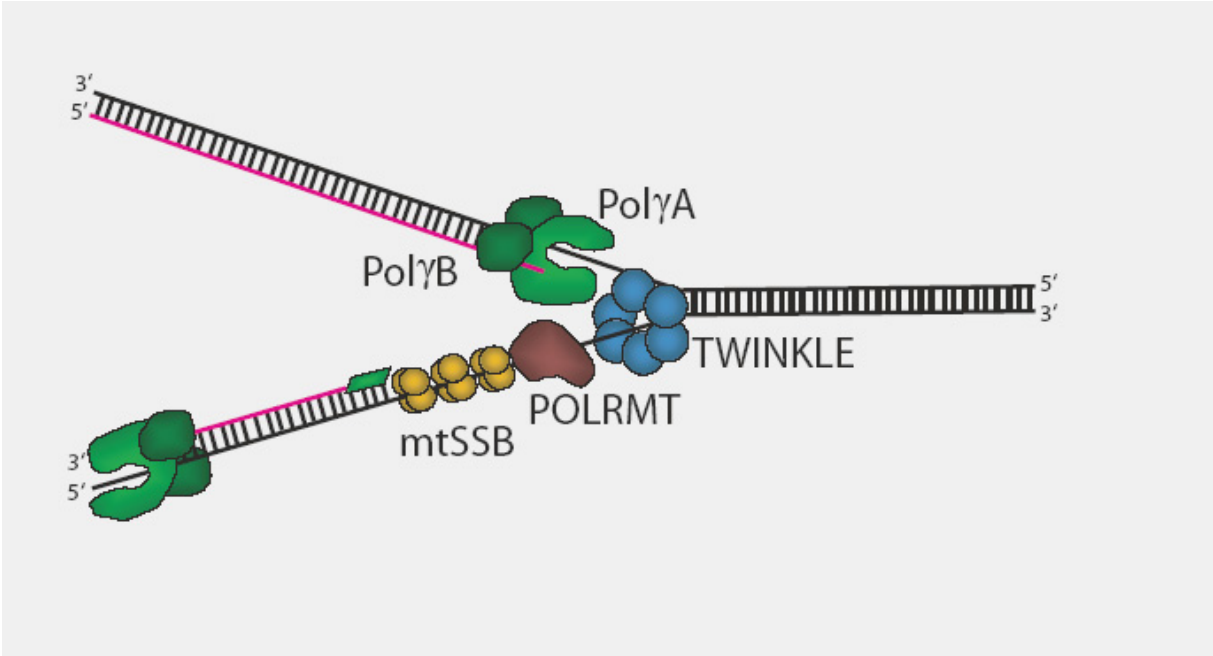


**Figure 7.** The mtDNA replication machinery. The TWINKLE helicase has 5’ to 3’ directionality and encircles the lagging strand whilst unwinding the duplex DNA template. The mtSSB protein stabilizes the unwound conformation and stimulates DNA synthesis by the POLγ holoenzyme. POLRMT synthesizes RNA primers on the lagging strand.

Four factors are required to perform the basic reactions of mammalian mitochondrial DNA replication; the DNA polymerase holoenzyme (POLγ) synthesizes a copy of the template strand, the helicase TWINKLE unwinds the duplex DNA ahead of the polymerase, mitochondrial single-stranded DNA binding protein (mtSSB) coats the ssDNA and prevents reannealing of separated strands, and finally POLRMT provides the RNA primers required for initiation of DNA replication at $O_H$ as well as $O_L$ (Fuste et al, 2009). (Figure 7). The mitochondrial replication machinery has been reconstituted *in vitro* with human recombinant proteins. POLγ and TWINKLE together form a processive replication machinery, a replisome, which can use duplex DNA as a template to synthesize ssDNA molecules of about 2 kb. Addition of mtSSB further stimulates the reaction, generating DNA products of about 16 kb, the size of the mammalian mtDNA molecule (Korhonen et al, 2004). When combined, POLRMT, POLγ, TWINKLE and mtSSB are capable of simultaneous leading- and lagging-strand DNA synthesis in vitro (Wanrooij et al, 2008). A close functional interaction between the mitochondrial replication factors is further demonstrated by the finding that several mutations in both TWINKLE and POLγA cause autosomal dominant progressive external ophtalmoplegia (adPEO) (Spelbrink et al, 2001), a human disorder characterized by multiple mtDNA deletions. Recently, a mutation in POLγB was also shown to cause the same disorder (Longley et al, 2006).

## POLγ – the mitochondrial DNA polymerase holoenzyme

POLγ is the only DNA polymerase so far reported in mitochondria, and is thus likely responsible for all DNA synthetic reactions in replication, recombination and repair events in the mitochondria. POLγ was first identified as an RNA-dependant DNA polymerase in human HeLa cells (Fridlender et al, 1972). POLγ is a highly efficient and processive polymerase with a high fidelity in nucleotide selection and incorporation. The mammalian mtDNA polymerase holoenzyme is a heterotrimer with one subunit of the catalytic POLγA, and two subunits of the processivity factor POLγB (Carrodeguas et al, 2001).

POLγA is a member of the family A class of DNA polymerases that also include the well characterized *E. coli* DNA polymerase I and the bacteriophage T7 DNA polymerase. Human POLγA is a 140 kDa peptide with 5´-3´polymerase, 3´-5´exonuclease, and 5´-deoxyribose phosphate (dRP) lyase activities. The third enzymatic activity is related to base excision repair. The polymerase and exonuclease activities reside in two distinct domains, separated by a large spacer region which also contains conserved sequence motifs (Kaguni, 2004). The function of the spacer is not yet clear, but mutational studies of *Drosophila* POLγA suggest it may be involved in DNA binding, subunit interactions and in the interplay between polymerase and exonuclease activities (Luo & Kaguni, 2005). The importance of the spacer is further demonstrated by findings that mutations in this region are linked to mitochondrial disorders (Luoma et al, 2005).

The accessory subunit of mtDNA polymerase, POLγB has a molecular mass of 55 kDa. POLγB is present in *Homo sapiens, Mus musculus, Drosophila melanogaster* and *Xenopus laevis* (Carrodeguas et al, 1999; Wang et al, 1997), but is notably absent in *Saccharomyces cerevisiae*. POLγA and POLγB form a heterodimer in *Drosophila*, but a heterotrimer ($POL\gamma AB_2$) in mammalian cells. This distinction may be explained by the fact that the *Drosophila* POLγB lacks many amino acid residues present in mouse and human POLγB, in

two motifs that are suggested to be involved in dimerization of the protein. POLγB has been shown to increase both catalytic activity and processivity of POLγA, an effect partially explained by enhanced DNA binding, but also by increased nucleotide binding (Carrodeguas et al, 2001).

POLγB shares both sequence and structural similarities with the anticodon-binding pocket of prokaryotic aminoacyl-tRNA synthetases (aaRS), a group of enzymes that catalyze the attachment of each amino acid to its corresponding tRNA. This finding lead to the suggestion that POLγB may play a role in primer recognition (Fan et al, 1999), alternatively it could be involved in directing the polymerase to the origins of replication, which have a high potential for forming complex secondary structures (Carrodeguas & Bogenhagen, 2000). There is also some structural similarity between a part of POLγB and thioredoxin, the accessory subunit of T7 DNA polymerase, although there is no amino acid sequence homology (Fan et al, 1999). POLγB has a dsDNA-binding activity that is however not essential for the stimulation of POLγA catalytic activity *in vitro*. It has been proposed that this function may play a role in other aspects of DNA replication or repair (Carrodeguas et al, 2002). The importance of this dsDNA-binding activity in mtDNA replication is addressed in Paper I of this thesis.

POLγ is clearly a vital protein for mitochondrial function. Mutations in both the catalytic and accessory subunits of POLγ cause mitochondrial disorders (Longley et al, 2006). It has also been shown that transgenic mice expressing an error-prone form of POLγ, lacking the 3´-5´-exonuclease activity, accumulate mutations and deletions in the mtDNA, and develop symptoms of premature ageing (Trifunovic et al, 2004).

### TWINKLE – the mitochondrial replicative helicase

The TWINKLE gene (*C10orf2*) was originally identified in a search for mutations associated with chromosome 10q24-linked autosomal dominant progressive external ophtalmoplegia (adPEO) (Spelbrink et al, 2001). The gene showed striking similarity with the bacteriophage T7 helicase-primase (gp4) and the protein was shown to co-localize with mtDNA in mitochondrial nucleoids. These data suggested that TWINKLE was the long-sought replicative helicase of mitochondria. Further studies demonstrated that human TWINKLE is indeed the helicase at the mitochondrial DNA replication fork (Korhonen et al, 2004). TWINKLE is a 684 amino acid protein with a predicted molecular weight of 77 kDa. cDNA amplification also revealed a splice variant of TWINKLE, named TWINKY. TWINKY is a 66 kDa peptide that lacks a part of the C-terminal helicase domain. It also shows mitochondrial localization, but its function is still unknown (Spelbrink et al, 2001).

TWINKLE, similar to T7 gp4, can be structurally divided into three distinctive regions; an amino-terminal domain with unknown function (corresponding to the primase domain of gp4), a carboxy-terminal helicase domain, and an intervening linker region. The helicase domain displays the same conserved motifs as T7 gp4 and other members of the RecA/DnaB superfamily. The similarities in the N-terminal part are however less obvious. Human TWINKLE contains five of six conserved primase motifs, but lacks essential cysteine residues of a $Zn^{2+}$ finger motif (motif I) that in gp4 is involved in nucleotide sequence recognition and transfer of the primer to the polymerase, and it also lacks essential amino acids required for

binding $Mg^{2+}$ ions. Bioinformatic analyses suggest that TWINKLE is an active primase in most eukaryotes, but that this activity has been lost in metazoan cells (Kusakabe & Richardson, 1996; Shutt & Gray, 2006b). To date, there have been no reports of primase activity in human TWINKLE. The observation that five mutations associated with adPEO are localized in the TWINKLE N-terminal domain strongly suggests that this region has some other important function involved in mitochondrial DNA replication. This issue is addressed in Paper III and IV of this thesis.

Similar to gp4 and other related ring helicases, TWINKLE forms a hexamer in solution (Farge et al, 2008; Spelbrink et al, 2001). Biochemical analyses have demonstrated that recombinant TWINKLE catalyzes the ATP-dependant unwinding of duplex DNA in a 5′ to 3′-direction and it requires specific substrates with a single-stranded 5′-loading site and a short 3′-tail to initiate unwinding. This substrate resembles the conformation of a replication fork, a structure with which TWINKLE would be expected to interact *in vivo*. mtSSB specifically stimulates TWINKLE unwinding activity. TWINKLE is however a poor helicase in isolation, since it can only unwind shorter stretches of duplex DNA (<55bp), even in the presence of mtSSB (Korhonen et al, 2003). The biochemical evidence for TWINKLE being the active helicase at the mtDNA replication fork was provided by reconstitution of a minimal mitochondrial replisome, where POLγ, together with TWINKLE and mtSSB, efficiently use dsDNA as a template to synthesize ssDNA products of up to 16 kb, which correspond to the size of the mammalian mitochondrial genome (Korhonen et al, 2004).

### mtSSB – the mitochondrial ssDNA binding protein

Replicative intermediates of mtDNA probably contain long single-stranded DNA regions. A single-stranded DNA binding protein (mtSSB) is present in mitochondria to maintain the integrity of these regions, and to increase the activity and fidelity of POLγ (Fernandez-Silva et al, 2003). MtSSB has been purified from several species and has a molecular weight of about 13-15 kDa, forming a homotetramer in solution. The crystal structure of mtSSB suggested that ssDNA wraps around the tetramer and mtSSB has been shown to coat the displaced strand that is the template for lagging strand synthesis. Unlike POLγ and TWINKLE, the other components of the minimal mtDNA replisome, mtSSB is not homologous to the T7 SSB, but is instead similar to the SSB present in *E. coli* (Kaguni, 2004). mtSSB has been shown to specifically stimulate the helicase activity of TWINKLE since it cannot be replaced with *E. coli* SSB (Korhonen et al, 2003).

## Mitochondrial genetics

In most multicellular organisms mtDNA is maternally inherited. There are several mechanisms to assure this, including simple dilution (an egg contains many more copies of mtDNA than a sperm cell), degradation of sperm mtDNA in the fertilized egg, and in some organisms failure of sperm mtDNA to enter the egg. The mitochondria from the sperm cell that enter the oocyte during fertilization are actively degraded by a ubiquitin-dependant mechanism (Sutovsky et al, 2000).

In normal cells all copies of mtDNA are assumed to be identical, a condition known as homoplasmy, but since mutations usually affect some but not all mtDNAs within a cell or tissue, the more common situation is heteroplasmy, with a mixture of wild-type and mutant mtDNA. The relative proportion of mutant and normal mtDNA in a tissue determines the clinical outcome of the mutation. A minimum critical number of mutated mtDNA molecules is required to cause respiratory chain dysfunction in a certain organ or tissue and ultimately lead to disease. This phenomenon is known as the threshold effect. The threshold at which symptoms arise depends on the energy demand of the affected tissue. Tissues with high energy demand, such as the central nervous system, skeletal muscle, heart, liver and kidney are particularly sensitive and therefore often affected in mitochondrial disorders (Graff et al, 1999).

Unlike nuclear DNA, the mtDNA is not replicated in synchrony with each cell division. Instead the mtDNA is continuously turned over and replicated during the entire cell cycle. Each cell contains hundreds or thousands of mtDNA molecules which are randomly distributed to the daughter cells at each division. The level of heteroplasmy may thus shift and the phenotype may change accordingly. This phenomenon, called mitotic segregation, explains how certain patients with mitochondrial disorders may shift clinical phenotypes as they grow older. A variable segregation early in embryogenesis may lead to different levels of mutated mtDNA between tissues. It can also occur in the germline, causing a dramatic change in levels of mutated mtDNA between generations, and siblings often inherit widely varying levels of mutated mtDNA (Dimauro & Davidzon, 2005; Graff et al, 1999).

During oogenesis, mitochondria are randomly segregated through tight “bottlenecks”; of the 150 000 mtDNA molecules estimated to be present in mammalian oocytes, only a small proportion is amplified and transmitted to the offspring. The random distribution of wild-type and mutant mtDNA during oogenesis creates a spectrum of heteroplasmy across the oocyte population, explaining the clinical variations of a mitochondrial disorder among siblings (Schapira, 2006).

The mutation rate of mtDNA is much faster than that of the nuclear genome. There are several possible explanations for this; mtDNA is less protected by proteins, it is close to the site of ROS production at the inner mitochondrial membrane, and it appears to have less efficient repair systems than the nucleus (Fernandez-Silva et al, 2003). The mitochondrial genetic bottleneck contributes to the rapid segregation of new genotypes, which are either lost during transmission or reach very high levels and affect the offspring (Cree et al, 2008). Recent studies demonstrate a strong purifying selection in the female germline to get rid of mtDNA aberrations that severely impair respiratory chain function (Stewart et al, 2008).

Another feature that distinguishes the mitochondrial genome from the nuclear genome is that mitochondrial genes are translated using a genetic code distinct from the universal genetic code. For example, UGA encodes tryptophan instead of being a stop codon. Mitochondria also have a simplified codon-anticodon pairing system that allows protein translation with only 22 tRNAs (Fernandez-Silva et al, 2003).

## Mitochondrial disorders

The first evidence of mitochondria being linked to human disease was presented almost 50 years ago in a patient with severe hypermetabolism (Luft et al, 1962). However, the genetic basis of mitochondrial disorders remained unclear until 1988 when the first disease-causing mutations of mtDNA were found (Holt et al, 1988; Wallace et al, 1988). Since then more than 150 point mutations and a large number of rearrangements (mainly deletions and duplications) have been associated with a variety of diseases (Dimauro & Davidzon, 2005). Mitochondrial diseases that were for many years considered being extremely rare, are now being described as among the most common genetic disorders, affecting at least 1 in 5000 people, when both nuclear and mtDNA mutations are considered (Schaefer et al, 2004). This could be an underestimate since mitochondrial disorders are often complex and difficult to diagnose.

Mitochondrial disorders show a wide spectrum of clinical presentation. They range from lesions of single tissues to complex multisystem syndromes with onset ranging from neonatal to adult life. Tissues with a high energy demand, such as neurons, cardiac and skeletal muscle, have a high density of mitochondria and are therefore extra sensitive to mitochondrial abnormalities. Consequently these tissues are often affected in mitochondrial disease. When a mitochondrial disorder is suspected, the diagnosis is typically confirmed by analyses of muscle biopsies. A classical hallmark of mitochondrial dysfunction is “ragged-red fibers” caused by accumulation of diseased mitochondria in muscle fibers. Also cytochrome-c-oxidase (COX) negative fibers are common. These features are however absent in several mitochondrial disorders, making diagnosis complicated (Schapira, 2006; Zeviani & Di Donato, 2004).

Abnormalities of mtDNA or OXPHOS activity has been demonstrated in several different neurodegenerative diseases, including Parkinson’s disease, Alzheimer’s disease and Friedreich’s ataxia (Schapira, 2006). It is unclear whether these abnormalities are the primary cause of disease or are defects not directly related to pathogenesis. If mitochondrial dysfunction does play a role in pathogenesis, improving mitochondrial function may also improve disease progression. Coenzyme $Q_{10}$ is an attractive target for treatment of Parkinson’s disease (PD), as it both enhances respiratory chain function and scavenges free radicals. A pilot study showed that Coenzyme $Q_{10}$ indeed had a beneficial effect upon PD pathogenesis in patients (Shults et al, 2002).

Mitochondrial dysfunction is also believed to contribute to human ageing. Progressive accumulation of somatic mtDNA mutations and reduction in respiratory chain function has been found to correlate with increasing age (Cottrell et al, 2000). In direct support of this theory, transgenic mice expressing an error-prone form of POLγ, lacking the 3´-5´-exonuclease activity, showed a `mutator´ phenotype with accumulation of mutations and deletions in the mtDNA, and they developed symptoms of premature ageing (Trifunovic et al, 2004). The causative role of these mtDNA mutations is however under debate, since mutation levels in tissues have been reported too low to cause respiratory chain function. It has been suggested that accumulation of mtDNA mutations may contribute to the ageing process by increasing cellular apoptosis (Kujoth et al, 2005). Fusion of mitochondria is thought to counteract the manifestation of respiratory chain deficiency caused by mutant mtDNAs because it allows complementation of mitochondrial genetic components

throughout the mitochondrial network (Sato et al, 2006). A decrease of mitochondrial fission and fusion activity has been observed in old endothelial cells, indicating that these processes are sensitive to ageing and may contribute to the accumulation of damaged mitochondria during ageing (Jendrach et al, 2005).

## Mitochondrial DNA defects

There are two major types of mtDNA defects that may occur: large rearrangements (including deletions and duplications) and point mutations. The defects can be present in the germline and be transmitted maternally, or appear randomly in somatic cells to cause sporadic cases of disease (Graff et al, 1999).

The first mtDNA defects associated with human disease were mtDNA deletions (Holt et al, 1988). Large-scale deletions have typically been associated with chronic progressive external ophtalmoplegia (CPEO), Kearns-Sayre syndrome and Pearson's syndrome, but have also been described in patients with MELAS (myopathy encephalopathy lactic acidosis and stroke-like episodes), diabetes or cardiomyopathy (Schapira, 2006).

Over 150 point mutations associated with human disease have been described in protein coding genes, tRNAs and rRNAs (Dimauro & Davidzon, 2005). They have a wide clinical presentation, including phenotypes like MELAS, MERRF (myoclonic epilepsy and ragged red fibers), NARP (neuropathy ataxia and retinitis pigmentosa), MILS (maternally inherited Leigh syndrome) and LHON (Leber hereditary optic neuropathy). MtDNA point mutations have also been reported in patients with for example diabetes, cardiomyopathy, sensorineural deafness, Parkinson's disease and ALS (amyotrophic lateral sclerosis) (Zeviani & Di Donato, 2004).

A disease expressed in both sexes but with no evidence of paternal transmission is strongly indicative of a mtDNA mutation. The pathogenesis of mtDNA mutations is complex. One single mutation can result in multiple phenotypes, and the same phenotype can be caused by many different mutations. It appears that multiple factors may influence the clinical outcome of a mutation, including tissue distribution, level of heteroplasmy, nuclear background, and the varying energy demand of different organs (Schapira, 2006).

## Nuclear gene mutations affecting mtDNA stability

MtDNA is dependent on nuclear DNA for the production of a range of proteins involved in its replication, transcription, translation, repair and maintenance. Mutations in nuclear genes affecting mtDNA transactions may cause mtDNA deletions, depletion, or point mutations, which leads to tissue dysfunction. These diseases thus clinically resemble those caused by mtDNA mutations, but follow a Mendelian pattern of inheritance.

Several diseases associated with multiple mtDNA deletions have been characterized, including MDS (mtDNA depletion syndrome), a severe disease of early childhood, MNGIE (mitochondrial neurogastrointestinal encephalomyopathy) and adPEO (autosomal dominant progressive external ophtalmoplegia) (Suomalainen & Kaukonen, 2001).

MNGIE is an autosomal recessive disorder characterized by symptoms like PEO, gastrointestinal dysmotility, peripheral neuropathy and mitochondrial dysfunction. The major cause of the disease is mutations in the thymidine phosphorylase (TP) gene. TP is a multifunctional enzyme which drives the thymidine salvage pathway. Defects of TP is thought to produce an excess of dTTP, causing an imbalance of the dNTP pools, which can reduce both the rate and fidelity of POLγ in mtDNA replication (Zeviani & Di Donato, 2004).

### *AdPEO*

Autosomal dominant progressive external ophtalmoplegia, or AdPEO, is an adult-onset human disorder characterized by multiple deletions in the mitochondrial DNA of skeletal muscle. Mutations in four different genes are reported to cause AdPEO; POLγA, POLγB, TWINKLE and adenine nucleotide translocator 1 (ANT1), the heart and muscle specific isoform of the ADP/ATP transporter (Agostino et al, 2003). ANT1 mutations probably affect the dATP pools within mitochondria. The functional interactions between POLγ and TWINKLE at the mtDNA replication fork may explain why mutations in these genes cause the same disorder.

Symptoms vary largely between and even within families, but the hallmark symptom is PEO (weakness of the external eye muscles), which causes ptosis and limitations of eye movements. Patients may also suffer from general muscle weakness and exercise intolerance. In addition, some patients develop peripheral neuropathy, deafness, ataxia, depression, cataracts, and other symptoms indicative of multiple organ involvement (Kiechl et al, 2004). A diagnose can be established by muscle biopsies showing ragged-red fibers, variable respiratory chain dysfunction and multiple mtDNA deletions (Van Goethem et al, 2001).

PEO is a common manifestation of mitochondrial myopathies and most cases appear to be sporadic and due to large scale rearrangements (deletions and duplications) of the mtDNA (Zeviani & Di Donato, 2004). Maternally inherited forms due to mtDNA point mutations have also been reported. 15% of PEO cases appear to be of autosomal dominant or autosomal recessive origin, and follow a Mendelian pattern of inheritance (Agostino et al, 2003).

In the nuclear encoded forms of PEO, deletion formation in the mtDNA is a secondary event caused by a primary nuclear defect. Deletion formation has been shown to be rare (Moslemi et al, 1996) but deleted small mtDNA molecules reportedly replicate faster than large wild-type molecules (Moraes et al, 1999) and may thus increase in proportion with age, until they reach a threshold for respiratory chain deficiency.

No effective treatment is yet available for mitochondrial disorders. Only palliative therapies are currently offered to patients. New interesting strategies are however being developed, mostly aimed at shifting the level of heteroplasmy in affected tissues (Dimauro & Davidzon, 2005). These strategies are still at an experimental level, but may offer some hope to the future for patients who suffer from mitochondrial disease.

# Aims of this thesis

The mitochondrial DNA replication system has been reconstituted *in vitro* with purified human recombinant proteins. The minimal replisome consists of DNA polymerase γ holoenzyme (including subunit A and B), the TWINKLE helicase and the single-stranded DNA binding protein mtSSB. These protein factors together form a processive replication machinery, a replisome, which can use duplex DNA as a template to synthesize ssDNA products of up to 16 kb, corresponding to the size of the mammalian mitochondrial genome (Korhonen et al, 2004). A close functional interaction between the mitochondrial replication factors may explain why mutations in both TWINKLE, POLγA and POLγB can cause the same disorder, adPEO (Longley et al, 2006; Spelbrink et al, 2001).

The general aim of this thesis has been to further investigate the molecular mechanisms of mitochondrial DNA replication, by studying both individual enzymatic activities as well as the combined actions of the different protein factors on the replication fork. The *in vitro* mtDNA replication system has been an essential tool in the biochemical characterization of the replication factors we have investigated in this thesis. The main focus has been on the mitochondrial helicase TWINKLE. The specific aims of each paper were:

**PAPER I**: To adress the importance of the smaller, accessory B subunit of DNA polymerase γ in the mtDNA replication machinery.

**PAPER II**: To investigate the molecular basis for autosomal dominant progressive exernal ophtalmoplegia (adPEO), by characterizing the biochemical consequences of seven different adPEO causing mutations in the linker region of TWINKLE.

**PAPER III**: To address the molecular role of the N-terminal domain of human TWINKLE.

**PAPER IV**: To investigate the molecular basis for autosomal dominant progressive exernal ophtalmoplegia (adPEO), by characterizing the biochemical consequences of four different adPEO causing mutations in the N-terminal region of TWINKLE.

# Results and discussion

### Paper I: The accessory subunit B of DNA polymerase gamma is required for mitochondrial replisome function

In mitochondria, the only known DNA polymerase is POLγ. The POLγ holoenzyme consists of a catalytic subunit (POLγA) with polymerase and 3'-5' exonuclease activities, and a smaller accessory subunit (POLγB). POLγB acts as a processivity factor, which increases the affinity of the polymerase for DNA and promotes tighter nucleotide binding, thereby increasing the polymerization rate (Carrodeguas et al, 2002; Kaguni, 2004).

In this work, we adressed the importance of the smaller, accessory B subunit of DNA polymerase γ in the mtDNA replication machinery. Four mutant versions of human POLγB were expressed and purified in recombinant form; IF, VV, RK and RKK, based on information from the crystal structure of POLγB (Carrodeguas et al, 2001). The IF construct was mutated in a conserved region required for interactions with POLγA and the VV mutation was expected to affect metal ion coordination. The RK and RKK mutations have previously been shown to abolish the dsDNA-binding activity of POLγB (Carrodeguas et al, 2002).

We showed that the dsDNA-binding activity of POLγB is not required to stabilize POLγA interactions with a primed DNA substrate, nor to stimulate POLγA processivity or DNA synthesis rate on a ssDNA template. POLγB was shown to block the exonuclease activity of POLγA, but the dsDNA-binding activity was not essential for this inhibition.

The role of POLγB for DNA synthesis on a dsDNA template was examined using the previously described minimal mtDNA replisome (Korhonen et al, 2004). This system requires the helicase TWINKLE in order for the polymerase to elongate through double-stranded regions. POLγB was shown to be absolutely required for mtDNA replisome function. Only the VV mutant was able to support elongation through dsDNA, concluding that both the ability of POLγB to interact with POLγA and the dsDNA-binding activity is required for full replisome function.

When the mitochondrial replisome synthesizes DNA on a duplex template, the speed of DNA synthesis is much slower than for the POLγ holoenzyme operating alone on an ssDNA template. This slower rate of the replisome may suggest that the TWINKLE helicase dictates the speed of the mtDNA replication fork and that the trailing POLγ holoenzyme is slowed down by the TWINKLE helicase ahead. We speculate that POLγB has two crucial roles in this process, explaining the characteristics of the accessory subunit that we report in this paper: If the polymerase is tailgating a slow moving helicase, it may be stimulated to enter the exonuclease mode. POLγB could efficiently prevent this by inhibiting the POLγA exonuclease activity. Second, the dsDNA-binding capacity of POLγB may be required to ensure that the polymerase stays bound to the template behind the slow-moving TWINKLE helicase.

### Paper II: Structure-function defects of the TWINKLE linker region in progressive external ophtalmoplegia

In this study we aimed to characterize seven naturally occurring AdPEO-causing mutations in the linker region of TWINKLE, to determine how they affect mtDNA replisome function. The proteins were expressed and purified in recombinant form, with the following amino acid changes: R354P, A359T, I367T, V368I, S369P, R374Q and L381P (Spelbrink et al, 2001). One protein, R354P, could not be obtained in soluble form and was thus not further studied. V368I behaved as wt TWINKLE in all our enzymatic assays. It has recently been reported that this mutation is in fact a nonpathogenic polymorphism that does not cause AdPEO as previously believed (Arenas et al, 2003).

In T7 gp4, the linker region between the primase and helicase domains is important for oligomerization and helicase activity (Guo et al, 1999; Lee & Richardson, 2004). The oligomeric state of the TWINKLE mutants was studied by size exclusion chromatography. All proteins except I367T and R374Q were able to form stable hexamers in isolation. I367T and R374Q could however form hexamers in the presence of wt TWINKLE. The mutants that could not oligomerize, I367T and R374Q, were essentially inactive in all further assays. It is very likely that TWINKLE, similarly to what has been reported for gp4, will only bind ssDNA in its hexameric form (Patel & Picha, 2000). The two mutants' inability to hexamerize therefore explains why they showed essentially no activity in all further assays; ssDNA binding, ssDNA-stimulated ATPase activity, helicase activity and rolling-circle replication assay.

The three remaining mutants showed a varying degree of activity. L381P, despite its ability to hexamerize, was totally inactive in all assays. S369P showed reduced levels of ssDNA binding, ssDNA-stimulated ATPase activity, helicase activity and it could not support replisome function. A359T bound ssDNA with similar efficiency as wt TWINKLE, showed only modest reductions in ATPase and helicase activities, and was able to support DNA synthesis, although at a slower rate.

The three-dimensional structure of TWINKLE has not been determined experimentally, so in order to better understand the molecular phenotypes of the different AdPEO mutations, we constructed a molecular model of TWINKLE, by homology modeling with T7 gp4 primase-helicase, for which the crystal structure has been solved (Toth et al, 2003). From the model, it is apparent that the TWINKLE linker region has two helical regions connected by a short turn. Mutations in the first helix may affect protein folding and thereby stability of the monomer. The second helix may instead be important for TWINKLE dimerization, and hence, hexamerization. In support of this explanation, the two mutations that disrupted multimerization, I367T and R374Q, are both present in the second helix. The unique structure of a proline residue does not fit well in alpha helical structures, and an R to P transition at position 354 in the first helical region may thus disrupt the structure and make the protein unstable. This may explain why we were unable to express the protein in soluble form. The structural model of TWINKLE suggests that L381 interacts specifically with residues in the conserved H1a motif of the neighboring subunit. The H1a motif is known to be involved in ATP binding, and we therefore believe that the L381P mutant may cause a defect in ATP binding or hydrolysis, explaining the total loss of all enzymatic activities tested.

From the results obtained in this paper it is evident that the vast knowledge that has been acquired about T7 gp4 can be used to explain and predict molecular phenotypes of mutations in the TWINKLE protein.

### Paper III: The N-terminal domain of TWINKLE contributes to single-stranded DNA binding and DNA helicase activities

In this study we investigated the molecular role of the N-terminal domain of human TWINKLE. This region shares some homology with the T7 gp4 primase, but lacks a functional zinc-finger motif and several amino acids essential for catalytic primase activity in gp4. To date, there have been no reports of primase activity in human TWINKLE, and bioinformatic studies suggest that TWINKLE is an active primase in most eukaryotes, but that this activity has been lost in metazoans (Shutt & Gray, 2006b). There have been reports of several mutations in the N-terminal part of TWINKLE that are associated with the disorder adPEO (Shutt & Gray, 2006b), clearly indicative of a functional importance of this region.

To address the role of the N-terminal region of TWINKLE we expressed and purified a series of deleted versions of the protein, lacking a part of (Δ1-121), or the entire N-terminus (Δ1-314), or lacking the C-terminus (Δ372-684). We also addressed the role of TWINKY, a smaller splice variant of TWINKLE lacking a part of the C-terminal helicase domain (Spelbrink et al, 2001).

Size exclusion chromatography revealed that TWINKLE, in contrast to T7 gp4, forms stable hexamers even in the absence of $Mg^{2+}$ or ATP. The C-terminal part of TWINKLE was absolutely required for oligomerization, but the N-terminal domain also contributed somewhat to this activity, as the N-terminally truncated proteins formed slightly more unstable hexamers.

In a gel-shift assay we showed that TWINKLE binds preferably to dsDNA. This is in contrast to T7 gp4 that binds tightly and preferably to ssDNA (Hingorani & Patel, 1993). The C-terminal helicase part was shown to be essential for DNA binding, but the N-terminal domain specifically contributed to the ssDNA binding activity.

TWINKLE was further shown to have a high intrinsic ability to hydrolyze ATP, and this activity was stimulated about 2-fold in the presence of ssDNA. TWINKLE is in this aspect distinct from T7 gp4 and other hexameric helicases which are usually stimulated a 100-fold in the presence of ssDNA (Patel & Picha, 2000; Washington et al, 1996). The N-terminal region of TWINKLE was not absolutely required for ATP hydrolysis; the Δ1-121 deletion mutant showed a mild reduction in ssDNA-dependant ATP hydrolysis, whereas the Δ1-314 deletion resulted in a more severe phenotype.

The ability of the wt and truncated proteins to unwind dsDNA was found to be proportional to the activities of ssDNA binding and DNA-dependent ATP hydrolysis.

The proteins ability to support rolling circle replication together with the POLγ holoenzyme and mtSSB on a double-stranded template was also assayed. Δ1-121 could support DNA synthesis at almost wt levels, but Δ1-314 could only support synthesis of shorter products,

indicating that the N-terminal domain of TWINKLE is required for full mtDNA replisome processivity. These findings are also consistent with studies of a related mutant, that when overexpressed in cell lines give rise to reduction of mtDNA copy number and stalling of the mtDNA replication fork (Wanrooij et al, 2007). This phenotype could be explained by reduced processivity of the mtDNA replisome.

Based on our findings that TWINKLE binds both ssDNA and dsDNA, but N-terminal truncations mainly affect the ssDNA binding ability, we propose that the dsDNA binding is localized to the C-terminal helicase domain, whereas ssDNA binding requires a larger protein surface, involving both the N-and C-terminal part of TWINKLE. These suggestions are supported by structural studies of the related E. coli primase DnaG, which reveal a basic cleft in the protein that is proposed to interact with the backbone of the ssDNA template (Keck et al, 2000; Podobnik et al, 2000). Mutational studies have also revealed catalytically essential basic residues in the corresponding region of T7 gp4 protein (Lee & Richardson, 2001). Molecular modeling of TWINKLE also suggests that the positively charged ssDNA-binding surface is conserved (Holmlund et al, 2009) (Paper IV).

The physiological role for the dsDNA binding ability of TWINKLE is so far unknown. The heptameric isoform of T7 gp4 appeared suitable to accommodate dsDNA in its central channel (Toth et al, 2003) and the bacterial helicase DnaB, as well as T7 gp4, have been shown to translocate actively on dsDNA, suggesting that this activity is universal among the replicative hexameric helicases. The dsDNA-translocation activity has been proposed to serve a function in driving branch migration of a Holliday junction (Kaplan, 2000; Kaplan & O'Donnell, 2002). No translocation of TWINKLE on dsDNA has yet been reported.

### Paper IV: Structure-function defects of the TWINKLE amino-terminal region in progressive external ophtalmoplegia

In this study we investigated the molecular phenotypes of four naturally occurring AdPEO-causing mutations, W315L, K319T, R334Q and P335L (Agostino et al, 2003; Deschauer et al, 2003; Lewis et al, 2002; Spelbrink et al, 2001), located in the N-terminal domain of TWINKLE. The exact function of this domain is not understood in detail, but we have previously shown that it is required for efficient binding to ssDNA and full replisome processivity (Farge et al, 2008) (Paper III).

The recombinant mutant proteins were expressed and purified, and their enzymatic properties were subsequently tested. The mutations did not affect hexamerization, but did however cause dramatic effects on ssDNA binding, ATP hydrolysis, helicase activity, and ability to support DNA replication on a mini-circle template. The most striking phenotype of the mutants was a dramatic decrease in ATPase activity, which could be partially overcome in the presence of ssDNA; K319T and R334Q reached up to almost wild-type levels. These defects in ATP hydrolysis were surprising, since the motifs known to be required for ATP binding and hydrolysis are located in the helicase domain (Spelbrink et al, 2001). The strong ssDNA-stimulation of ATPase activity was also unexpected since most of the mutants had a poor ssDNA-binding capacity.

To better understand the molecular phenotypes observed for the TWINKLE mutants, we used the previously published homology model of TWINKLE, based on the T7 gp4 crystal structure (Korhonen et al, 2008) (Paper II). The electrostatic surface features of the protein shows that the N-terminal domain contains a conserved region with significant electropositive potential, which in structurally related gp4 and *E. coli* primase DnaG is proposed to interact with the backbone of the ssDNA template (Kato et al, 2003; Keck et al, 2000; Podobnik et al, 2000). The residues W315, K319, R334 and P335 are all located in this region and may thus bind ssDNA during mtDNA replication. In support of this idea, all the mutants, except R334Q, were defective in ssDNA binding. Because of the low homology between TWINKLE and gp4 in the N-terminal region, it is difficult to draw specific conclusions about the different mutations. However, we were able to speculate about the structural effects that the mutations may cause.

Based on our biochemical data in combination with the molecular model of TWINKLE, we propose a model in which the AdPEO-causing mutations in the N-terminal domain impair the interplay between ssDNA-binding and ATP hydrolysis, an essential element of the catalytic cycle of related hexameric helicases. The structural model of TWINKLE suggests specific interactions between amino acids in the linker region and residues in the conserved H1a motif that is involved in ATP binding. In addition, the model displays a complex web of specific interactions between residues in the linker region and the N-terminal ssDNA-binding region. Mutations in the ssDNA-binding region may thus directly influence the structure of the linker region, and in turn affect ATP-binding and hydrolysis. Addition of ssDNA could counteract the negative structural impact of such mutations and stabilize the active conformation of the hexamer.

Our observations give further support to our hypothesis in a previous report (Farge et al, 2008) (Paper III) that this region in TWINKLE is important for binding to ssDNA.

# Concluding remarks

The work presented in this thesis provides a deeper understanding of the basic molecular mechanisms of mtDNA replication. By using an *in vitro* replication system we have been able to investigate different aspects of the accessory B subunit of the POLγ holoenzyme and in particular the mitochondrial replicative helicase TWINKLE.

We showed that POLγB efficiently blocks the exonuclease activity of POLγA. In addition, the dsDNA-binding activity of POLγB was demonstrated to be required for the TWINKLE-dependant stimulation of the POLγ holoenzyme. Attempts to detect direct physical interactions between these proteins have been unsuccessful, but we have clearly demonstrated the absolute interdependence of these replication factors for replisome function *in vitro*. Also, mutations in TWINKLE, as well as POLγA and POLγB, have been reported to cause the same disorder, adPEO (Longley et al, 2006; Spelbrink et al, 2001).

Because of the homology with the T7 gp4 helicase-primase, it is tempting to speculate that the N-terminal domain of TWINKLE contains the long-sought primase activity in human mitochondria. However, previous efforts to detect such an activity have remained fruitless (unpublished observations), and bioinformatic analyses suggest that although TWINKLE appears to be an active primase in most eukaryotes, this activity has been lost in metazoans (Shutt & Gray, 2006b). The functional role of the TWINKLE N-terminus has thus remained elusive. In this work we demonstrated that this domain contributes to ssDNA-binding and helicase activities of TWINKLE, and is ultimately required for full replisome activity. The primase domains of gp4 and structurally related DnaG primase of *E. coli* contain an elongated region with significant electropositive potential, suggested to interact with ssDNA (Kato et al, 2003; Keck et al, 2000). Our molecular model of TWINKLE indicates that this region is conserved, thus supporting our biochemical observations, and suggesting that the TWINKLE N-terminus is indeed involved in ssDNA-binding.

To investigate the molecular mechanisms behind the human disorder adPEO, we performed a detailed biochemical analysis on eleven different adPEO-causing TWINKLE mutations, seven in the linker-region and four in the N-terminal domain. Distinct molecular phenotypes were observed, with individual consequences for multimerization, ATPase activity, helicase activity and ability to support mtDNA synthesis *in vitro*. It became apparent that even relatively mild changes in TWINKLE activity may disturb the coordination with POLγ at the replication fork. The different molecular phenotypes could be interpreted using a structural model of TWINKLE that was based on homology with the T7 gp4 protein. These studies provided a molecular understanding of adPEO and demonstrated that knowledge about the bacteriophage T7 replication machinery, despite clear differences, may be used to understand the molecular phenotypes of mutations in the human mtDNA replication apparatus.

In a recent study it was found that addition of POLRMT to the minimal mitochondrial replisome leads to simultaneous leading- and lagging-strand DNA synthesis *in vitro*, suggesting that this polymerase is the lagging-strand primase in mammalian mitochondria (Wanrooij et al, 2008). A subsequent report, using an $O_L$-containing template, demonstrated $O_L$ as a preferred initiation site for POLRMT-dependent primer synthesis (Fuste et al, 2009). In future research it would be interesting to explore the action of the full replisome

(including POLRMT) on dsDNA templates containing all the critical sequence elements for mtDNA replication; $O_H$, CSBII, TAS and $O_L$. These experiments may help to clarify the mode of mammalian mtDNA replication.

TWINKLE could be characterized further to investigate other aspects of the protein: What is the role of the dsDNA-binding activity of TWINKLE? Can TWINKLE translocate on dsDNA, and if so, could it be involved in DNA recombination or repair? How is TWINKLE loaded on to the closed circular mtDNA molecule that would be the substrate *in vivo*? Does it use a ring-opening mechanism as T7 gp4 (Ahnert et al, 2000) or does it need additional factors to assist the loading process? Furthermore, it would be interesting to use the *in vitro* system to search for additional factors in mitochondrial extracts that could stimulate or repress mtDNA replication.

# Acknowledgements

The four years since I started my PhD-training has been a great experience, both scientifically and personally. I would like to take this opportunity to thank a number of people who made this thesis work possible.

First of all I would like to thank my supervisor **Maria Falkenberg, Mia**, for accepting me as your PhD-student. You have been an excellent supervisor! Thank you for believing in me and for your tremendous support. You were nearly always present in the lab which made it easy to discuss results and experiments on a daily basis, something I very much appreciate. You were never too busy to take a moment to chat, whether it was about science or personal matters. Your never-ending enthusiasm gave me fresh motivation during dark moments of science. Thank you for pushing me forward so I could finish in such a short time.

My co-supervisor, **Claes Gustafsson**… We haven't had that many scientific discussions together, but I think you may have been more involved in my work than I have realized; I believe you and Mia do not cease to discuss science outside of office hours… Your enthusiasm and optimistic attitude has been very inspiring over the years. Thank you for your constructive ideas and comments when reading through this thesis and the papers before printing.

My co-supervisor, **Nils-Göran Larsson**; your dedication to science has been a great inspiration for me! Thank you for your constructive criticism and for your many excellent lectures!

All my past and present colleagues/friends at the Division of Metabolic Diseases, it has been great to meet and get to know so many nice and friendly people. I hope our paths will cross again in the future.

A special thanks goes to **Jenny Wihlén** (formerly **Korhonen**), my closest friend and collaborator in the lab. You were very helpful and teaching me a lot of technical skills when I was new in the lab, this I am very grateful of. It has been fun to get to know you personally, and I hope we will stay in touch even though you will be moving up north soon.

The lovely **Géraldine**… It has been great working with you over the years! Thank you for all your advice and support. Our collaboration eventually resulted in two very nice papers!

**Mina**, thank you for all your help during the last few months when I have been preparing my thesis defense! You have given me a lot of helpful advice and I very much appreciate you taking the time to read through my thesis. We have had some nice chats, and you helped creating a lively atmosphere in the lab.

Other members of the lab; **Martina, Yonghong, Jordi, Nelli, Sjoerd** and **Xavi**, thank you for all your scientific advice and support. You are all very nice people and I am glad I had the opportunity to be a part of your lab.

All the people in **Claes** lab; **Olga, Vera, Julia, Olivera, Tomas, Zsolt, Xuefeng, Paulina**, thank you for your help with practical matters, finding things in the lab and help with the PCR-

machines! And **Fariba**, your calm and friendly ways always made me happy, and I have missed you since you left.

All the people in the large office; **Yolanda, Mina, Ivana, Ana, Mügen, Yunjian, Marjan, Christos, Metodi, Marie** and **Claire**! It has been fun to sit in the same office with you all, listening to the lively discussions taking place in there.

**Jim** and **Zekiye**, thank you for helping me out with sequencing, and **Stina, Venus** and **Lene** for help with administrative matters. **Fredrik Tingstedt**, thank you for dealing with my computer problems during the time of writing this thesis.

**Emmanouella, Chris, Anna, Daniel, Fredrik, Camilla, Mia, Benedetta, Marjan and others**, taking part in often very lively lunch room discussions. The lunch breaks have always been a lot of fun with you there.

My parents, **Carin** and **Bertil**, and my sister **Helena**! Thank you for always believing in me and for your unconditional love and encouragement.

Last but not least, my own little family, my husband **Fredrik** and my son **Hugo**! I love you both so very much! Having a child in the middle of my PhD-studies of course changed my life in a fantastic way, and soon we will have another little member added to our family, which is very exciting. Fredrik, thank you for always listening to my disappointments and frustrations, for taking care of my computer so I could write this thesis, and for taking care of our cars (!!!) so I could get to work… Your love and support means the world to me.

# References


Agostino A, Valletta L, Chinnery PF, Ferrari G, Carrara F, Taylor RW, Schaefer AM, Turnbull DM, Tiranti V, Zeviani M (2003) Mutations of ANT1, Twinkle, and POLG1 in sporadic progressive external ophthalmoplegia (PEO). *Neurology* **60**(8)**:** 1354-1356

Ahnert P, Picha KM, Patel SS (2000) A ring-opening mechanism for DNA binding in the central channel of the T7 helicase-primase protein. *EMBO J* **19**(13)**:** 3418-3427

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* **290**(5806)**:** 457-465

Andersson SG, Karlberg O, Canback B, Kurland CG (2003) On the origin of mitochondria: a genomics perspective. *Philos Trans R Soc Lond B Biol Sci* **358**(1429)**:** 165-177; discussion 177-169

Arenas J, Briem E, Dahl H, Hutchison W, Lewis S, Martin MA, Spelbrink H, Tiranti V, Jacobs H, Zeviani M (2003) The V368i mutation in Twinkle does not segregate with AdPEO. *Ann Neurol* **53**(2)**:** 278

Benkovic SJ, Valentine AM, Salinas F (2001) Replisome-mediated DNA replication. *Annu Rev Biochem* **70:** 181-208

Bogenhagen D, Clayton DA (1977) Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. *Cell* **11**(4)**:** 719-727

Bogenhagen DF, Clayton DA (2003a) Concluding remarks: The mitochondrial DNA replication bubble has not burst. *Trends Biochem Sci* **28**(8)**:** 404-405

Bogenhagen DF, Clayton DA (2003b) The mitochondrial DNA replication bubble has not burst. *Trends Biochem Sci* **28**(7)**:** 357-360

Bogenhagen DF, Rousseau D, Burke S (2008) The layered structure of human mitochondrial DNA nucleoids. *J Biol Chem* **283**(6)**:** 3665-3675

Bowmaker M, Yang MY, Yasukawa T, Reyes A, Jacobs HT, Huberman JA, Holt IJ (2003) Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone. *J Biol Chem* **278**(51)**:** 50961-50969

Brown TA, Cecconi C, Tkachuk AN, Bustamante C, Clayton DA (2005) Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes Dev* **19**(20)**:** 2466-2476

Brown TA, Tkachuk AN, Clayton DA (2008) Native R-loops persist throughout the mouse mitochondrial DNA genome. *J Biol Chem* **283**(52)**:** 36743-36751

Camasamudram V, Fang JK, Avadhani NG (2003) Transcription termination at the mouse mitochondrial H-strand promoter distal site requires an A/T rich sequence motif and sequence specific DNA binding proteins. *Eur J Biochem* **270**(6)**:** 1128-1140

Carrodeguas JA, Bogenhagen DF (2000) Protein sequences conserved in prokaryotic aminoacyl-tRNA synthetases are important for the activity of the processivity factor of human mitochondrial DNA polymerase. *Nucleic Acids Res* **28**(5)**:** 1237-1244

Carrodeguas JA, Kobayashi R, Lim SE, Copeland WC, Bogenhagen DF (1999) The accessory subunit of Xenopus laevis mitochondrial DNA polymerase gamma increases processivity of the catalytic subunit of human DNA polymerase gamma and is related to class II aminoacyl-tRNA synthetases. *Mol Cell Biol* **19**(6)**:** 4039-4046

Carrodeguas JA, Pinz KG, Bogenhagen DF (2002) DNA binding properties of human pol gammaB. *J Biol Chem* **277**(51)**:** 50008-50014

Carrodeguas JA, Theis K, Bogenhagen DF, 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**(1)**:** 43-54

Chen XJ, Butow RA (2005) The organization and inheritance of the mitochondrial genome. *Nat Rev Genet* **6**(11)**:** 815-825

Clayton DA (1982) Replication of animal mitochondrial DNA. *Cell* **28**(4)**:** 693-705

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

Cotney J, Shadel GS (2006) Evidence for an early gene duplication event in the evolution of the mitochondrial transcription factor B family and maintenance of rRNA methyltransferase activity in human mtTFB1 and mtTFB2. *J Mol Evol* **63**(5)**:** 707-717

Cottrell DA, Blakely EL, Borthwick GM, Johnson MA, Taylor GA, Brierley EJ, Ince PG, Turnbull DM (2000) Role of mitochondrial DNA mutations in disease and aging. *Ann N Y Acad Sci* **908:** 199-207

Cree LM, Samuels DC, de Sousa Lopes SC, Rajasimha HK, Wonnapinij P, Mann JR, Dahl HH, Chinnery PF (2008) A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nat Genet* **40**(2)**:** 249-254

Deschauer M, Kiefer R, Blakely EL, He L, Zierz S, Turnbull DM, Taylor RW (2003) A novel Twinkle gene mutation in autosomal dominant progressive external ophthalmoplegia. *Neuromuscul Disord* **13**(7-8)**:** 568-572

Desjardins P, Morais R (1990) Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. *J Mol Biol* **212**(4)**:** 599-634

Dimauro S, Davidzon G (2005) Mitochondrial DNA and disease. *Ann Med* **37**(3)**:** 222-232

Ekstrand MI, Falkenberg M, Rantanen A, Park CB, Gaspari M, Hultenby K, Rustin P, Gustafsson CM, Larsson NG (2004) Mitochondrial transcription factor A regulates mtDNA copy number in mammals. *Hum Mol Genet* **13**(9)**:** 935-944

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

Falkenberg M, Larsson NG, Gustafsson CM (2007) DNA replication and transcription in mammalian mitochondria. *Annu Rev Biochem* **76:** 679-699

Fan L, Sanschagrin PC, Kaguni LS, Kuhn LA (1999) The accessory subunit of mtDNA polymerase shares structural homology with aminoacyl-tRNA synthetases: implications for a dual role as a primer recognition factor and processivity clamp. *Proc Natl Acad Sci U S A* **96**(17)**:** 9527-9532

Farge G, Holmlund T, Khvorostova J, Rofougaran R, Hofer A, Falkenberg M (2008) The N-terminal domain of TWINKLE contributes to single-stranded DNA binding and DNA helicase activities. *Nucleic Acids Res* **36**(2)**:** 393-403

Fernandez-Silva P, Enriquez JA, Montoya J (2003) Replication and transcription of mammalian mitochondrial DNA. *Exp Physiol* **88**(1)**:** 41-56

Fisher RP, Lisowsky T, Parisi MA, Clayton DA (1992) DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. *J Biol Chem* **267**(5)**:** 3358-3367

Fridlender B, Fry M, Bolden A, Weissbach A (1972) A new synthetic RNA-dependent DNA polymerase from human tissue culture cells (HeLa-fibroblast-synthetic oligonucleotides-template-purified enzymes). *Proc Natl Acad Sci U S A* **69**(2)**:** 452-455

Fuste JM, Wanrooij S, Jemt E, Granycome CE, Cluett TJ, Shi Y, Atanassova N, Holt IJ, Gustafsson CM, Falkenberg M (2009) Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. *Manuscript*

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

Graff C, Clayton DA, Larsson NG (1999) Mitochondrial medicine--recent advances. *J Intern Med* **246**(1)**:** 11-23

Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. *Science* **283**(5407)**:** 1476-1481

Guo S, Tabor S, Richardson CC (1999) The linker region between the helicase and primase domains of the bacteriophage T7 gene 4 protein is critical for hexamer formation. *J Biol Chem* **274**(42)**:** 30303-30309

Habib SJ, Neupert W, Rapaport D (2007) Analysis and prediction of mitochondrial targeting signals. *Methods Cell Biol* **80:** 761-781

Hingorani MM, Patel SS (1993) Interactions of bacteriophage T7 DNA primase/helicase protein with single-stranded and double-stranded DNAs. *Biochemistry* **32**(46)**:** 12478-12487

Holmlund T, Farge G, Pande V, Korhonen J, Nilsson L, Falkenberg M (2009) Structure-function defects of the twinkle amino-terminal region in progressive external ophthalmoplegia. *Biochim Biophys Acta* **1792**(2)**:** 132-139

Holt IJ, Harding AE, Morgan-Hughes JA (1988) Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature* **331**(6158)**:** 717-719

Holt IJ, Jacobs HT (2003) Response: The mitochondrial DNA replication bubble has not burst. *Trends Biochem Sci* **28**(7)**:** 355-356

Holt IJ, Lorimer HE, Jacobs HT (2000) Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell* **100**(5)**:** 515-524

Jendrach M, Pohl S, Voth M, Kowald A, Hammerstein P, Bereiter-Hahn J (2005) Morpho-dynamic changes of mitochondria during ageing of human endothelial cells. *Mech Ageing Dev* **126**(6-7)**:** 813-821

Johnson DE, Takahashi M, Hamdan SM, Lee SJ, Richardson CC (2007) Exchange of DNA polymerases at the replication fork of bacteriophage T7. *Proc Natl Acad Sci U S A* **104**(13)**:** 5312-5317

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

Kaplan DL (2000) The 3'-tail of a forked-duplex sterically determines whether one or two DNA strands pass through the central channel of a replication-fork helicase. *J Mol Biol* **301**(2)**:** 285-299

Kaplan DL, O'Donnell M (2002) DnaB drives DNA branch migration and dislodges proteins while encircling two DNA strands. *Mol Cell* **10**(3)**:** 647-657

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

Kaufman BA, Durisic N, Mativetsky JM, Costantino S, Hancock MA, Grutter P, Shoubridge EA (2007) The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Mol Biol Cell* **18**(9)**:** 3225-3236

Keck JL, Roche DD, Lynch AS, Berger JM (2000) Structure of the RNA polymerase domain of E. coli primase. *Science* **287**(5462)**:** 2482-2486

Kiechl S, Horvath R, Luoma P, Kiechl-Kohlendorfer U, Wallacher-Scholz B, Stucka R, Thaler C, Wanschitz J, Suomalainen A, Jaksch M, Willeit J (2004) Two families with autosomal dominant progressive external ophthalmoplegia. *J Neurol Neurosurg Psychiatry* **75**(8)**:** 1125-1128

Knott AB, Bossy-Wetzel E (2008) Impairing the mitochondrial fission and fusion balance: a new mechanism of neurodegeneration. *Ann N Y Acad Sci* **1147:** 283-292

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

Korhonen JA, Pande V, Holmlund T, Farge G, Pham XH, Nilsson L, Falkenberg M (2008) Structure-function defects of the TWINKLE linker region in progressive external ophthalmoplegia. *J Mol Biol* **377**(3)**:** 691-705

Korhonen JA, Pham XH, Pellegrini M, Falkenberg M (2004) Reconstitution of a minimal mtDNA replisome in vitro. *EMBO J* **23**(12)**:** 2423-2429

Krause F, Reifschneider NH, Vocke D, Seelert H, Rexroth S, Dencher NA (2004) "Respirasome"-like supercomplexes in green leaf mitochondria of spinach. *J Biol Chem* **279**(46)**:** 48369-48375

Kucej M, Butow RA (2007) Evolutionary tinkering with mitochondrial nucleoids. *Trends Cell Biol* **17**(12)**:** 586-592

Kujoth GC, Hiona A, Pugh TD, Someya S, Panzer K, Wohlgemuth SE, Hofer T, Seo AY, Sullivan R, Jobling WA, Morrow JD, Van Remmen H, Sedivy JM, Yamasoba T, Tanokura M, Weindruch R, Leeuwenburgh C, Prolla TA (2005) Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science* **309**(5733)**:** 481-484

Kusakabe T, Richardson CC (1996) The role of the zinc motif in sequence recognition by DNA primases. *J Biol Chem* **271**(32)**:** 19563-19570

Larsson NG, Clayton DA (1995) Molecular genetic aspects of human mitochondrial disorders. *Annu Rev Genet* **29:** 151-178

Larsson NG, Wang J, Wilhelmsson H, Oldfors A, Rustin P, Lewandoski M, Barsh GS, Clayton DA (1998) Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nat Genet* **18**(3)**:** 231-236

Lecrenier N, Van Der Bruggen P, Foury F (1997) Mitochondrial DNA polymerases from yeast to man: a new family of polymerases. *Gene* **185**(1)**:** 147-152

Lee SJ, Richardson CC (2001) Essential lysine residues in the RNA polymerase domain of the gene 4 primase-helicase of bacteriophage T7. *J Biol Chem* **276**(52)**:** 49419-49426

Lee SJ, Richardson CC (2002) Interaction of adjacent primase domains within the hexameric gene 4 helicase-primase of bacteriophage T7. *Proc Natl Acad Sci U S A* **99**(20)**:** 12703-12708

Lee SJ, Richardson CC (2004) The linker region between the helicase and primase domains of the gene 4 protein of bacteriophage T7. Role in helicase conformation and activity. *J Biol Chem* **279**(22)**:** 23384-23393

Lewis S, Hutchison W, Thyagarajan D, Dahl HH (2002) Clinical and molecular features of adPEO due to mutations in the Twinkle gene. *J Neurol Sci* **201**(1-2)**:** 39-44

Linder T, Park CB, Asin-Cayuela J, Pellegrini M, Larsson NG, Falkenberg M, Samuelsson T, Gustafsson CM (2005) A family of putative transcription termination factors shared amongst metazoans and plants. *Curr Genet* **48**(4)**:** 265-269

Longley MJ, Clark S, Yu Wai Man C, Hudson G, Durham SE, Taylor RW, Nightingale S, Turnbull DM, Copeland WC, Chinnery PF (2006) Mutant POLG2 disrupts DNA polymerase gamma subunits and causes progressive external ophthalmoplegia. *Am J Hum Genet* **78**(6)**:** 1026-1034

Luft R, Ikkos D, Palmieri G, Ernster L, Afzelius B (1962) A case of severe hypermetabolism of nonthyroid origin with a defect in the maintenance of mitochondrial respiratory control: a correlated clinical, biochemical, and morphological study. *J Clin Invest* **41:** 1776-1804

Luo N, Kaguni LS (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**(4)**:** 2491-2497

Luoma PT, Luo N, Loscher WN, Farr CL, Horvath R, Wanschitz J, Kiechl S, Kaguni LS, Suomalainen A (2005) Functional defects due to spacer-region mutations of human mitochondrial DNA polymerase in a family with an ataxia-myopathy syndrome. *Hum Mol Genet* **14**(14)**:** 1907-1920

Mackintosh SG, Raney KD (2006) DNA unwinding and protein displacement by superfamily 1 and superfamily 2 helicases. *Nucleic Acids Res* **34**(15)**:** 4154-4159

Madsen CS, Ghivizzani SC, Hauswirth WW (1993) Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. *Mol Cell Biol* **13**(4)**:** 2162-2171

McCulloch V, Seidel-Rogol BL, Shadel GS (2002) A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine. *Mol Cell Biol* **22**(4)**:** 1116-1125

McCulloch V, Shadel GS (2003) Human mitochondrial transcription factor B1 interacts with the C-terminal activation region of h-mtTFA and stimulates transcription independently of its RNA methyltransferase activity. *Mol Cell Biol* **23**(16)**:** 5816-5824

Metodiev MD, Lesko N, Park CB, Camara Y, Shi Y, Wibom R, Hultenby K, Gustafsson CM, Larsson NG (2009) Methylation of 12S rRNA is necessary for in vivo stability of the small subunit of the mammalian mitochondrial ribosome. *Cell Metab* **9**(4)**:** 386-397

Montoya J, Christianson T, Levens D, Rabinowitz M, 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**(23)**:** 7195-7199

Montoya J, Gaines GL, Attardi G (1983) The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. *Cell* **34**(1)**:** 151-159

Moraes CT, Kenyon L, Hao H (1999) Mechanisms of human mitochondrial DNA maintenance: the determining role of primary sequence and length over function. *Mol Biol Cell* **10**(10)**:** 3345-3356

Moslemi AR, Melberg A, Holme E, Oldfors A (1996) Clonal expansion of mitochondrial DNA with multiple deletions in autosomal dominant progressive external ophthalmoplegia. *Ann Neurol* **40**(5)**:** 707-713

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

Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290**(5806)**:** 470-474

Park CB, Asin-Cayuela J, Camara Y, Shi Y, Pellegrini M, Gaspari M, Wibom R, Hultenby K, Erdjument-Bromage H, Tempst P, Falkenberg M, Gustafsson CM, Larsson NG (2007) MTERF3 is a negative regulator of mammalian mtDNA transcription. *Cell* **130**(2)**:** 273-285

Patel SS, Picha KM (2000) Structure and function of hexameric helicases. *Annu Rev Biochem* **69:** 651-697

Pellegrini M, Asin-Cayuela J, Erdjument-Bromage H, Tempst P, Larsson NG, Gustafsson CM (2009) MTERF2 is a nucleoid component in mammalian mitochondria. *Biochim Biophys Acta* **1787**(5)**:** 296-302

Pham XH, Farge G, Shi Y, Gaspari M, Gustafsson CM, Falkenberg M (2006) Conserved sequence box II directs transcription termination and primer formation in mitochondria. *J Biol Chem* **281**(34)**:** 24647-24652

Podobnik M, McInerney P, O'Donnell M, Kuriyan J (2000) A TOPRIM domain in the crystal structure of the catalytic core of Escherichia coli primase confirms a structural link to DNA topoisomerases. *J Mol Biol* **300**(2)**:** 353-362

Pomerantz RT, O'Donnell M (2007) Replisome mechanics: insights into a twin DNA polymerase machine. *Trends Microbiol* **15**(4)**:** 156-164

Richardson CC (1983) Bacteriophage T7: minimal requirements for the replication of a duplex DNA molecule. *Cell* **33**(2)**:** 315-317

Saraste M (1999) Oxidative phosphorylation at the fin de siecle. *Science* **283**(5407)**:** 1488-1493

Sato A, Nakada K, Hayashi J (2006) Mitochondrial dynamics and aging: Mitochondrial interaction preventing individuals from expression of respiratory deficiency caused by mutant mtDNA. *Biochim Biophys Acta* **1763**(5-6)**:** 473-481

Sawaya MR, Guo S, Tabor S, Richardson CC, Ellenberger T (1999) Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell* **99**(2)**:** 167-177

Schaefer AM, Taylor RW, Turnbull DM, Chinnery PF (2004) The epidemiology of mitochondrial disorders--past, present and future. *Biochim Biophys Acta* **1659**(2-3)**:** 115-120

Schafer E, Seelert H, Reifschneider NH, Krause F, Dencher NA, Vonck J (2006) Architecture of active mammalian respiratory chain supercomplexes. *J Biol Chem* **281**(22)**:** 15370-15375

Schagger H, Pfeiffer K (2000) Supercomplexes in the respiratory chains of yeast and mammalian mitochondria. *EMBO J* **19**(8)**:** 1777-1783

Schapira AH (2006) Mitochondrial disease. *Lancet* **368**(9529)**:** 70-82

Seidel-Rogol BL, McCulloch V, Shadel GS (2003) Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. *Nat Genet* **33**(1)**:** 23-24

Shadel GS, Clayton DA (1997) Mitochondrial DNA maintenance in vertebrates. *Annu Rev Biochem* **66:** 409-435

Shang J, Clayton DA (1994) Human mitochondrial transcription termination exhibits RNA polymerase independence and biased bipolarity in vitro. *J Biol Chem* **269**(46)**:** 29112-29120

Shults CW, Oakes D, Kieburtz K, Beal MF, Haas R, Plumb S, Juncos JL, Nutt J, Shoulson I, Carter J, Kompoliti K, Perlmutter JS, Reich S, Stern M, Watts RL, Kurlan R, Molho E, Harrison M, Lew M (2002) Effects of coenzyme Q10 in early Parkinson disease: evidence of slowing of the functional decline. *Arch Neurol* **59**(10)**:** 1541-1550

Shutt TE, Gray MW (2006a) Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends Genet* **22**(2)**:** 90-95

Shutt TE, Gray MW (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**(5)**:** 588-599

Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* **76:** 23-50

Sinha NK, Morris CF, Alberts BM (1980) Efficient in vitro replication of double-stranded DNA templates by a purified T4 bacteriophage replication system. *J Biol Chem* **255**(9)**:** 4290-4293

Skulachev VP (2001) Mitochondrial filaments and clusters as intracellular power-transmitting cables. *Trends Biochem Sci* **26**(1)**:** 23-29

Spelbrink JN, Li FY, Tiranti V, Nikali K, Yuan QP, Tariq M, Wanrooij S, Garrido N, Comi G, Morandi L, Santoro L, Toscano A, Fabrizi GM, Somer H, Croxen R, Beeson D, Poulton J, Suomalainen A, Jacobs HT, Zeviani M, Larsson C (2001) Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. *Nat Genet* **28**(3)**:** 223-231

Stewart JB, Freyer C, Elson JL, Wredenberg A, Cansu Z, Trifunovic A, Larsson NG (2008) Strong purifying selection in transmission of mammalian mitochondrial DNA. *PLoS Biol* **6**(1)**:** e10

Suomalainen A, Kaukonen J (2001) Diseases caused by nuclear genes affecting mtDNA stability. *Am J Med Genet* **106**(1)**:** 53-61

Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C, Schatten G (2000) Ubiquitinated sperm mitochondria, selective proteolysis, and the regulation of mitochondrial inheritance in mammalian embryos. *Biol Reprod* **63**(2)**:** 582-590

Taanman JW (1999) The mitochondrial genome: structure, transcription, translation and replication. *Biochim Biophys Acta* **1410**(2)**:** 103-123

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

Tiranti V, Savoia A, Forti F, D'Apolito MF, Centra M, Rocchi M, Zeviani M (1997) Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the Expressed Sequence Tags database. *Hum Mol Genet* **6**(4)**:** 615-625

Toth EA, Li Y, Sawaya MR, Cheng Y, Ellenberger T (2003) The crystal structure of the bifunctional primase-helicase of bacteriophage T7. *Mol Cell* **12**(5)**:** 1113-1123

Trifunovic A, Wredenberg A, Falkenberg M, Spelbrink JN, Rovio AT, Bruder CE, Bohlooly YM, Gidlof S, Oldfors A, Wibom R, Tornell J, Jacobs HT, Larsson NG (2004) Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**(6990)**:** 417-423

Walberg MW, Clayton DA (1981) Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic Acids Res* **9**(20)**:** 5411-5421

Wallace DC, Singh G, Lott MT, Hodge JA, Schurr TG, Lezza AM, Elsas LJ, 2nd, Nikoskelainen EK (1988) Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science* **242**(4884)**:** 1427-1430

Van Goethem G, Dermaut B, Lofgren A, Martin JJ, Van Broeckhoven C (2001) Mutation of POLG is associated with progressive external ophthalmoplegia characterized by mtDNA deletions. *Nat Genet* **28**(3)**:** 211-212

Wang Y, Bogenhagen DF (2006) Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane. *J Biol Chem* **281**(35)**:** 25791-25802

Wang Y, Farr CL, Kaguni LS (1997) Accessory subunit of mitochondrial DNA polymerase from Drosophila embryos. Cloning, molecular analysis, and association in the native enzyme. *J Biol Chem* **272**(21)**:** 13640-13646

VanLoock MS, Chen YJ, Yu X, Patel SS, Egelman EH (2001) The primase active site is on the outside of the hexameric bacteriophage T7 gene 4 helicase-primase ring. *J Mol Biol* **311**(5)**:** 951-956

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

Wanrooij S, Goffart S, Pohjoismaki JL, Yasukawa T, Spelbrink JN (2007) Expression of catalytic mutants of the mtDNA helicase Twinkle and polymerase POLG causes distinct replication stalling phenotypes. *Nucleic Acids Res* **35**(10)**:** 3238-3251

Washington MT, Rosenberg AH, Griffin K, Studier FW, Patel SS (1996) Biochemical analysis of mutant T7 primase/helicase proteins defective in DNA binding, nucleotide hydrolysis, and the coupling of hydrolysis with DNA unwinding. *J Biol Chem* **271**(43)**:** 26825-26834

Wenz T, Luca C, Torraco A, Moraes CT (2009) mTERF2 regulates oxidative phosphorylation by modulating mtDNA transcription. *Cell Metab* **9**(6)**:** 499-511

Wong TW, Clayton DA (1985) Isolation and characterization of a DNA primase from human mitochondria. *J Biol Chem* **260**(21)**:** 11530-11535

Yaffe MP (1999) The machinery of mitochondrial inheritance and behavior. *Science* **283**(5407)**:** 1493-1497

Yang MY, Bowmaker M, Reyes A, Vergani L, Angeli P, Gringeri E, Jacobs HT, Holt IJ (2002) Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. *Cell* **111**(4)**:** 495-505

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

Yasukawa T, Yang MY, Jacobs HT, Holt IJ (2005) A bidirectional origin of replication maps to the major noncoding region of human mitochondrial DNA. *Mol Cell* **18**(6)**:** 651-662

Youle RJ, Karbowski M (2005) Mitochondrial fission in apoptosis. *Nat Rev Mol Cell Biol* **6**(8)**:** 657-663

Zeviani M, Di Donato S (2004) Mitochondrial disorders. *Brain* **127**(Pt 10)**:** 2153-2172