

Aus dem Lehrstuhl Physiologische Chemie, Biomedizinisches Centrum,  
Institut der Ludwig-Maximilians-Universität München  
Vorstand: Prof. Andreas Ladurner, PhD



***Regulation of Mitochondrial DNA Replication by MacroD1 and  
ADP-ribosylation***

Dissertation  
zum Erwerb des Doktorgrades der Naturwissenschaften  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität zu München

vorgelegt von  
Flavia Söllner

aus  
Salzgitter

Jahr  
2020

Mit Genehmigung der Medizinischen Fakultät  
der Universität München

Betreuer(in): Prof. Andreas G. Ladurner, PhD

Zweitgutachter(in): Prof. S. Meiners

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 17.09.2021

## Acknowledgements

First and foremost, I would like to acknowledge my supervisor Andreas G. Ladurner. Thank you for supporting me and standing behind me when it most mattered. You encouraged me to work with MacroD1 at a crucial point in my PhD, thereby opening the gateway to mitochondrial biology to me and with that extraordinary amounts of scientific excitement which I would have otherwise missed.

I would like to acknowledge Christiane Kotthoff. I never expected to be allowed to work so closely with someone as talented and experienced at the start of this journey – and I would like you to know that I am truly thankful for every day I was allowed to work with you. I am thankful not only for your input, for your encouragement and believe in me, but also for calling me out when you thought I could do better.

To my thesis advisory committee: Ragnhild Eskeland, Sandra Hake, Christof Osman, Tobias Straub and Bernhard Lüscher – thank you for your time and advising me throughout my PhD thesis. I am very thankful for all your input. I thank Aurelio P. Nardoza for his earlier research into the mitochondrial, MacroD1-dependend ADP-ribosylome, which served as a starting point for this thesis. Additionally, I would like to express my gratitude to each and every one of our collaborators. Carla Margulies and Maria Spletter, thank you for inducting me into the *Drosophila* world.

I would like to thank Miriam Valera-Alberna, Caroline Prell-Schicker, Christine Werner, Anton Eberharter, Lisa Richter, Charlotte Blessing, Claudia Gonzales-Leal, Tia Tyrsett Kuo, Sandra Esser and Thomas Pysik for the interesting conversations and advice I have received from you over the years. Marcel Gilbert Genge, Umut Gonsel, Moritz Resch, Dejana Mokranjac and Kai Hell, you and MitoClub have given me so much joy. I do believe I will still try to dial in to the weekly MitoClubs, you might have to block my Zoom account if you want to get rid of me.

Julia Preißer & Shao-Yen Kao: microscopy, fly-lines, pre-Yoga dinners, Yoga, Karate, ice cream, more ice cream, ice skating and for so much more – thank you. Rebecca Smith, you are the PostDoc every student should meet during their PhD. I vividly remember you coming up to me after my first departmental talk, passing me a note with some wonderful input and thoughts you had throughout, while also offering me some of your cell lines. It was the first of many times you shared invaluable insights and reagents with me. I am very happy to have the privilege of calling you my friend.

To my family; my grandparents Magdalena and Josef, Stefan and Maria, my parents Hildegard and Walter, my brother Immo, my sister Antonia, my wonderful nephews Joshi and Jonah, and my uncle Alfred. I am so lucky you are the group of people I landed in. Without you, life would be lonely. Thank you for humouring me when I need reassurance and for not getting too fed up with me when I do. Last but not least, I want to thank Benjamin – not only for suggesting you could write your own acknowledgement to ensure it sounds appropriate, but also for all the other stuff I won't list, even though you suggested I do. Adventure awaits.

## Abstract

Mammalian mitochondria, the double-membrane-bound organelles found in most eukaryotic cells contain between 100 and several 100000 copies of mitochondrial DNA (mtDNA). Human mtDNA spans 16569 base pairs and codes for 37 genes: 13 protein coding, 2 rRNAs and 22 tRNAs. Both rRNAs and tRNAs are required for the synthesis of mitochondrially encoded proteins, which are all vital subunits of the electron transport chain. Hence, the maintenance and regulation of the mitochondrial genome is essential for normal cellular function and metabolism. Transcription and replication of mtDNA is constant, simultaneous and uncoupled from the progression of the cell cycle. The proteins involved in the replication and repair of mtDNA are encoded in the nucleus and shuttled into mitochondria. Dysregulation of the mitochondrial genome is implicated in a variety of neurological diseases and mitochondrial myopathies. The aetiologies behind these diseases are often either mutations in the mitochondrial genome, or mutations of the nuclear genome sequences coding for factors involved in the maintenance of mtDNA. The regulation of replication, and thereby of copy number, can be modulated and controlled at various points; at mtDNA accessibility, at transcription initiation, and at the switch between transcription initiation and replication. The transcription-initiation machinery for mtDNA consists of: POLRMT, TFB2M and TFAM; the core replication factors are: POLG1, POLG2, POLRMT, TWINKLE, LIG3, mtSSB, and topoisomerases. How the interplay between transcription initiation, transcription and replication is organized, modulated and regulated however, is largely unknown. ADP-ribosylation is a post-translational modification (PTM) heavily involved in the maintenance of the nuclear genome, but whether it is implicated in mtDNA homeostasis has not been thoroughly elucidated.

Research into mtDNA homeostasis, replication and regulation is limited by a lack in DNA methodologies focused on the DNA within mitochondria, instead of in the nucleus. Further, very little is known about which mitochondrial factors are ADP-ribosylated and how this modification affects mitochondrial biology.

In this thesis I have i) developed and tested a novel fluorescence-based assay to monitor mitochondrial DNA replication *in cellulo* and ii) investigated ADP-ribosylation in the context of MacroD1, a mitochondrial ADP-binding protein, in relation to mtDNA homeostasis.

# Zusammenfassung

Säugetier-Mitochondrien, die doppelmembrangebundenen Organellen, die in den meisten eukaryotischen Zellen gefunden werden, enthalten zwischen hundert und mehreren hunderttausend Kopien mitochondrialer DNA (mtDNA). Die humane mtDNA umfasst 16569 Basenpaare und kodiert 37 Gene: 13 Proteinkodierungen, 2 rRNAs und 22 tRNAs. Sowohl rRNA als auch tRNA werden für die Synthese von mitochondrial kodierten Proteinen, die alle lebenswichtige Untereinheiten der Elektronentransportkette sind, benötigt. Daher ist die Aufrechterhaltung und Regulation des mitochondrialen Genoms für die normale Zellfunktion sowie den Stoffwechsel von wesentlicher Bedeutung. Die Transkription und Replikation von mtDNA ist konstant, zeitgleich und vom Fortschreiten des Zellzyklus entkoppelt. Die an der Replikation und Reparatur von mtDNA beteiligten Proteine werden im Kern kodiert und in die Mitochondrien transportiert. Eine Dysregulation des mitochondrialen Genoms ist an einer Vielzahl von neurologischen Erkrankungen und mitochondrialen Myopathien beteiligt. Die Ursachen für diese Krankheiten sind häufig entweder Mutationen im mitochondrialen Genom oder Mutationen der Kerngenomsequenzen, die für Faktoren kodieren, welche an der Aufrechterhaltung der mtDNA beteiligt sind. Die Regulierung der Replikation, und damit der Kopienzahl, kann an verschiedenen Punkten moduliert und gesteuert werden; über die Zugänglichkeit der mtDNA, bei Transkriptionsinitiierung und beim Wechsel zwischen Transkriptionsinitiierung und Replikation. Die Transkriptionsinitiationsmaschinerie für mtDNA besteht aus POLRMT, TFB2M und TFAM, die Kernreplikationsfaktoren sind POLG1, POLG2, POLRMT, TWINKLE, LIG3, mtSSB und Topoisomerasen. Wie das Zusammenspiel von Transkriptionsinitiierung, Transkription und Replikation organisiert, moduliert und reguliert wird, ist jedoch weitgehend unbekannt. Die ADP-Ribosylierung ist eine posttranslationale Modifikation (PTM), die stark an der Aufrechterhaltung des Kerngenoms beteiligt ist. Ob sie jedoch an der mtDNA-Homöostase beteiligt ist, ist noch nicht ausreichend geklärt.

Die Erforschung der mtDNA-Homöostase, -Replikation und -Regulierung ist durch einen Mangel an Methoden begrenzt, die sich auf die DNA in den Mitochondrien anstatt im Kern konzentrieren. Bislang ist sehr wenig darüber bekannt, welche mitochondrialen Faktoren ADP-ribosyliert sind und wie diese Modifikation die mitochondriale Biologie beeinflusst.

In dieser Arbeit habe ich i) ein neuartiges fluoreszenzbasiertes Verfahren zur Überwachung der mitochondrialen DNA-Replikation in cellulo entwickelt und getestet sowie ii) die ADP-Ribosylierung im Zusammenhang mit MacroD1, einem mitochondrialen ADP-bindenden Protein, in Bezug auf die mtDNA-Homöostase erforscht.



# Table of Contents

<b>Acknowledgements</b> .....	<b>3</b>
<b>Abstract</b> .....	<b>5</b>
<b>Zusammenfassung</b> .....	<b>7</b>
<b>Table of Contents</b> .....	<b>9</b>
<b>Figures &amp; Tables</b> .....	<b>13</b>
<b>1. Introduction</b> .....	<b>15</b>
<b>1.1 Mitochondria</b> .....	<b>15</b>
1.1.1 The origins of mitochondria date back approximately 2 billion years .....	15
1.1.2 The form of mitochondria have impact on their function .....	17
1.1.3 Mitochondria contain an ancient genome.....	18
1.1.4 The mitochondrial genome is distinctly organised .....	21
1.1.5 The replication and transcription of mtDNA are interconnected .....	22
1.1.6 Mitochondrial transcription and replication are regulated in numerous ways .....	29
1.1.7 There is co-regulation between the mitochondrial and the nuclear genome .....	31
1.1.8 Mitochondrial DNA maintenance is disease-relevant.....	31
<b>1.2 ADP-ribosylation</b> .....	<b>35</b>
1.2.1 How is ADP-ribosylation regulated and what is the function? .....	35
1.2.2 Which factors write, read or erase the ADP-ribosyl modification? .....	37
1.2.3 The regulation and function of mitochondrial ADP-ribosylation .....	41
<b>1.3 MacroD1</b> .....	<b>45</b>
1.3.1 What are the molecular properties of MacroD1?.....	45
1.3.2 Is MacroD1 disease-relevant? A critical view at the literature .....	46
<b>1.4 Aim of the study</b> .....	<b>48</b>
Aim I: Mitochondrial DNA - a novel fluorescence-based replication assay .....	48
Aim II: Analysis of MacroD1 interaction partners.....	50
Aim III: Biological function of MacroD1 and ADP-ribosylation in mitochondria.....	51
<b>2. Methods</b> .....	<b>52</b>
<b>2.1 Nucleic acid techniques</b> .....	<b>52</b>
2.1.1 Transformation of bacterial cells .....	52
2.1.2 DNA extraction from bacterial cells .....	52
2.1.3 DNA extraction from mammalian cells .....	53
2.1.4 DNA concentration assessment.....	54
2.1.5 DNA restriction digest, separation and purification from agarose gel .....	54
2.1.6 DNA ligation .....	55
2.1.7 Polymerase chain reaction (PCR) .....	56
2.1.8 DNA mutagenesis.....	56
2.1.9 Quantitative PCR measurements (qPCR) .....	57
2.1.10 Semi-quantitative long-amplicon PCR .....	59
2.1.11 DNA sequencing.....	61
2.1.12 Mitochondrial mutation load assay .....	61
<b>2.2 Protein and mammalian cell culture techniques</b> .....	<b>62</b>
2.2.1 Stable, controlled, expression in HEK293 -T-Rex cells .....	62
2.2.2 CRISPR/Cas9 MacroD1 knockout in HEK293 -T-Rex & U2-OS cells .....	62
2.2.3 Stable expression of proteins in MD1KO & WT U2-OS cells .....	63
2.2.4 Routine maintenance of cells.....	64
2.2.5 Counting cells.....	65
2.2.6 Transfecting cells .....	65

2.2.7 Cell-treatment with ddC .....	67
2.2.8 Cell proliferation rate analysis .....	67
2.2.9 Seahorse XF cell mito stress analysis .....	67
2.2.10 Subcellular localization - fractionation of cells .....	68
2.2.11 Submitochondrial localization: proteinase k treatment of mitochondria .....	69
2.2.12 Immunoprecipitation .....	69
2.2.13 Western blot analysis.....	70
2.2.14 Far western blotting.....	71
2.2.15 Mitochondrial DNA immunoprecipitation (MIP) assay.....	72
2.2.16 Immunofluorescent staining and confocal imaging.....	72
2.2.17 Immunofluorescent staining and flow cytometric analysis .....	73
2.2.18 Immunoprecipitation, MS sample preparation and analysis .....	74
2.2.19 Mass spectrometry data analyses.....	76
2.2.20 POLG2 assay for relative quantification of active mtDNA replication forks .....	77
2.2.21 Flow-cytometric imaging to analyse <i>de novo</i> mtDNA synthesis .....	77
<b>2.3 Fly work .....</b>	<b>79</b>
2.3.1 Routine maintenance of fly lines .....	79
2.3.2 Immunofluorescent staining and confocal imaging of <i>Drosophila melanogaster</i> flight muscle.....	79
2.3.3 Flight test .....	80
<b>2.4 Statistical &amp; biostatistical analyses.....</b>	<b>81</b>
2.4.1 Multiple sequence alignments.....	81
2.4.2 Sequence searches.....	81
2.4.3 Calculations.....	81
<b>3. Results .....</b>	<b>82</b>
<b>3.1 Mitochondrial DNA: a novel fluorescence-based replication assay .....</b>	<b>82</b>
3.1.1 The synthesis and characterization of monoclonalised NGFP-POLG2 U2-OS cell lines .....	83
3.1.2 ddC-treatment leads to a reduction of NGFP-POLG2, while cell size and mitochondrial content remain unchanged.....	87
3.1.3 Overexpression of the exonuclease pMA3790-mCherry in mitochondria leads to a reduction in NGFP-POLG2 fluorescence .....	89
3.1.4 Overexpression of POLG1 leads to an increase in NGFP-POLG2 fluorescence .....	91
3.1.5 Overexpression of TWINKLE leads to an increase in NGFP-POLG2 fluorescence .....	93
<b>3.2 Analysis of MacroD1 interaction partners .....</b>	<b>95</b>
3.2.1 Analysis of MacroD1 interaction partners .....	95
3.2.2 MacroD1 interactome: MacroD1 interacts with the mitochondrial DNA replication and repair factors. ....	97
3.2.3 MacroD1-based ADP-ribosylome: MacroD1 interaction with the mtDNA replication factors depends on ADP-ribosylation .....	101
3.2.4 MacroD1 interacts with factors involved in phosphorylation reactions.....	103
3.2.5 Generating and characterising monoclonalised U2-OS cell lines expressing core mtDNA replication factors.....	105
3.2.6 Verifying ADP-ribosylation status of mitochondrial DNA replication and repair factors...107	
<b>3.3 The biological function of MacroD1 and ADP-ribosylation in mitochondria .....</b>	<b>110</b>
3.3.1 The characterization of MacroD1 knockout cells: MacroD1 impacts mitochondrial and cellular homeostasis .....	113
3.3.2 Rescue by gene supplementation depends on ADPr-binding capacity of MacroD1. Overexpression of ADPr-binding mutant exhibits dominant negative effect.....	117
3.3.3 MacroD1 affects complex formation of core mitochondrial DNA replication factors on mitochondrial DNA .....	121
3.3.4 MacroD1 deletion in U2-OS cells does not affect mtDNA mutation load.....	123
3.3.5 Deletion of MacroD1 decreases active mtDNA replication forks and mtDNA synthesis ...125	
3.3.6 MacroD1 affects mtDNA copy number in an ADPr-dependent manner.....	129
3.3.7 Depletion of a mitochondrial macrodomain protein impacts on the organismal level.....131	

<b>4. Discussion .....</b>	<b>134</b>
4.1 The novel POLG2-assay: monitoring mtDNA replication .....	134
4.2 MacroD1 interacts with the core mtDNA replication machinery.....	135
4.3 MacroD1 is a vital factor in the regulation of mtDNA replication.....	137
<b>5. Summary of findings .....</b>	<b>141</b>
<b>6. Future work.....</b>	<b>143</b>
6.1 Future applications for the POLG2-assay .....	143
6.2 What are the effects of stable overexpression of mtDNA replication factors? .....	144
6.3 Which stage of mtDNA replication is affected by ADP-ribosylation?.....	145
6.4 Which enzymes MAR- and /or PARylate mitochondrial proteins?.....	145
6.5 Which residues are targeted by mitochondrial ADP-ribosylation? .....	147
<b>References .....</b>	<b>149</b>
<b>Appendices .....</b>	<b>163</b>
Appendix 1: Affidavit .....	164
Appendix 2: Abbreviations.....	165
Appendix 3: Recipes.....	169
Appendix 4: Antibodies.....	173
Appendix 5: Primers .....	174
Appendix 6: Plasmids.....	175
Appendix 7: Cell lines.....	177
Appendix 8: Fly lines .....	178
Appendix 9: Data Links & MTS Predictions .....	179
Original Mass Spectrometry Analyses Sheets.....	179
Data Links.....	179
MTS predictions: MS Data .....	180
Volcano Plot.....	184
MTS predictions: PARPs .....	185



## Figures & Tables

<b>Figure 1.1: Mitochondrial Morphology.</b> .....	16
<b>Figure 1.2: The Human Mitochondrial Genome and its Products.</b> .....	20
<b>Figure 1.3: Mitochondrial DNA Replication Fork.</b> .....	24
<b>Figure 1.4: Mechanism of reversible target protein ADP-ribosylation.</b> .....	34
<b>Tables 1.1: Identified ADP-Ribosyltransferases and their enzymatic activity.</b> .....	36
<b>Table 1.2: ADP-Ribosylhydrolases, their specificity and enzymatic activity.</b> .....	36
<b>Table 1.3: Mitochondrial ADP-ribosyltransferases and their enzymatic activity.</b> .....	40
<b>Table 1.4: Mitochondrial ADP-ribosylhydrolases and enzymatic activity.</b> .....	40
<b>Figure 1.5: The macrodomain protein MacroD1.</b> .....	44
<b>Table 2.1: Restriction reaction</b> .....	55
<b>Table 2.2: Ligation reaction</b> .....	56
<b>Table 2.3: Site-directed mutagenesis reaction.</b> .....	57
<b>Table 2.4: Master mix (per sample)</b> .....	58
<b>Table 2.5: Cycle parameters for qPCR amplification</b> .....	58
<b>Table 2.6: Master mix – long amplicon</b> .....	60
<b>Table 2.7: Cycle parameters – long amplicon PCR</b> .....	60
<b>Table 2.8: Master mix – short amplicon</b> .....	60
<b>Table 2.9: Cycle parameters – short amplicon PCR.</b> .....	60
<b>Table 2.10: HEK293 -T-Rex cells are seeded in accordance with the specifications below:</b> .....	66
<b>Table 2.11: U2-OS cells are seeded in accordance with the specifications below:</b> .....	66
<b>Figure 2.1: Gating example for flow cytometric analysis.</b> .....	74
<b>Figure 2.2: Gating for flow cytometric analysis in imaging flow cytometric assay in the IDEAS software.</b> .....	78
<b>Figure 2.3: Tube for flight test.</b> .....	80
<b>Figure 3.1: Stable NGFP-POLG2 Expressing U2-OS Cell Lines for POLG2 Assay.</b> .....	84
<b>Figure 3.2: Pharmacological inhibition of mtDNA replication.</b> .....	86
<b>Figure 3.3: Pharmacological inhibition of mtDNA replication measured by flow cytometry on living cells.</b> .....	88
<b>Figure 3.4: Overexpression of the mitochondrially targeted exonuclease mUNG1 (pMA3790 mCherry) reduces active mitochondrial replication forks.</b> .....	90
<b>Figure 3.5: Overexpression of wild-type POLG1 and exonuclease-deficient POLG1 increases active mitochondrial replication forks and mtDNA copy number.</b> .....	92
<b>Figure 3.6: Overexpression of TWINKLE increases active mitochondrial replication forks.</b> .....	93
<b>Figure 3.7: ShinyGO hierarchical clustering tree of the MacroD1-enriched interactors implicating MacroD1 in mtDNA homeostasis and mitochondrial gene expression.</b> .....	94
<b>Figure 3.8: The MacroD1 Interactome.</b> .....	96
<b>Figure 3.9: The mtTARG1 Interactome.</b> .....	98
<b>Figure 3.10: The MacroD1-specific ADP-ribosylome &amp; ADPr-binding-independent interaction partners of MacroD1.</b> .....	100
<b>Figure 3.11: Mitochondrial Phosphorylation Factors Interact with MacroD1.</b> .....	102
<b>Figure 3.12: U2-OS cells stably expressing mtGFP and the mtDNA replication factors: POLG1-GFP, TWINKLE-GFP and LIG3-GFP.</b> .....	104
<b>Figure 3.13: ADP-ribosylation Verification.</b> .....	106
<b>Figure 3.14: Essential mtDNA packaging factor TFAM is ADP-ribosylated &amp; MacroD1 impacts mitochondrial, not nuclear, ADP-ribosylation.</b> .....	108
<b>Figure 3.15: CRISPR/Cas9 deletion of the MacroD1 gene in U2-OS cells.</b> .....	112
<b>Figure 3.16: MacroD1 knockout cells exhibit a reduction in respiration and a slower proliferation rate.</b> .....	114
<b>Figure 3.17: The phenotypic changes are MacroD1-dependent and rescuable by gene-supplementation.</b> .....	116
<b>Figure 3.18: Rescue by gene supplementation depends on ADPr-binding capacity of MacroD1.</b> .....	118
<b>Figure 3.19: Characterization of MacroD1-rescue cell line.</b> .....	119
<b>Figure 3.20: MacroD1 affects complex formation of core mitochondrial DNA replication factors on mitochondrial DNA.</b> .....	120

<b>Figure 3.21: Change in DNA and DNA factor complex formation does not lead to a significant change in mtDNA integrity or increased mutation load.</b> .....	122
<b>Figure 3.22: Deletion of MacroD1 decreases active mtDNA replication forks.</b> .....	124
<b>Figure 3.23: Deletion of MacroD1 lowers de novo mtDNA synthesis.</b> .....	126
<b>Figure 3.24: MacroD1 affects mtDNA copy number in an ADPr-dependent manner.</b> .....	128
<b>Figure 3.25: CG33054, mitochondrial macrodomain protein in <i>Drosophila melanogaster</i>.</b> .....	130
<b>Figure 3.26: Mitochondrial Flight Muscle and Flight phenotypes upon MacroD1-like mitochondrial macrodomain protein knockdown in <i>Drosophila melanogaster</i>.</b> .....	132
<b>Figure 4.1: MacroD1 &amp; ADP-ribosylation regulate mtDNA replication, altering mitochondrial morphology and metabolism.</b> .....	138
<b>Figure 5.1: The POLG2, mtDNA Replication Assay.</b> .....	142
<b>Table 6.1: Which Enzyme MAR- and /or PARylate Mitochondrial Proteins?</b> .....	146
<b>Table 7.1: Nucleotide Abbreviations</b> .....	167
<b>Table 7.2: Amino Acid Abbreviations</b> .....	168
<b>Table 7.3: Antibodies used in this thesis.</b> .....	173
<b>Table 7.4: Primers used in this thesis.</b> .....	174
<b>Table 7.5: Plasmids used in this thesis.</b> .....	175
<b>Table 7.6: Cell lines used and/or generated in this thesis.</b> .....	177
<b>Table 7.7: Fly lines used in this thesis.</b> .....	178

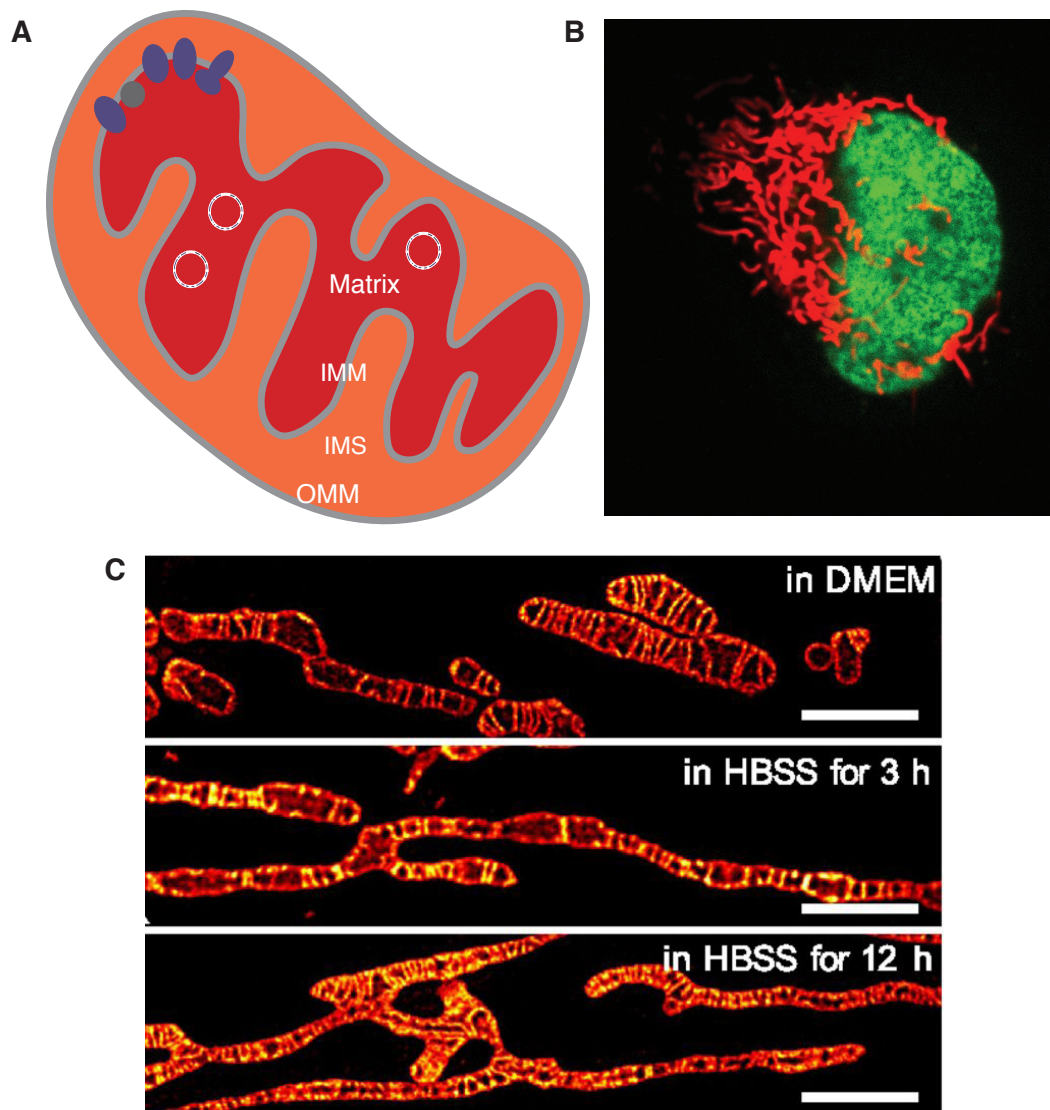
# 1. Introduction

## 1.1 Mitochondria

Mitochondria are double-membrane-bound organelles which are found in most eukaryotic cells. They generate the majority of the cells chemical energy and are therefore often referred to as “the powerhouse of the cell”. While mitochondria are important to the cells energy flux, their abundance, activity and turnover are highly regulated. This section focuses on the origin of mitochondria (section 1.1.1) and on their form and function (section 1.1.2). Special attention is drawn to the mitochondrial genome (section 1.1.3 to 1.1.6) and the specificities of mitochondrial myopathies (section 1.1.8).

### 1.1.1 The origins of mitochondria date back approximately 2 billion years

A high amount of energy is required to sustain multicellular organisms. Energy that must be made available, stored, distributed and transformed. This flux of energy is required to fuel, organise and maintain the intricate and extensive systems of eukaryotic cells and the organisms made up of these cells, such as our human form. The origin for this production and storage of energy is anchored in a single endosymbiotic event dating back approximately 2 billion years, when a heterotrophic anaerobe host cell related to the Asgard Archaea engulfed an aerobic  $\alpha$ -proteobacterium (Roger et al., 2017; Wallace, 2010), the birth of a subcellular double-membrane bound organelle, the mitochondrion (Gray, 1992; Mereschkowsky, 1910; Sagan, 1967).



**Figure 1.1: Mitochondrial Morphology.** A) Schematic drawing of a mitochondrion with its four compartments: OMM (grey), IMS (orange), IMM (grey) and matrix (red). The electron transport chain components are situated in the IMM (top left) (See text for further description). B) Confocal image of the mitochondrial network (red) surrounding the nucleus (green) of a human osteosarcoma cell (U2-OS cell line). C) Deconvoluted STED images showing changes in the mitochondrial morphology upon nutrition starvation (HBSS containing  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ , Scale bar:  $2\ \mu\text{m}$ ). Figure from (Wang et al., 2019), permission through PNAS.



The origin of these mitochondria, the engulfment of one single-membrane bound organism by another, has imbued mitochondria (as well as chloroplasts) with distinct features amongst the organelles of eukaryotic cells. One of these features are the four discrete compartments, their form and function. The innermost of these four compartments, the matrix, contains the other distinct feature of mitochondria, a separate genome, mitochondrial DNA (**mtDNA**) (see section 1.1.3).

### 1.1.2 The form of mitochondria have impact on their function

Textbook illustrations of mitochondria traditionally show a bean-shaped, single organelle (see Figure 1.1 A). In reality, mitochondria come in a plethora of shapes and sizes depending on organism, tissue and cell type. In humans mitochondria form an extensive filamentous network throughout the cell body (see Figure 1.1 B). Additionally, mitochondrial networks are not static, their constant movement (Lewis and Lewis, 1914) is the ongoing fusion and fission of mitochondria within this network (Wang et al., 2019) (see Figure 1.1 C).

Mitochondria are divided into four discrete compartments: outer mitochondrial membrane (**OMM**), intermembrane space (**IMS**), inner mitochondrial membrane (**IMM**) and matrix. The OMM is formed like a protective skin around the rest of the mitochondrion. The narrow IMS is flanked by the OMM and the IMM. The IMM with its folded layers forms the cristae structures, which hold the megadalton electron transport chain (**ETC**) complexes, surrounding the aqueous matrix in which the ancient mitochondrial genome is found (Figure 1.1).

Each of these compartments contains highly specialized proteins and plays vital roles in a number of different cellular processes. The foremost function is the production of adenosine triphosphate (**ATP**), the organic compound providing the energy required by living cells. ATP is produced from glucose and oxygen via two processes, the citric acid cycle and oxidative phosphorylation (**OXPHOS**) via the ETC (Andersson et al.,

2003). The ETC is composed of complexes I–V. Complexes I – IV set up an electrochemical gradient which drives complex V, the ATP synthase, catalysing the synthesis of ATP from ADP and phosphate (see Figures 1.1 A and 1.2 B). This energy production can be redistributed within the cell due to the motility and spread of the mitochondrial filamentous network, supplying the energy where it is needed.

Hence, the involvement of mitochondria as regulators of metabolism, cell-cycle and developmental control, antiviral response and cell death, collectively making mitochondria into coordinators of cellular stress responses (Eisner et al., 2018; McBride et al., 2006; Spinelli and Haigis, 2018). Another mitochondrial function is the maintenance and regulation of its ancient genome, which codes for key components of aerobic respiration and translation, rendering it indispensable for aerobic metabolism.

### 1.1.3 Mitochondria contain an ancient genome

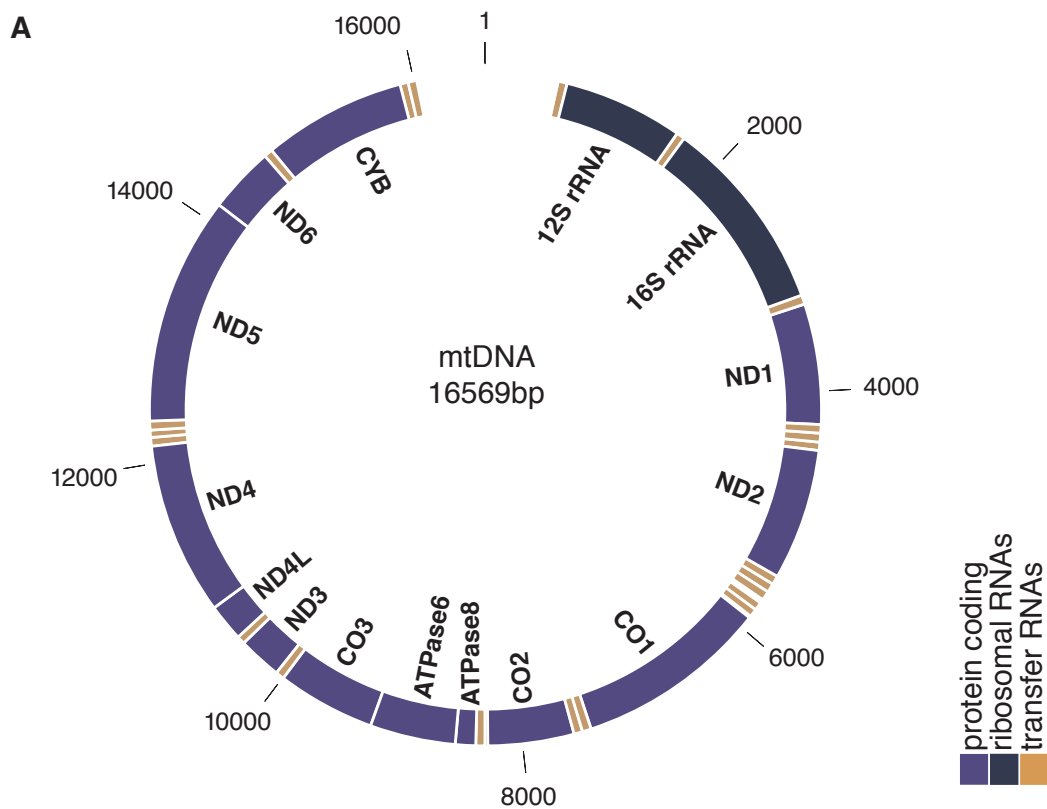
The first observations of mitochondrial DNA, and with that the existence of an extranuclear genome as a general phenomenon, were made by electron microscopy of chick embryo mitochondria (Nass and Nass, 1963). This was closely followed by biochemical studies of DNA found in mitochondrial fractions from *Saccharomyces cerevisiae* (Schatz et al., 1964). Subsequent studies revealed the sequence of human mtDNA (Anderson et al., 1981), as well as its  $\alpha$ -proteobacterium ancestral origin (Friedman and Nunnari, 2014; Stewart and Chinnery, 2015).

Through evolution, much of the ancestral endosymbionts genome has been passed to the host's chromosomes. This mitochondrial genome reduction is most readily explained by endosymbiotic gene transfer (Stewart and Chinnery, 2020; Timmis et al., 2004), also elucidating why the majority of mitochondrial proteins (estimated at around ~1,500) are encoded by the nuclear genome (not on mtDNA) and translated in

the cytoplasm before being shuttled across the OMM, or the OMM and the IMM into mitochondria and the mitochondrial matrix.

Due to their separate evolutionary origins, there a number of other disparities between nuclear DNA (**nDNA**) and mtDNA: i) the mitochondrial genome is circular, ii) lacks introns and ii) is organized loosely into nucleoids – 1 to 2 DNA molecules packed tightly into DNA-protein complexes. iv) Mitochondria are polyploid, each cell can have between 100 and several 100000 independent mtDNA copies and v) their genetic code is different (UGA is a nuclear stop codon but in mitochondria codes for tryptophan) (Gustafsson et al., 2016; Stewart and Chinnery, 2015).

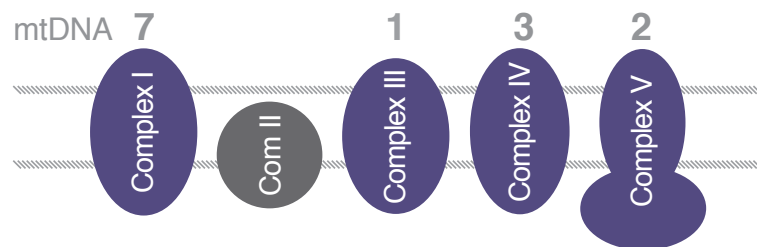
The mitochondrial genome retained by humans is formed of an inner **L** (light) and an outer **H** (heavy)-strand, the naming of which is based on respective molecular weight (Stewart and Chinnery, 2015) and spans 16569 base pairs (**bp**), containing 37 genes. These genes encode for 13 mitochondrial proteins (all subunits of the ETC), 2 ribosomal RNAs and 22 transfer RNAs (see Figure 1.2). This means that the gene products of mitochondria are either directly or indirectly (via translation) involved in energy metabolism (see Figure 1.2 B) and therefore essential for the wellbeing of the entire organism. It should come as no surprise that dysregulation of the mitochondrial genome is implicated in a variety of neurological diseases and mitochondrial myopathies. The aetiologies behind these are often either mutations in the mitochondrial genome, or mutations of the nuclear genome sequences coding for factors involved in the maintenance of mtDNA (Nunnari and Suomalainen, 2012; Stewart and Chinnery, 2015; Taylor and Turnbull, 2005; Young and Copeland, 2016). Mitochondrial disorders are further discussed in section 1.1.8.



**B**

**37 mitochondrial genes**

13	protein coding	OXPHOS system
2	ribosomal RNAs	
22	transfer RNAs	



**Figure 1.2: The Human Mitochondrial Genome and its Products.** A) Human mtDNA spans 16569 base pairs and codes for 13 proteins (purple), 2 ribosomal RNAs (blue) and 22 transfer RNAs (ochre). B) There are 37 mitochondrial genes, 13 of which code for essential subunits of four of the five complexes in the ETC. The ETC is shown as a schematic, complexes containing mitochondrial-encoded subunits are purple. Complexes I – IV set up an electro-chemical gradient. Complex V is the ATP synthase.

#### 1.1.4 The mitochondrial genome is distinctly organised

While there are usually two copies of nDNA per cell, human cells contain between 100 and several 100000 copies of mtDNA (Gustafsson et al., 2016). The DNA in the nucleus is highly organized, nDNA wraps around eight histone proteins - forming the nucleosome. At low salt conditions, *in vitro*, these nucleosomes interspersed on DNA look like 'beads-on-a-string', at physiological conditions the 'beads-on-a-string' can adopt higher grades of compaction and more condensed structures (Olins and Olins, 2003). This nDNA compaction state is modulated and controlled by chemical modifications, regulating accessibility and with that transcription and replication (Kouzarides, 2007; Li et al., 2007).

Mitochondrial DNA on the other hand was long postulated to be free-floating and unprotected in the mitochondrial matrix and under constant assault from reactive oxygen species (**ROS**), by-products of the ETC. This production and accumulation of ROS in the mitochondrial matrix has historically been linked to an increase or accumulation of mtDNA mutations during ageing. It was believed that mtDNA is free and uncompacted making it a target for ROS. However, this assumption is contentious – and mtDNA is not unprotected or loose. We now know that mtDNA is packaged into tight nucleoprotein complexes called nucleoids. Nucleoids (in humans) consist of 1 (seldomly 2) mtDNA molecules tightly coated by approximately 1000 mitochondrial transcription factor A (**TFAM**) proteins (1 TFAM per 16 bp mtDNA). Each of these high-mobility group domain proteins are capable of inducing 180° DNA U-turns (Bonekamp and Larsson, 2018). However, overall the compaction of nucleoids is not as organized as the chromatin found in the nucleus. Instead of the higher-order structure of a 30 nm-fibre, the protein-DNA complexes in mitochondria can be envisioned more like a tightly packed, crunched and balled up piece of paper, with a slightly elongated shape spanning 80 × 80 × 100 nm (Bonekamp and Larsson, 2018; Gustafsson et al., 2016).

### 1.1.5 The replication and transcription of mtDNA are interconnected

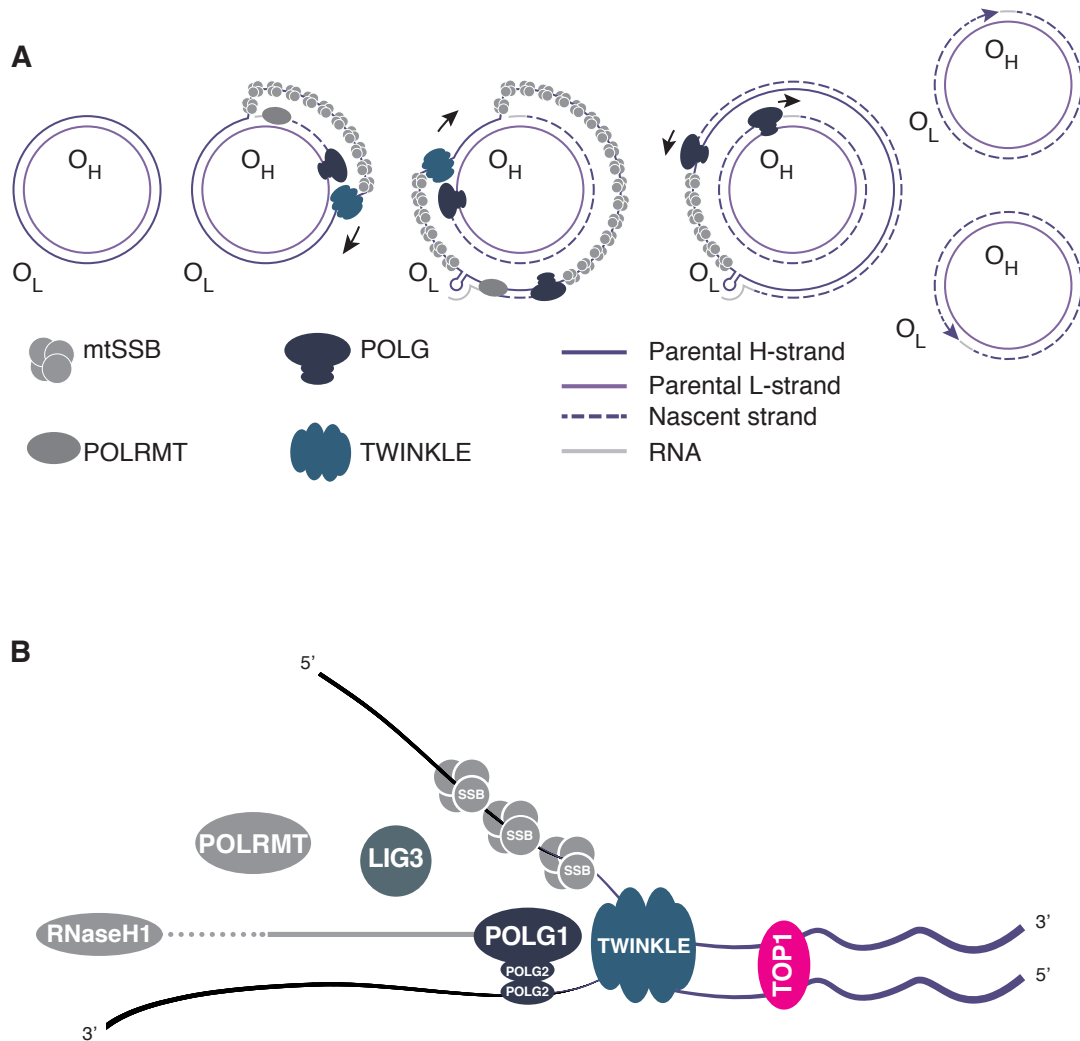
TFAM is not only the sole mitochondrial protein fulfilling the criteria of an mtDNA packaging factor (Bonekamp and Larsson, 2018; Kukat et al., 2011a; Park and Larsson, 2011), but also essential in transcription and mtDNA replication. All of the proteins involved in transcription and mtDNA replication are encoded in the nucleus and shuttled into mitochondria, while the ribosomal and transfer RNAs, required for translation of the transcribed DNA, are mtDNA-encoded. In mitochondria, transcription and replication of mtDNA is constant, simultaneous and uncoupled from the progression of the cell cycle. The transcription-initiation machinery for mtDNA consists of: POLRMT, TFB2M and TFAM. The core replication factors are: POLG, POLRMT, TWINKLE, LIG3, mtSSB, and topoisomerases (Falkenberg, 2018; Gustafsson et al., 2016; Holt and Reyes, 2012; Yasukawa and Kang, 2018). Some of the factors involved in transcription are also core replication factors, but what does the interplay between transcription initiation, transcription and replication consist of?

As discussed previously (section 1.1.3 Figure 1.2), human mitochondrial DNA contains the coding regions for 37 essential genes, all this information is very tightly packed into the 16569 bp small genome. There is only one longer noncoding region the **NCR**. Within the NCR there are two promoters, one for each strand, the light-strand and the heavy-strand promoter (**LSP** and **HSP**). The mitochondrial DNA-directed RNA polymerase (**POLRMT**) binds to these promoter regions in a sequence specific manner to then initiate transcription together with TFAM and mitochondrial transcription factor B2 (**TFB2M**). Transcription from these promoter regions produces polycistronic transcripts of the entire genome stopping just before the NCR. Through RNA processing of these primary near-genome-length transcripts, individual RNA molecules are released (Gustafsson et al., 2016), some of which are the tRNAs or rRNAs required to translate the remaining RNA molecules into the subunits of the ECT. However, POLRMT is not only essential in transcription of mtDNA. In fact, what makes replication of the mitochondrial genome so unique is that it is primed by extension of

processed RNA transcripts synthesized by that same RNA, thereby unequivocally interlinking mtDNA transcription with mtDNA replication.

The current understanding of mtDNA synthesis is that it proceeds continuously on both the H- and L-strand, much like the strand displacement model (Robberson et al., 1972). According to this model (see Figure 1.3 A), there is a dedicated origin of replication on each strand, denoted as **OH** and **OL**. Replication initially starts at OH, synthesizing a new H-strand. At this initial stage there is no L-strand replication and the displaced parental H-strand is coated in mitochondrial single strand binding proteins (**mtSSBs**), thereby preventing POLRMT from initiating random RNA synthesis. When the replication fork passes the OL, the H strand folds over into a stem-loop blocking mtSSBs, thus leaving a short region of single-stranded DNA accessible for POLRMT to prime for L-strand synthesis. After about 25 nucleotides (**nts**) DNA polymerase  $\gamma$  (**POLG**), which is at the core of mtDNA replication, takes over and both H- and L-strand synthesis proceed until both strands have reached full circle (Falkenberg, 2018).

Besides POLG, active replication forks require the interplay between POLRMT, TWINKLE, LIG3, mtSSB, and topoisomerases (Falkenberg, 2018; Gustafsson et al., 2016; Holt and Reyes, 2012; Yasukawa and Kang, 2018) (see Figure 1.3 B). TWINKLE is the mitochondrial DNA helicase that unwinds the mitochondrial genome ahead of POLRMT and POLG (Korhonen et al., 2003; Peter and Falkenberg, 2020). Mitochondrial SSBs coat single-strand DNA protecting it against nucleases (Falkenberg, 2018). Topoisomerases (such as TOP1 and TOP2A) relieve the torsional tension created by DNA unwinding (Zhang et al., 2001). Ligation of newly synthesized mtDNA requires the only known mitochondrial DNA ligase, LIG3 (Simsek et al., 2011). Essential transcription and replication factors, POLRMT, TFAM, TFB2M, POLG, TWINKLE, Topoisomerases and LIG3 will be discussed individually in more depth.



**Figure 1.3: Mitochondrial DNA Replication Fork.** A) Replication of mtDNA initiates at OH and produces a full-length nascent H-strand. Mitochondrial SSBs bind and protect the parental H-strand. As the replisome passes OL, a stem-loop structure forms, blocking binding for mtSSBs, thereby presenting a single-stranded loop-region from which POLRMT can synthesize new primer RNA. After about 25 nt, POLG replaces POLRMT and DNA synthesis commences on the L-strand. Synthesis of the H- and L-strands continues in an uninterrupted manner, until two double-stranded DNA molecules have been formed. The figure is based on three reviews (Falkenberg, 2018; Falkenberg and Gustafsson, 2020; Gammage and Frezza, 2019). B) The core replisome consists of a number of factors. TWINKLE unwinds mtDNA ahead of POLRMT and POLG. POLG extends the POLRMT generated RNA primers thereby synthesising DNA. RNaseH1 digests and removes the RNA. Topoisomerases relieve the torsional tension created by DNA unwinding, while mitochondrial SSBs coat single-strand DNA. LIG3 is the only known mitochondrial DNA ligase, ligating newly synthesized DNA fragments.



#### *1.1.5.1 The mitochondrial DNA-directed RNA polymerase is involved in both transcription and replication of mtDNA*

The mitochondrial RNA polymerase (**POLRMT**) was first discovered in humans based on its resemblance to the yeast homologue RNAP (Hillen et al., 2017; Ringel et al., 2011), both of which share high sequence similarity with the T3 and T7 bacteriophage polymerases (Masters et al., 1987). POLRMT is encoded in the nucleus and spans 1230 amino acids. It is a large (140 kDa) single subunit enzyme, essential in both mitochondrial gene expression and DNA synthesis (Falkenberg, 2018; Gustafsson et al., 2016; Kasiviswanathan et al., 2012). The loss of POLRMT in a conditional mouse model leads to embryonal lethality at E8.5 (Kuhl et al., 2016). The first 41 amino acids of POLRMT make up the mitochondrial targeting sequence (**MTS**), cleaved off upon mitochondrial import. The mature form of POLRMT interacts with mtDNA in a sequence specific manner at the mtDNA promoter regions and catalytic nucleotide incorporation. However, POLRMT alone is not able to initiate transcription or replication. It requires the interplay with the transcription and packaging factors TFB2M and TFAM.

#### *1.1.5.2 The mitochondrial transcription factor A packages mtDNA into nucleoids*

The mitochondrial transcription factor A (TFAM) is encoded in the nucleus and spans 246 amino acids, the first 42 of which make up the MTS, cleaved off upon mitochondrial import. TFAM belongs to the high mobility group-box domain family. In mitochondria two main functions are attributed to TFAM i) the activation of transcription at the two promoter regions (LSP and HSP) and ii) the coating and packaging of the entire mitochondrial genome into nucleoids. In support of these two functions, TFAM has been shown to bind promoter regions specifically, while also coating mtDNA in an unspecific binding manner inducing a U-turn with an overall bend of 180° (Rubio-Cosials and Sola, 2013). Upon binding mtDNA, TFAM induces unwinding of part of the surrounding DNA, generating bubbles which might drive protein aggregation into these patches (Gustafsson et al., 2016). It has further been

argued that this is what ultimately makes TFAM indispensable for transcription. TFAM unwinds mtDNA at the promoter region, if one of the TFAM-induced bubbles at the promoter meets with a POLRMT:TF2BM bubble, transcription from this region can take place (Gustafsson et al., 2016). If the separate TFAM-induced patches do not coalesce and don't form transcription bubbles with POLRMT:TF2BM transcription does not take place. The coating of the mitochondrial genome by TFAM is due to unspecific binding along the entire molecule with a ratio of 1 TFAM protein per 16 bp of mtDNA. This coating, with the subsequent U-turns, induces the tight compaction of mtDNA inherent in nucleoid compaction and is believed to protect the mitochondrial genome from degradation (Gustafsson et al., 2016; Kanki et al., 2004; Kukat et al., 2015; Kukat et al., 2011a; Wang et al., 2013b). TFAM and how it regulates transcription and replication of mtDNA is further discussed in subsection 1.1.6.

#### *1.1.5.3 The mitochondrial transcription factor B2 enables transcriptional initiation*

The Mitochondrial transcription factor B2 (TFB2M) is an essential component of the mitochondrial transcription initiation complex, and like TFAM and POLRMT encoded by the nuclear genome and transported into mitochondria. TFB2M spans 396 amino acids, the first 19 of which encode the MTS, cleaved upon import into mitochondria. Mature TFB2M functions as a homodimer (Hillen et al., 2017) and interacts with both TFAM (McCulloch and Shadel, 2003) and POLRMT (Falkenberg et al., 2002) within the transcriptional initiation complex. In this complex TFB2M functions by inducing POLRMT structural changes enabling promoter opening by POLRMT and transcription initiation (Hillen et al., 2017). Once POLRMT leaves the promoter region and the elongation phase of transcription is reached, TFB2M is believed to disassociate (Gustafsson et al., 2016).

#### *1.1.5.4 DNA polymerase $\gamma$ is the sole mtDNA polymerase*

DNA polymerase  $\gamma$  (POLG) is at the core of mtDNA replication, as the only replicative polymerase in mitochondria (Falkenberg, 2018; Hudson and Chinnery, 2006; Yasukawa and Kang, 2018). In humans POLG consists of the catalytic POLG1 subunit (Lecrenier et al., 1997; Ropp and Copeland, 1996) and the homodimeric POLG2 accessory subunit (Lim et al., 1999; Young et al., 2015). POLG1 (spanning 1239 amino acids) and POLG2 (spanning 485 amino acids) are encoded in the nucleus on two different chromosomes and transported into mitochondria where, in humans, they function as heterotrimers. The catalytic POLG1 subunit carries out three enzymatic activities: i) DNA polymerase ii) 3' to 5' exonuclease and iii) 5'-deoxyribose phosphate lyase activities (Krasich, 2017). While the homodimeric accessory subunit, POLG2, is required for both processivity of mtDNA replication and for the activity of the catalytic subunit, POLG1 (Fan et al., 2006; Young et al., 2015). While in vitro assays demonstrated that POLG1 alone is sufficient for DNA synthesis, the synthesized DNA did not exceed ~150 nucleotides. However, upon the addition of POLG2, extension of DNA could be increased to the full-length of the mitochondrial genome (Young et al., 2015). Additional experiments in cell culture have shown that although POLG2 does not incorporate in all mtDNA nucleoids, it is associated with all replicative mtDNA (Lewis et al., 2016), making it an essential factor of the active replisome.

A function unrelated to transcription or replication, is the degradation of linearized mtDNA. It has been shown that the 3' to 5' exonuclease activity of POLG1 is essential in degrading mtDNA linearized due to double strand breakage (Peeva et al., 2018). This could circumvent the reported lack of efficient double strand repair systems in mitochondria (Alexeyev et al., 2013) and dilute damaged mtDNA.

The essential role of POLG1 and POLG2 has been demonstrated in a number of model organisms. POLG1 KO in mice, for example, resulted in lethality at E7.5-8.5 (Hance et al., 2005). While POLG2 KO was embryonic lethal at E8-8.5 (Humble et al., 2013).

#### *1.1.5.6 The mtDNA helicase TWINKLE unwinds dsDNA*

The mitochondrial DNA helicase TWINKLE, so called due to its appearance when TWINKLE is GFP-tagged - shining twinkling foci throughout the mitochondrial network as stars would in the night sky (Spelbrink et al., 2001). TWINKLE spans 684 amino acids and is part of the SF4 superfamily. Like the other members of this family, TWINKLE functions in a ring-shaped helicase fashion. TWINKLE monomers polymerise into a higher order conformation, normally as hexamers although heptamers have also been reported (Peter and Falkenberg, 2020; Spelbrink et al., 2001).

TWINKLE is essential in both transcription and replication of the mitochondrial genome, as it is the only mitochondrial replicative DNA helicase, unwinding DNA ahead of POLRMT and POLG in the active replication fork (Korhonen et al., 2003; Peter and Falkenberg, 2020). Recently, TWINKLE has also been implicated as somewhat of a “bottleneck” of mtDNA replication, meaning that alterations in levels of TWINKLE lead to an alteration in mtDNA copy number (Ikeda et al., 2015; Peter and Falkenberg, 2020; Tynismaa et al., 2004).

#### *1.1.5.7 Mitochondrial topoisomerases relieve mtDNA torsional tension*

The mitochondrial topoisomerases (**TOPOs**) are essential in replication and gene expression as they allow changes and adjustments in the topology of mtDNA. Various topoisomerases are present in mitochondria, such as: TOP1, TOP3A, TOP2A and TOP2B relieving the torsional tension created by DNA unwinding (Gustafsson et al., 2016; Zhang et al., 2001). These topoisomerases are classified into different categories: type IA, IB and IIA depending on whether they catalyse transient single- or double-strand breaks. TOP1 belongs to the type IB topoisomerase transiently breaking DNA strands one at a time relaxing both negatively and positively supercoiled DNA (Dalla Rosa et al., 2014; Zhang et al., 2001). TOP3A belongs to type IA which can only relax negatively supercoiled DNA. TOP2A and B belong to the topoisomerase IIA family

cleaving both strands of the DNA duplex thereby leading to the relaxation of both negatively and positively supercoiled DNA (Zhang et al., 2014).

#### *1.1.5.8 Mitochondrial ligase 3 is critical for mtDNA integrity*

Ligase 3 (**LIG3**) is 1009 amino acids long, the first 42 of which are indicated as the MTS, cleaved off upon import. Although LIG3 has both nuclear and mitochondrial forms, it is the only DNA ligase known to exist in mitochondria, making it essential in ligating newly synthesized mtDNA fragments both during the termination of mtDNA replication as well as during mtDNA repair (Akbari et al., 2014; Lakshmipathy and Campbell, 1999; Simsek et al., 2011).

#### 1.1.6 Mitochondrial transcription and replication are regulated in numerous ways

Having discussed mtDNA (section 1.1.3), how it is packaged (section 1.1.4), transcribed and replicated, as well as the factors essential in performing these actions (section 1.1.5), it becomes apparent that gene expression and mtDNA replication are inextricably linked. The transcriptional initiation machinery (TFAM, TFB2M, POLRMT) is simultaneously also the machinery that initiates mtDNA synthesis and is required before the core replication factors (POLG, TWINKLE, TOPOs, mtSSB and LIG3) can actively synthesise DNA (Falkenberg, 2018; Gustafsson et al., 2016; Holt and Reyes, 2012; Yasukawa and Kang, 2018). However, the regulation of replication and thereby of copy number still has many open questions and points of investigation. What is known is that mtDNA replication can be modulated and controlled at various points; at mtDNA accessibility, at transcription initiation, and at the switch between transcription initiation and replication (Falkenberg, 2018; Fuste et al., 2010). The actual interface between transcription and mtDNA replication, when POLRMT primes for DNA synthesis and when it generates near-genome-length RNA transcripts is not very well understood.

One proposed layer of regulation is by epigenetic modification, both by direct DNA methylation marks and by modification of the proteins involved in the packaging of this DNA, such as TFAM. However, as with many other DNA processes, epigenetic modifications and regulation are studied mainly in the nucleus, with their elucidation in the mitochondrion still facing various obstacles (Pawar and Eide, 2017; Stimpfel et al., 2018). Therefore, we know surprisingly little about the role of DNA modifications and PTMs in the regulation of the location, activity and turnover of mtDNA replication and mitochondrial homeostasis.

One of the few mitochondrial DNA-interacting proteins that has been studied more extensively in regard to PTM modification, as well as the impact this modification has on the maintenance and regulation of the mitochondrial genome, is TFAM. Possibly due to both i) the abundance on and ii) importance of TFAM in the maintenance of mtDNA. The modifications of this mitochondrial transcription factor are reported to include O-linked glycosylation, acetylation and phosphorylation (Dinardo et al., 2003; King et al., 2018; Lu et al., 2013; Suarez et al., 2008; Wang et al., 2014). Phosphorylation of TFAM Ser residues, for example, has been shown to have regulatory function in the TFAM-DNA interaction mechanism (King et al., 2018) as well as being able to abrogate transcriptional activation (Lu et al., 2013). So far, not much is known about ADP-ribosylation of mitochondrial DNA replication and maintenance factors.

In this thesis the interaction between the ADP-ribosyl-hydrolase MacroD1 (subchapter 1.3) with various components of mtDNA synthesis and gene expression is investigated in the context of mtDNA regulation. We ask whether mitochondrial ADP-ribosylation and the factors involved in the writing, reading and erasing of this post-translational modification might be the orchestrators, regulating when initiation by POLRMT, TFAM and TFB2M leads to gene expression and when to genome replication.

### 1.1.7 There is co-regulation between the mitochondrial and the nuclear genome

The ETC consists of dual-origin complexes. That they are made up of not only subcomponents encoded by mtDNA, but also nDNA. As imbalanced subunit production frequently culminates in OXPHOS defects and subsequent malignancies, it is important to maintain a certain co-regulation of mitochondrial and nuclear transcriptional programs. This tight co-regulation is achieved through mitochondrial – nuclear signalling pathways, as well as clustered expression of nDNA encoded OXPHOS subcomponents (Heddi et al., 1993; Isaac et al., 2018; Jazwinski, 2013).

### 1.1.8 Mitochondrial DNA maintenance is disease-relevant

The mitochondrial genome along with the compaction, transcription and replication factors involved in the maintenance of mtDNA perform vital contributions to oxidative phosphorylation and with that to the bioenergetic profile and homeostasis of the organism. Thus, rendering the maintenance and faithful translation of this small genome essential for most multicellular aerobic life. Therefore, it should come as no surprise that disturbances or defects in the replication or maintenance of mtDNA causing deletions, mutations or copy number changes can lead to respiratory disturbances and is implicated in a variety of neurological diseases and mitochondrial myopathies. The aetiologies behind these diseases are usually either mutations in the mitochondrial genome, or mutations of the nuclear genome sequences encoding the mtDNA replication and repair factors involved in the maintenance of mtDNA (Nunnari and Suomalainen, 2012; Stewart and Chinnery, 2015; Suomalainen and Battersby, 2018; Taylor and Turnbull, 2005; Young and Copeland, 2016). The common feature being that they mainly affect tissues with high energy demand. As changes in mtDNA copy number lead to changes in ETC complexes in the IM and with that to changes in the bioenergetic system, so do mutations in mtDNA, be it in the coding sequences for the ETC subcomplexes directly or in the tRNAs and rRNAs required for their synthesis.

Myopathies with changes to copy number of mtDNA or deletions within the mitochondrial genome are normally due to mutations in the genes coding for the proteins necessary for mtDNA replication and maintenance. Examples are POLG1 mutations in Alpers-Huttenlocher syndrome and POLG2 or TWINKLE mutations in progressive external ophthalmoplegias (**PEO**) (Copeland, 2012; Young and Copeland, 2016). In fact, disease causing mutations in POLG, as it is the sole replicative mitochondrial polymerase, are so well studied that a number of databases are dedicated solely to POLG mutations and corresponding diseases (Nurminen et al., 2017; Rahman and Copeland, 2019).

While myopathies caused by mutations in the core factors of mtDNA maintenance follow either an autosomal dominant or recessive inheritance pattern, the inheritance patterns of myopathies caused by mutations in mtDNA vastly differ. The mitochondrial genome is inherited from a single parent, the mother. Normally, patients with harmful mutations in their mtDNA present clinical symptoms once the fraction of deleterious mtDNA molecules pass a certain threshold level, often reported at > 60% with increasing severity depending on copy number (Stewart and Chinnery, 2015; Suzuki et al., 2011).

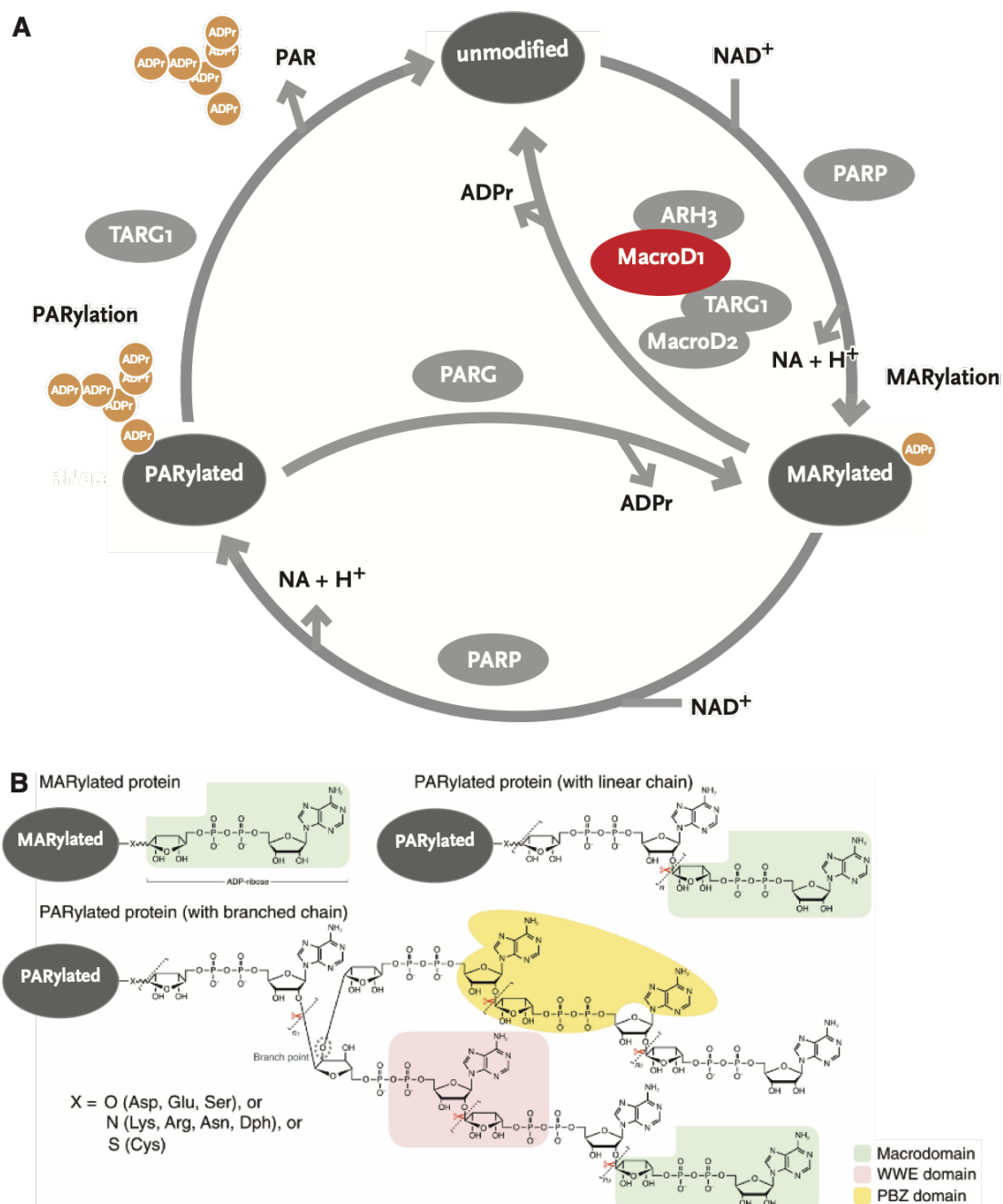
A treatment approach for these spontaneous or inherited mtDNA mutations has proven elusive, with treatment focused mainly on palliative care. However, recently there have been three paths of investigation one prenatally, aiming to avert transmission of deleterious mtDNA prior to fertilization, the other two postnatally by targeted degradation of inherited deleterious mtDNA (Bacman et al., 2018; Gammage et al., 2018) and mtDNA editing using a new CRISPR-free mtDNA base -editing tool (Mok et al., 2020).

Prenatally, transmission of deleterious mtDNA can be reduced by mitochondrial replacement therapy (**MRT**). In MRT the nDNA from the egg of the “mother” is transplanted into an enucleated egg from an unaffected female donor, prior to



fertilization (Greenfield et al., 2017; Stewart and Chinnery, 2015). Thus, resulting in a “three-parent child”, the first of which was born in 2017 (Zhang et al., 2017). As these methods are relatively new, the long-term consequences have not been fully explored, neither in regards to nuclear-mitochondrial genome mismatch (Stewart and Chinnery, 2015) nor in regards to observations regarding reversion to the deleterious mitochondrial genome (Hudson et al., 2019).

Transfecting nucleic acids into mitochondria, as well as recombining them into endogenous mtDNA is notoriously tricky, due to both the morphology of mitochondria and an inefficient mtDNA repair system (Lightowers, 2011). Therefore, CRISPR-editing of DNA, a method explored in the nuclear context, is not a viable option in mitochondria. There are currently two methods under investigation which show promise in editing or depleting detrimental mtDNA in humans postnatally, both of which employ synthetic protein constructs, artificially targeted to mitochondria. The first employs a bacterial cytidine deaminase toxin which has been shown to function in targeted mutations of mtDNA in cultured cells (Mok et al., 2020). The second method, which has shown some promise in a pathogenic mtDNA mouse model (Kauppila et al., 2016), is the exploitation of the ability of the mitochondrial replisome to efficiently degrade linearized mtDNA. TALENs and ZNFs are targeted to mitochondria, cleaving specifically the mutant DNA, thereby selectively reducing the amount of this mtDNA (Bacman et al., 2018; Gammage et al., 2018). Although the melioration of the disease phenotype was modest in these studies, the method might herald a first approach to target pathogenic mtDNA in an organism postnatally.



**Figure 1.4: Mechanism of reversible target protein ADP-ribosylation.** A) Some PARPs are able to catalyse MARYlation on target proteins. Some PARPs are subsequently able to further polymerize this initial MARYlation of the protein and form PAR chains - PARylating the target protein. PARG degrades PAR chains, leaving the initial modification – the mono-ADP-ribose moiety. To fully reverse the protein to its unmodified state mono-ADP-ribosyl-hydrolases such as MacroD1 (red), MacroD2 or TARG1 are required. The figure is based on the Barkauskaite review (Barkauskaite et al., 2013). B) Proteins can MARylated and/ or PARylated by linear or branched PAR chains. Shown are examples of different parts of protein-conjugated ADP-ribose and the specific protein domains that recognize them. Figure with minor alterations from (Leung, 2017), permission granted through the Copyright Clearance Center.

## 1.2 ADP-ribosylation

A post-translational modification (**PTM**) is the covalent modification of a protein, subsequent to protein synthesis. In this subchapter we will discuss one of these PTMs, the rapid and reversible addition of one or more ADP-ribose moieties onto a target protein (ADP-ribosylation). This section focuses on how ADP-ribosylation is regulated (section 1.2.1), the factors involved in ADP-ribosylation (section 1.2.2), as well as the seldomly discussed subset of ADP-ribosylation happening outside the nucleus – within the mitochondrion (section 1.2.3).

### 1.2.1 How is ADP-ribosylation regulated and what is the function?

ADP-ribosylation is the transfer of ADP-ribose from NAD<sup>+</sup> to the side chains of amino acids. In this fully reversible post-translational modification, mono- and poly-ADP-ribose (**MAR & PAR**) moieties can be added, recognized and removed from substrate proteins (see figure 1.4). This highly dynamic addition and removal of ADP-ribose moieties often induces an alteration in the function, activity, or interaction of the substrate protein (Barkauskaite et al., 2013). As a PTM, ADP-ribosylation has been shown to be heavily involved in the maintenance of the nuclear genome, specifically in chromatin structure (Poirier et al., 1982), transcriptional regulation (Kraus, 2008) and DNA damage repair (D'Amours et al., 1999; Pascal and Ellenberger, 2015).

Gene name	Alt. name	Activity
PARP1	ARTD1	PARylation
PARP2	ARTD2	PARylation
PARP3	ARTD3	MARylation
PARP4	vPARP/ARTD4	PARylation
PARP5a	Tankyrase1/ARTD5	PARylation
PARP5b	Tankyrase2/ARTD6	PARylation
PARP9	ARTD9/BAL1	inactive
PARP14	ARTD8/BAL2	MARylation
PARP15	ARTD7/BAL3	MARylation
PARP7	TiPARP/ARTD14	MARylation
PARP12	ARTD12	MARylation
PARP13	ZAP/ARTD13	MARylation
PARP10	ARTD10	MARylation
PARP11	ARTD11	MARylation
PARP6	ARTD17	MARylation
PARP8	ARTD16	MARylation
PARP16	ARTD15	MARylation
ARTC1		MARylation
ARTC3		inactive
ARTC4		inactive
ARTC5		MARylation
SIRT4		MARylation
SIRT6		MARylation

**Tables 1.1: Identified ADP-ribosyltransferases and their enzymatic activity.** Some PARPs are able to catalyze MARylation, some PARylation. The Sirtuins SIRT4 and SIRT6 have also been shown to possess MARylation capacity. This Table is based on Table 1 in (Luscher et al., 2018).

Gene name	Specificity	Activity
PARG		PAR-hydrolase
ARH3	Ser	MAR- & PAR-hydrolase
MacroD1	Glu	MAR-hydrolase
MacroD2	Asp/Glu	MAR-hydrolase
TARG1	Asp/Glu	MAR- & PAR-hydrolase

**Table 1.2: ADP-ribosylhydrolases, their specificity and enzymatic activity.** ADP-ribosylhydrolases functioning either on PAR- or MARylated substrates, reverting proteins to either MARylated or unmodified states.

### 1.2.2 Which factors write, read or erase the ADP-ribosyl modification?

During reversible mono- and poly-ADP-ribosylation (MARylation & PARylation), a plethora of proteins have to work in concert to catalyse the transfer of ADP-ribose moieties onto target proteins, to remove them and to imbue them with specificity for their substrates. Poly-ADP-ribose polymerases (**PARPs**) are the major family of enzymes, catalysing the transfer of ADPr moieties from NAD<sup>+</sup> to the substrate proteins, 'writing' the modification. The 'readers' are proteins that can bind to MAR- and PARylated protein substrates: PAR-binding linear motifs (**PBM**s), PAR-binding zinc-fingers (**PBZ**s), WWE domains and macrodomain proteins (Barkauskaite et al., 2013). The macrodomains are of special interest as they are not only capable of binding ADP-ribosylated substrates as 'readers', but some of them are also capable of removing – 'erasing' – the PTM from modified proteins.

#### *1.2.2.1 ADP-ribosyltransferases can transfer either mono- or poly-ADPribose*

PARPs and also some members of the sirtuin (**SIRT**) family are the enzymes MAR- and PARylating target substrate proteins. There are many identified PARP proteins (see Table 1.1), some of which are capable of transferring poly-ADPr, some mono-ADPr and some being catalytically inactive. As such, PARPs is a bit of a misnomer, another name that is also frequently ascribed to these enzymes is ADP-ribosyltransferases (**ART**s) (Luscher et al., 2018). PARPs modify proteins at specific sites. The residues accepting the ADPr-modification have been shown to include: serine (Ser), cysteine (Cys), aspartate (Asp), glutamate (Glu), asparagine (Asn), lysine (Lys) or arginine (Arg) residues (Altmeyer et al., 2009; Cervantes-Laurean et al., 1995; Laing et al., 2011; Leidecker et al., 2016; Manning et al., 1984; McDonald and Moss, 1994; Palazzo et al., 2018).

Glu- and Asp-residues were long-believed to be the main modification targets in substrate proteins (Cervantes-Laurean et al., 1995). Until recently, when it was shown that Ser-residues are the main ADPr-modification sites (Larsen et al., 2018). Not only were they the most enriched targets, identified in human osteosarcoma cells by MS analysis, the modification on Ser-residues exhibited a 10- to 27-fold increase after oxidative DNA-damage treatment (Leidecker et al., 2016; Palazzo et al., 2018).

An additional interesting interplay between Ser-phosphorylation and Ser-ADP-ribosylation has also been highlighted, with these two modifications apparently co-targeting in the DNA-damage response (Larsen et al., 2018). The PARPs identified as 'writers' for Ser-ADPr are PARP1 and PARP2 (Bonfiglio et al., 2017a; Leidecker et al., 2016), their specificity for Ser-residues is conferred by their interaction with Histone PARylation Factor 1 (**HPF1**) (Bonfiglio et al., 2017b). While PARP1 and PARP2 would normally modify substrates on Glu- and Asp- residues, interaction with HPF1 changes their specificity to Ser-residues. It was further discussed that since PARP1 exists in much higher abundance in cells than HPF1, PARP1 might interact in concert with other factors imbuing specificity for additional/different ADPr-modification residues (Bonfiglio et al., 2017b).

#### *1.2.2.2 Readers contain ADP-ribose recognition motifs*

The PAR-binding proteins or 'readers' of protein-ADPr modifications are proteins containing motifs with which they can bind to MAR- and PARylated substrates. These ADPr-binding domains are classified into four separate classes: PBMs, PBZs, WWE and macrodomain proteins. Macrodomain proteins are the only ones to interact specifically with MARylated substrates (Barkauskaite et al., 2013; Luscher et al., 2018).

The first known ADPr-binding domain, the PBM, was identified 20 years ago and consists of a loosely defined 20 aa long stretch of hydrophobic amino acids interspersed by basic residues (Teloni and Altmeyer, 2016). PBMs have been found to

exist in over 500 proteins. They were the first to implicate the importance of ADPr in DNA replication, DNA repair, cell cycle regulation and chromatin modification (Teloni and Altmeyer, 2016). The importance of ADPr in DNA-damage signalling and repair was validated by the emergence of the PAR-binding motif, the PBZ – a Cys2-His2 type zinc finger motif spanning approximately 30 aa, which binds two consecutive ADPr-moieties (Ahel et al., 2008). WWE domain-containing proteins are named after their most conserved residues: Tryptophan (Trp or W) and Glu (E). These motifs have been found in 12 human proteins, specifically in PARP/ART family members and E3 ubiquitin ligases (Teloni and Altmeyer, 2016). This suggested an involvement of WWE domains in both ubiquitylation and PARylation pathways (Barkauskaite et al., 2013).

Proteins containing PBMs, PBZs and WWE domains are all capable of interacting with PARylated protein substrates. But there is one more ADPr-protein ‘reader’ - the macrodomain. Macrodomains are named after the histone variant containing one of these globular modules – macroH2A –which plays a significant role in X- inactivation (Lo Re and Vinciguerra, 2017). Initially identified in single-strand RNA viruses (including coronaviruses), these conserved protein domains have since been identified in bacteria, archaea, and eukaryotes (Kraus, 2009). It is believed that the viral macrodomains were possibly highjacked from the host vector genomes (Neuvonen and Ahola, 2009). Macrodomains are large 130 – 190 aa reader modules adopting a distinct fold made-up of a central  $\beta$ -sheet surrounded by 4 to 6  $\alpha$ -helices capable of binding to O-acetyl-ADP-ribose, as well as MARYlated and PARylated target proteins (Barkauskaite et al., 2013; Karras et al., 2005; Luscher et al., 2018). Besides being the only known domain shown to bind MARYlated protein substrates, some macrodomain containing proteins are also able to process this modification and remove it via hydrolysis.

Gene name	Alt. name	Activity	Localization
PARP1	ARTD1	PARylation	mitochondria ?
SIRT 4		MARylation	mitochondria

**Table 1.3: Mitochondrial ADP-ribosyltransferases and their enzymatic activity.**

Gene name	Alt. name	Activity	Localization
PARG	PARG55 isf	inactive	mitochondria
PARG	PARG60 isf	inactive	mitochondria
ARH3		MAR - & PAR-hydrolase	mitochondria
MacroD1		MAR-hydrolase	mitochondria

**Table 1.4: Mitochondrial ADP-ribosylhydrolases and enzymatic activity.**



### *1.2.2.3 ADP-ribosylhydrolases function as erasers of the post-translational modification*

The mechanism of target protein ADP-ribosylation is fully reversible, some of the macrodomain containing ADPr-interaction partners of ADP-ribosylated proteins not only bind to ADPr, but also function as hydrolases, reversing the PTM (see Table 1.2).

Poly-ADP-ribose Glycohydrolase (**PARG**) was the first identified enzyme shown to degrade PAR chains, hydrolysing the ribose-ribose O-glycosidic bonds between ADPr units in PAR chains, leading to MARylated proteins (see Figure 1.4). PARG, ADP-ribose protein glycohydrolase TARG1 (**TARG1**) and the only known non-macrodomain-containing ADP-ribosylhydrolase: ADP-ribosylhydrolase 3 (**ARH3**) are the ADP-ribosylhydrolases able to degrade PAR chains (Barkauskaite et al., 2013; Luscher et al., 2018). TARG1 and ARH3, unlike PARG, can fully reverse the substrate proteins to their unmodified state. ARH3, however, functions exclusively on Ser-ADPr modifications while TARG1 targets Asp/Glu ADPr-modified substrates (Fontana et al., 2017; Sharifi et al., 2013). MacroD1 has been shown to function on Glu-modified residues and MacroD2 on Asp/Glu MARylated substrates, fully reversing the ADP-ribosylation cycle to an unmodified protein product (Jankevicius et al., 2013; Rosenthal et al., 2013).

### 1.2.3 The regulation and function of mitochondrial ADP-ribosylation

The mechanisms and functions of ADP-ribosylation have historically been investigated primarily in a nuclear context (Krishnakumar and Kraus, 2010) and although mitochondria contain roughly 70% of the total cellular NADH/NAD<sup>+</sup> pool (Alano et al., 2007), intramitochondrial ADP-ribosylation has not been studied as extensively. The presence of a mitochondrial enzyme system able to add and remove ADP-ribose moieties to a target protein was described over 40 years ago. However, in 1975 it was unclear which mitochondrial enzyme performed this function. What was clear, was

that the isolated protein was 100 kDa big and could disassociate into two 50 kDa subunits (Kun et al., 1975).

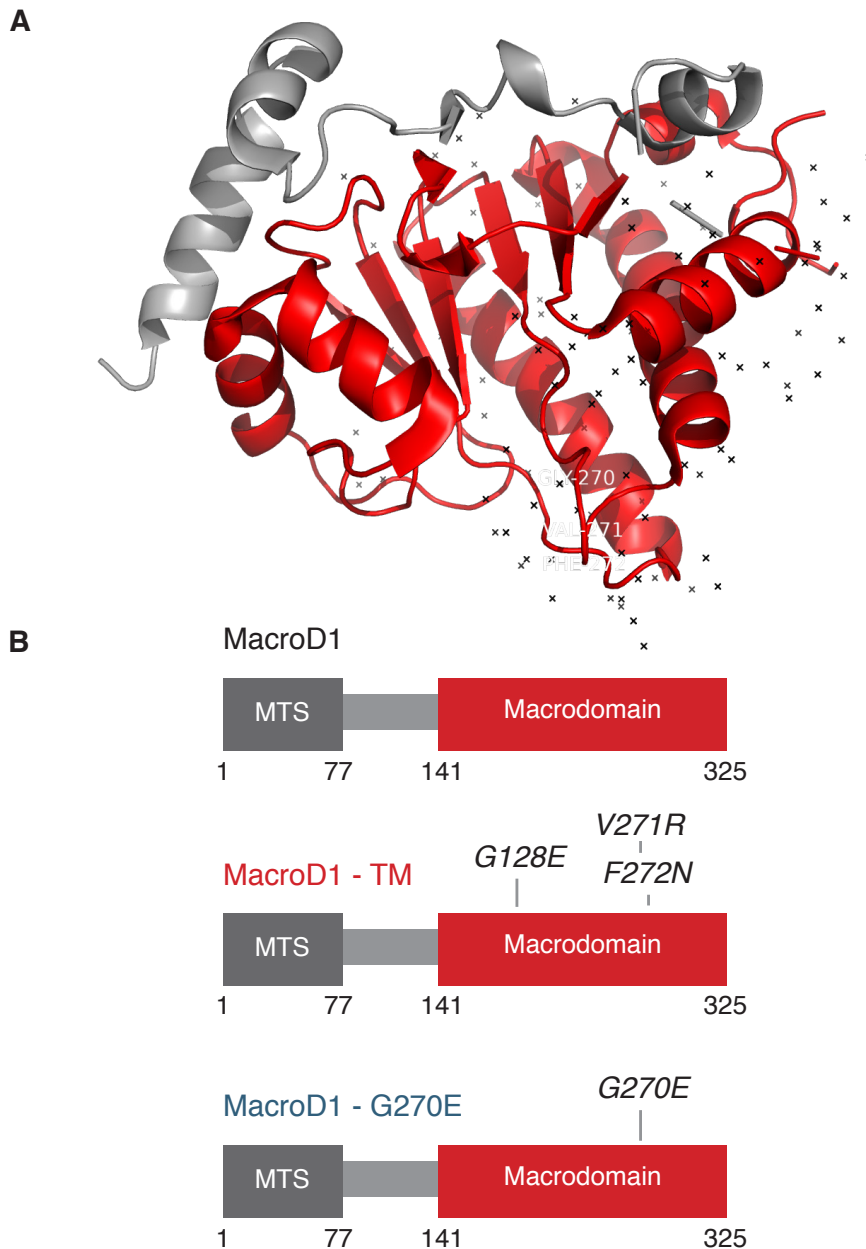
Since then two ADP-ribosyltransferases have been identified in mitochondria SIRT4 and PARP1 (Haigis et al., 2006; Rossi et al., 2009) (see Table 1.3), with the mitochondrial localization of PARP1 being somewhat contentious. It has been argued by people in the PARP field that the presence of PARP1 in mitochondria is unlikely, and that detection of this protein within mitochondria is merely due to the sheer abundance of it, additionally PARP1 is devoid of an MTS sequence. However, only a subset of mitochondrially-targeted proteins have a canonical MTS, as the presence of this sequence is only one of a number of mitochondrial import mechanisms (Calvo and Mootha, 2010). Additionally, PARP1 has been detected in mitochondria by mitochondrial fractionation and immunoprecipitation, as well as in immunofluorescent imaging (Lai et al., 2008; Rossi et al., 2009; Szczesny et al., 2008).

Although two 'writers' of the ADPr-modification have been identified in this organelle, to date only few mitochondrial proteins have been shown to be ADP-ribosylated (Gagne et al., 2008; Gagne et al., 2012; Haigis et al., 2006; Herrero-Yraola et al., 2001; Lai et al., 2008; Szczesny et al., 2014). The most frequently discussed is probably glutamate dehydrogenase (**GLUD1**), a metabolic enzyme catalysing the conversion of glutamate to  $\alpha$ -ketoglutarate. MARYlation of GLUD1, by SIRT4, has been shown to negatively modulate its enzymatic activity (Haigis et al., 2006; Herrero-Yraola et al., 2001). This means that GLUD1 is not only ADPr-modified, but also indicates a clear function of this modification on this protein in mitochondria.

If there are mitochondrially localized enzymes putting the ADP-ribosylation modification on target proteins, as well as proteins found to be modified by ADPr – are there enzymes found in mitochondria that reverse the modification?

Two isoforms of PARG have been shown to localize to mitochondria, PARG55 and PARG60 (Meyer et al., 2007; Niere et al., 2008). However, it appears that these isoforms are catalytically inactive, with PARG requiring the now cleaved off MTS region for its enzymatic function (Niere et al., 2012). ARH3, on the other hand, localizes to mitochondria and cleaves both MAR and PAR of ADPr-Ser residues (Niere et al., 2012). More recently, the Glu-MAR-specific ADP-ribosylhydrolase, MacroD1, has been identified to exclusively localize to mitochondria (Agnew et al., 2018; Neuvonen and Ahola, 2009; Žaja et al., 2020).

PARP1 ADP-ribosylates protein substrates on Glu/Asp-residues, unless PARP1 interacts with HPF1. There are no known reports of HPF1 in mitochondria. If PARP1 does in fact localize to mitochondria, it would stand to reason that in mitochondria its main target might be Glu/Asp-ADPr sites. ARH3 however, is incapable of reversing ADPr-modifications unless they are on a Ser-residue of the target protein. Might MacroD1 be the main ADP-ribosylhydrolase reversing mitochondrial ADP-ribosylation?



**Figure 1.5: The macrodomain protein MacroD1.** A) MacroD1 is shown from residue 91-325. The macrodomain is in red. The Figure is the from the crystal structure (PDB Entry - 2X47) deposited by (Chen et al., 2011). B) Schematic MacroD1, which spans 325 amino acid residues. The mitochondrial targeting sequence, cleaved of upon mitochondrial import, consists of the first 77 amino acid residues. The macrodomain of MacroD1 extends from residue 141 to 325. With specific mutations (G128E/V271R/F272N) in the macrodomain it is possible to generate a trapping mutant (MacroD1 – TM). Through a different mutation (G270E) it is possible to generate a MacroD1 mutant incapable of ADPr-binding (MacroD1 - G270E).

## 1.3 MacroD1

That the ADP-ribosylation reaction is catalysed by PARPs/ARTs has long been established (Chambon et al., 1963), so has the degradation of PAR chains from PARylated substrate proteins via PARG (Miwa and Sugimura, 1971; Miwa et al., 1974), as reviewed by W. Lee Kraus in his review *PARPs and ADP-Ribosylation: Fifty Years... and Counting* (Kraus, 2015). What wasn't known was whether and how the complete reversal from PARylated to unmodified protein takes place until the more recent discoveries of ARH3 and the macrodomain proteins: MacroD1, MacroD2 and TARG1 (Chen et al., 2011; Jankevicius et al., 2013; Niere et al., 2008; Peterson et al., 2011; Rosenthal et al., 2013; Sharifi et al., 2013) and with that enzymes capable of fully reversing the ADP-ribosylation modification. Amongst the four macrodomain-containing proteins known to possess ADP-ribosylhydrolase activity (MacroD1, MacroD2, TARG1 and PARG), MacroD1 stands apart for two main reasons: i) it localises exclusively to mitochondria (Agnew et al., 2018; Neuvonen and Ahola, 2009; Žaja et al., 2020) and ii) it is capable of fully reversing Glu-ADPr in vitro (Jankevicius et al., 2013).

### 1.3.1 What are the molecular properties of MacroD1?

MacroD1 is a relatively small protein, spanning 325 amino acid residues, containing a long N-terminal stretch consisting of the 77 amino acid long MTS and a C-terminal macrodomain (see Figure 1.5).

#### *1.3.1.1 MacroD1 contains a canonical mitochondrial targeting sequence*

The MTS was initially predicted using the MitoProt algorithm and tested on ectopically expressed protein constructs (Neuvonen and Ahola, 2009), subsequently this localisation was confirmed by mutation truncation studies as well as with MacroD1-specific antibodies on endogenous MacroD1 (Agnew et al., 2018). It was further

shown that MacroD1 is enriched in tissue with high energy requirements, such as human and mouse skeletal muscle (Agnew et al., 2018; Žaja et al., 2020), correlating with high concentrations of mitochondria in these tissues. Thus, firmly demonstrating an exclusively mitochondrial localisation, despite earlier reports attributing it to the nucleus.

#### *1.3.1.2 The macrodomain of MacroD1 can be mutated to abolish ADPr-binding capacity and catalytic activity*

The macrodomain on the C-terminus of MacroD1 is capable of binding ADPr (Neuvonen and Ahola, 2009), *in vitro* experiments have further shown that MacroD1 is also capable of processing the ADPr-moiety (Jankevicius et al., 2013; Neuvonen and Ahola, 2009). The interaction between MacroD1 and ADP-ribosylated substrate proteins is MAR-Glu specific (Jankevicius et al., 2013).

A single point mutation of the Glycine residue 270 into either Tyr or Glu is capable of abolishing the binding capacity of MacroD1 for ADPr (Chen et al., 2011; Neuvonen and Ahola, 2009). By performing three separate point mutations within the macrodomain of MacroD1, it is further possible to generate a MacroD1 triple mutant (here called MacroD1 – TM) with abolished MacroD1 ADP-ribosyl-hydrolase activity but still binding ADPr and its derivatives (Kistemaker et al., 2016), thereby trapping substrate proteins.

#### 1.3.2 Is MacroD1 disease-relevant? A critical view at the literature

Very little definitive is known about the relevance of MacroD1 in human health and disease. Although there are some reports of both overexpression and depletion of MacroD1 being cancer relevant, many of these reports appear nuclear centric (Feijs et al., 2020), referring to MacroD1 as a nuclear protein (Mo et al., 2018; Zang et al., 2018). However, this reported nuclear localisation of MacroD1 is contradictory to data

obtained from our own lab (presented in this thesis), as well as data from a number of other labs focusing concerted efforts on MacroD1-biology (Agnew et al., 2018; Neuvonen and Ahola, 2009; Žaja et al., 2020). Publicly available databases such as MS data from MitoMiner v4.0 (Smith and Robinson, 2019) and human MitoCarta2.0 (Calvo et al., 2016), as well as RNA-sequencing data from FANTOM5 available through the Human Protein Atlas <http://www.proteinatlas.org> (Noguchi et al., 2017), all accessed in 2020, further corroborate the mitochondrial localisation of endogenous MacroD1.

From the RNA-seq data sets one can additionally see that MacroD1 is highly enriched in both heart and skeletal muscle, which are mitochondria-rich tissue with elevated energy demand. Although it is impossible to exclude that in certain pathologies a subfraction of MacroD1 erroneously localises to the nucleus, its main function appears to be mitochondrial. Mitochondrial localisation of MacroD1 does not exclude oncopotential, as functional mitochondria are essential to cancer progression (Wallace, 2012; Zong et al., 2016). If MacroD1 is essential for mitochondrial homeostasis, alterations in MacroD1 might have tumorigenic effects.

Considering the rapidly emerging role of macrodomain proteins, similar to MacroD1, as mediators of the replication of viral genomes (Fehr et al., 2016; Fehr et al., 2018; Frick et al., 2020; Grunewald et al., 2019; Srinivasan et al., 2020). As well as the involvement of nuclear macrodomain proteins in DNA-damage response and transcription, we wonder whether the role of MacroD1 in mitochondria might be related to the replication or maintenance of the mitochondrial genome.

## 1.4 Aim of the study

There are three main aims of this thesis outlined below. These aims will be addressed in the same order in both the results and the discussion sections.

### Aim I: Mitochondrial DNA - a novel fluorescence-based replication assay

There are several methods to monitor mitochondrial DNA replication in mammalian cells (Prole et al., 2020; Uhler J.P. and M., 2021). The most commonly-used methods to quantify mtDNA synthesis are by EdU-incorporation (Haines et al., 2010) or EtBr-staining (Villa et al., 2012) and subsequent imaging and foci counting. Alternatively, it is also possible to assess mtDNA content by quantitating the mtDNA/nDNA ratio (Venegas and Halberg, 2012) and making subsequent mtDNA replication efficiency assumptions in relation to mtDNA copy number changes. Both of these types of approaches have a number of different drawbacks. While EdU-staining is very specific and widely-used, standard confocal microscopes do not have the resolution required to count individual foci (Kukat et al., 2015; Kukat et al., 2011b), meaning that it becomes impossible to get a clear picture of foci count – specifically if a certain mutation or drug-treatment introduces changes in mtDNA segregation or compaction of nucleoids. The main drawback of qPCR is that cells have to be harvested and lysed, total DNA has to be isolated and qPCR has to be performed.

During the work for this thesis it became necessary to monitor mitochondrial replication in mammalian cells. We wanted to know the mtDNA/nDNA ratio – but also whether this ratio is directly associated with mtDNA replication, or whether it is a matter of mtDNA integrity. To directly measure relative mtDNA replication we developed an innovative flow-cytometry-based method, which in conjunction with our stably expressing U2-OS NGFP-POLG2 cell lines culminates in a novel mtDNA replication assay, the POLG2 mtDNA replication assay, with which it becomes possible to measure and quantitate changes in mtDNA replication *in vivo*.



The principle behind this method is as follows: POLG is the replicative DNA polymerase in mitochondria (Hudson, G. and Chinnery, P. F., 2006, Falkenberg, M., 2018, Yasukawa, T. and D. Kang, 2018) and in humans consists of the catalytic POLG1 (Ropp, P. A. and Copeland W. C., 1996, Lecrenier, N., et al., 1997) and the homodimeric POLG2 accessory subunits (Lim, S. E., et al., 1999, Young, M. J., et al., 2015). In mtDNA replication the accessory subunit, POLG2, is required for both processivity of mtDNA replication and activity of the catalytic subunit, POLG1 (Fan, L., et al., 2006, Young, M. J., et al. 2015). In *in vitro* assays POLG1 alone has been shown to synthesise a maximum of ~150 nucleotides. Upon the addition of POLG2 (N'-terminally fused to GFP) extension was increased to the full-length of mtDNA (Young, M. J., et al. 2015). Young et al. also demonstrated that stable insertion of NGFP-POLG2 into HEK293 cells did not alter the bioenergetic profile of the cells and that although NGFP-POLG2 always co-localized with mtDNA, not all mtDNA contained POLG2 (Young, M. J., et al. 2015). Subsequently, Lewis et al. used the same construct, and coupled with EdU-incorporation, demonstrated that all active mtDNA replication forks in transfected U2-OS cells contained NGFP-POLG2 (Lewis, S. C., et al., 2016).

We decided to use this same construct (kind gift from W. Copeland) to generate the cell-lines which are the main tool in the POLG2 mtDNA replication assay and stably integrated it into U2-OS cells – generating NGFP-POLG2 U2-OS cell lines (see subchapters 3.1 and 4.1). With these cell lines we are able to use a flow cytometric analyser to comparatively quantify active mitochondrial replication forks, while simultaneously measuring changes to both cell size and mitochondrial content. As this assay can be performed on both living cells as well as fixed cells, using immunostaining and further gating, on for example expression levels of certain proteins, is an added level of investigation.

We have tested our new method using both negative and positive controls (see subchapter 3.1) and believe that we have been successful in generating an addition to the expanding toolkit used in the analysis of the black box that is mitochondrial replication.

## Aim II: Analysis of MacroD1 interaction partners

Both writers and erasers of the ADP-ribosyl modification mark are reported to localise to mitochondria (Agnew et al., 2018; Neuvonen and Ahola, 2009; Niere et al., 2008; Niere et al., 2012; Žaja et al., 2020). Some mitochondrial proteins, such as GLUD1 (Haigis et al., 2006; Herrero-Yraola et al., 2001), are known to be ADP-ribosylated. Although, while ADP-ribosylation in the nucleus is heavily involved in the maintenance of the nuclear genome (Barkauskaite et al., 2013), very little about its targets or functions in the mitochondrion have been elucidated.

The only protein involved in the ADP-ribosylation machinery that exclusively localizes to mitochondria is MacroD1 (Agnew et al., 2018; Neuvonen and Ahola, 2009; Žaja et al., 2020). Armed with this knowledge Aurelio Nardoza, a former postdoc in the Ladurner Lab, set out to generate a MacroD1-specific interactome and ADP-ribosylome, in collaboration with Fabian Hosp from the Matthias Mann Lab. To further strengthen the proteomic analysis, Aurelio and Fabian performed an identical set of experiments using an additional macrodomain as bait, artificially targeting TARG1 to mitochondria (mtTARG1). The data sets that were generated during these mass spectrometric experiments are analysed in this thesis with expert support from Benjamin E. Oliver.

Based on our analysis, we see two areas of ADP-ribosylation research opening up in this double-membrane-bound organelle: i) we observe a number of serine threonine kinases and their interaction partners amongst the MacroD1-substrates hits and ii) we identify core mtDNA replication and repair factors, namely: POLG1, LIG3, TOP1, TOP2A and TWINKLE. In an orthologous Far-western approach we biochemically validate some of the identified hits for this thesis – focusing on the mitochondrial replication factors.

### Aim III: Biological function of MacroD1 and ADP-ribosylation in mitochondria

The identification of both mtDNA maintenance factors, as well as factors involved in the phosphorylation of Ser-residues suggests a regulatory role of MacroD1 and ADP-ribosylation in the maintenance and homeostasis of mtDNA. To further investigate the importance of MacroD1 on the mitochondrial genome and in cellular homeostasis we use CRISPR/Cas9 MacroD1 knockout cell lines generated by Rebecca Smith, as well as MacroD1-overexpressing cells. In experiments assessing mtDNA copy number, mtDNA mutation load and mtDNA integrity we test how the deletion of MacroD1 or overexpression of the ADPr-binding-deficient MacroD1 mutant affects the homeostasis of the mitochondrial genome. Subsequently we investigate whether MacroD1 directly affects mtDNA replication using EdU-incorporation and our novel POLG2 assay.

Finally, we assess the impact these MacroD1-specific changes to the mitochondrial genome have for not only the physiology of the cell, but also on an organismal level, using the *Drosophila melanogaster* model.

## 2. Methods

### 2.1 Nucleic acid techniques

#### 2.1.1 Transformation of bacterial cells

Transformation is the introduction of naked, usually recombinant, DNA into bacterial cells across the cell membrane, to maintain and amplify said DNA. During this work, heat shock competent *E. coli* DH5 $\alpha$  cells were transformed with different DNA plasmids.

In brief, approximately 10 pg to 100 ng DNA are incubated with 50  $\mu$ L thawed competent cells for 20 minutes on ice. Subsequently the transformation mixture is heat-shocked at sub-lethal temperature, 42°C for 80 seconds, before being placed on ice for 2 minutes. If the antibiotic in question is ampicillin, cells are plated on agar plates containing the antibiotic and incubated at 37°C for 12-16 hours. If the antibiotic in question is not ampicillin (e.g. kanamycin) an additional expression step will have to be included in the protocol by adding ~ 800  $\mu$ L LB-medium and incubating at 37°C for 30 minutes, prior to plating.

#### 2.1.2 DNA extraction from bacterial cells

To extract plasmid DNA from the transformed bacteria, putative positive clones have to be determined and picked from the incubated agar dishes (section 2.1.1). The chosen clones are inoculated in a medium enriched with the appropriate antibiotic. The volume of the medium is dependent on whether a mini- or midiprep is to be performed. For both of these procedures our lab uses commercial kits. For minipreps the *mi*-Plasmid miniprep kit from metabion and for midipreps The PureYield™ Plasmid Midiprep System from Promega were used. The procedure was carried out in accordance with the manufacturer's recommendations.

### 2.1.3 DNA extraction from mammalian cells

To extract both mitochondrial and genomic DNA from mammalian cells we use commercial kits. To enrich mitochondrial DNA specifically we used the NucleoSpin®Plasmid kit from Macherey-Nagel, to extract both genomic and mitochondrial DNA we use the Relia-Prep-Kit from Promega.

To enrich mitochondrial DNA during DNA preps from mammalian cells we utilize the principles of DNA plasmid preps, this approach also works with mitochondrial DNA as mitochondrial genomes are small and circular, much like plasmids. From the three commercial kits we have tested for this specific application the NucleoSpin®Plasmid kit from Macherey-Nagel far outperforms others (Metabion & Qiagen). To isolate mtDNA using this kit we follow the manufacturer instructions with one small alteration, in the initial step instead of pelleting and washing bacterial cells, we pellet and wash mammalian cells.

We have also optimized our procedure for use with the Relia-Prep-Kit from Promega to our specific purpose. In brief, we harvest cells, wash pellets twice in 400  $\mu$ L 1x PBS, add 20  $\mu$ L Proteinase K and 400  $\mu$ L Cell Lysis Buffer then vortex for 10 seconds. We incubate the samples at 56°C for at least 30 minutes on the ThermoBlock with continuous shaking at 1000 rpm. Subsequently, we add 500  $\mu$ L Binding Buffer and vortex the samples for 10 seconds before loading onto the provided collection columns. The columns are centrifuged for 1 second on full speed and the flow through is discarded. The columns are washed 4x with 500  $\mu$ L column wash buffer with full speed centrifugation for 2 seconds per wash. After the last centrifugation, we change the collection column and centrifuge for 4 seconds, prior to eluting sample into a clean microcentrifuge tube using 200  $\mu$ L H<sub>2</sub>O. Subsequent to DNA isolation, concentration and quality of DNA is assessed (section 2.1.4).

#### 2.1.4 DNA concentration assessment

Nucleic acid concentration and quality is routinely measured using the NanoDrop2000 UV-Vis spectrophotometer from Thermo Scientific. The NanoDrop2000 measures the UV-absorption at different wavelengths; then calculates the nucleic acid concentration using a modified Beer-Lambert equation. Using the 260nm/280nm ratio, the NanoDrop2000 can also be used to assess the purity of the nucleic acid solution, with a ratio of  $\sim 1.8$  being considered as pure DNA.

For assays that require exact equivalence of starting material from multiple starting samples (e.g. qPCR) we use the Tecan plate reader infiniteM1000 with PicoGreen. We use the 96 well-plates from Corning to measure our samples consisting of: 95  $\mu\text{L}$  1xTE, 3  $\mu\text{L}$  DNA or PCR sample and 100  $\mu\text{L}$  1x PicoGreen. Fluorescent intensity is measured, samples are diluted and equilibrated until homogenous values are reached.

#### 2.1.5 DNA restriction digest, separation and purification from agarose gel

We use restriction endonucleases and alkaline phosphatases to cleave DNA at specific sites and to dephosphorylate the 5'-ends, respectively.

To set up restriction digests we mix the separate components as indicated in table 2.1, using  $\sim 500$  ng DNA to set up verification restriction reactions and approximately 2  $\mu\text{g}$  when DNA is needed for further subcloning. The restriction mixtures are incubated at the optimal temperature required for the enzyme in question. The reaction is stopped either by direct heat inactivation or with the addition of a DNA loading reagent.

To analyse restriction digests or to isolate DNA molecules of specific lengths, agarose gel electrophoresis is used, with agarose gels between 0.7% and 2%. The migration front is observed with either orange G or bromophenol blue dye (see Buffers Appendix 3 for

specifications), DNA itself is visualized by ethidium bromide (EtBr) intercalation under UV light. To purify DNA molecules from agarose gel the *mi-Gel* Extraction Kit from metabion is used.

**Table 2.1: Restriction reaction**

Reagent	Amount
H2O	
Buffer 10x	2 – 5 $\mu$ L
Plasmid DNA	500 ng - 2 $\mu$ g
Enzyme 4000 – 10000 U/ mL	0.5 – 1 $\mu$ L
SAP or CIP (when needed)	0.5 – 1 $\mu$ L
Total volume	20-50 $\mu$ L

### 2.1.6 DNA ligation

The ligation reactions for this thesis were performed using the T4 DNA ligase from NEB. The T4 DNA ligase joins cohesive and blunt ends in an ATP dependent reaction. We set up ligations in three reactions: control (no insert), 1:5 and 1:10. The ligation reaction was mixed as indicated in table 2.2. The insert amounts are calculated using the following equation:

$$\text{ng insert} = ((\text{ng vector} * \text{size (bp) insert}) / \text{size (bp) vector}) * \text{ratio}$$

The ligation reactions are incubated either at room temperature for 30 to 120 minutes or at 16<sup>o</sup>C overnight. Heat-inactivation of the ligation reaction is performed at 65<sup>o</sup>C for 10 minutes. Subsequently, the reactions are transformed into competent cells and streaked onto appropriate agar plates (see section 2.1.1). The control sample with no added insert is used in the assessment of re-ligation frequency and thereby relative amounts of false positives.

**Table 2.2: Ligation reaction**

Reagent	Amount
H2O	
T4 DNA Ligase Buffer 10x	2 $\mu$ L
Vector	100 ng
Insert	Calculated in samples / none in control
T4 DNA Ligase 40 U/ $\mu$ L	0.5 $\mu$ L
Total volume	20 $\mu$ L

### 2.1.7 Polymerase chain reaction (PCR)

During polymerase chain reaction the enzymatic activity of a thermostable DNA polymerase is exploited to amplify specific regions of DNA in vitro. We used a variety of different DNA polymerases depending on target length and fidelity requirements. As targets we used nuclear and mitochondrial genomic DNA, as well as plasmids purified from bacterial cultures (appendices 6 and 7 for plasmids and primers).

During this project the principle of PCR technology was applied for: subcloning and DNA mutagenesis (section 2.1.5 and 2.1.8), quantitative PCR measurements (section 2.1.9), semi-quantitative long-amplicon PCR (section 2.1.10) and Mitochondrial DNA Immunoprecipitation (section 2.2.15).

### 2.1.8 DNA mutagenesis

To introduce desired changes into the DNA sequence of plasmids, site-directed mutagenesis is performed. This mutation method is based on the principles of PCR, with small alterations. The two primers used, are complimentary to each other and complimentary to the intended template, except for the nucleotide changes desired in the target. The DNA polymerase, PfuUltra™ from Stratagene, has a very low error-rate. After approximately 18 PCR cycles the samples are taken out of the thermo cycler. The reactions now contain mutated linear DNA and non-mutated parental DNA templates. The parental strands were cultivated in E. coli and are thus methylated, the novel DNA strands are not. It is therefore possible to selectively



digest the parental templates using the methyl sensitive DpnI restriction enzyme, leaving only the novel DNA strands which are repaired and amplified after transformation into DH5 $\alpha$  E. coli cells (digestion and transformation is performed as in subsections 2.1.5 and 2.1.1).

For this project site-specific mutagenesis was used to generate different MacroD1 point mutants.

**Table 2.3: Site-directed mutagenesis reaction**

Reagent	Amount
H2O	
PfuUltra™ Buffer x10	5 $\mu$ L
Template	50 ng
Primer 1	125 ng
Primer 2	125 ng
dNTP 5 mM	1 $\mu$ L
Glycerol	5 $\mu$ L
PfuUltra™ Polymerase	1 U
Total volume	50 $\mu$ L

### 2.1.9 Quantitative PCR measurements (qPCR)

To determine relative copy number changes of mitochondrial DNA in relation to nuclear DNA, we performed quantitative real time PCR, using primer sets against two mitochondrial (tRNA<sup>Leu(UUR)</sup> and 16S rRNA) and one nuclear ( $\beta$ 2- microglobulin) gene sequences. The method is as described by Venegas and Halberg (Venegas and Halberg, 2012) with small alterations, the primers can be found in the primer table ( Appendix 5).

In brief, cells are harvested, washed twice by pelletation in 1xPBS, total DNA is isolated using the Relia-Prep-Kit (section 2.1.3) and brought to identical concentrations (2 ng/  $\mu$ L) with the Tecan plate reader and PicoGreen (section 2.1.4). The Master mix (see table 2.4) and the template DNA are pipetted into 96 well plates and kept in a cool dark place prior to measurements. We use SYBR SuperMix (iTaQ SYBR Green Supermix with ROX, Bio- Rad) and

the QuantStudio 3 Real-Time PCR System (ThermoFisher) for qPCR reactions (see table 2.5 for the cycle parameters). Subsequently,  $C_T$  values are exported from the analysis software and relative mtDNA content (in relation to nuclear DNA) is calculated with the following formula:

$$(\beta 2M \text{ average } C_T) - (tRNA^{Leu(UUR)} \text{ average } C_T) = \Delta C_T$$

$$2^{-\Delta C_T} = \text{mtDNA content}$$

$$\text{Relative mtDNA content} = (\text{mtDNA content cell line or treatment in question} / \text{mtDNA content control cell line or treatment})$$

**Table 2.4: Master mix (per sample)**

Reagent	Amount
SYBR SuperMix	5 $\mu$ L
Primer Mix	2 $\mu$ L (5 $\mu$ M)
H <sub>2</sub> O	1 $\mu$ L
<i>DNA template</i>	2 $\mu$ L
<b>Total volume</b>	<b>10 <math>\mu</math>L</b>

**Table 2.5: Cycle parameters for qPCR amplification**

Stages	Duration	Temperature
1	120 s	50°C
2	20 s	95°C
3	15 s	95°C
3	30 s	62°C ( $T_m$ )
4	15 s	95°C
4	20 s	50°C
4	15 s	95°C

### 2.1.10 Semi-quantitative long-amplicon PCR

Semi-quantitative long-amplicon PCR (LA-PCR) is used to determine relative mtDNA-integrity/ability of DNA-repair, subsequent to oxidative DNA-damage by H<sub>2</sub>O<sub>2</sub> -treatment. With this method single strand breaks (SSBs) are introduced into the DNA, the control cells are not treated. The basis of this assay is that the SSBs will halt the polymerase during the elongation step and mtDNA with more damage will thereby result in fewer of the long amplicons. We perform this assay as described by Szczesny, B., et al. (Szczesny et al., 2014) with some small changes.

In short: cells are incubated with 0.5 mM H<sub>2</sub>O<sub>2</sub> for 1 hour to induce DNA damage. Afterwards, cells are harvested and total DNA is extracted and brought to the same concentration (approximately 20 ng/ μL) (section 2.1.4). The damage to the mtDNA is then determined by amplifying both long (8.9-kb) and short (221-bp) amplicons (for primers see Appendix 5) from the mitochondrial DNA. In this assay we use two different polymerases, the Q5 High-Fidelity DNA Polymerase (NEB) for long amplicons and the Phusion-HF polymerase with GC-buffer for short amplicons. The relative amounts of amplicons are measured using the Tecan plate reader and PicoGreen (see section 2.1.4), the fluorescent intensity values generated from this plate reader are also used to calculate differences in PCR amplicon abundance. To correct for difference in mtDNA copy number per cell, we normalized the obtained data to the read out of the short amplicons using the following equation:

$$PCR\ product\ normalized = ((long\ fragment - long\ fragment_{T0}) / (short\ fragment - short\ fragment_{T0}))$$

We calculate the relative DNA damage/DNA integrity with the following equation:

$$FC = median\ H_2O_2\ treatment / median\ non-treatment$$

**Table 2.6: Master mix – long amplicon**

Reagent	Amount
5 x buffer	8 $\mu$ L
dNTPs (NEB)	0.8 $\mu$ L
8.9 kb Primer Mix	4 $\mu$ L
Template	5 $\mu$ L
Q5 Polymerase	0.4 $\mu$ L
H <sub>2</sub> O	21.8 $\mu$ L
<b>Total volume</b>	<b>40 <math>\mu</math>L</b>

**Table 2.7: Cycle parameters – long amplicon PCR**

Repeat		Duration	Temperature
1 x	initial denaturation	60 s	98 <sup>o</sup> C
34 x	denaturation	10 s	98 <sup>o</sup> C
	annealing	30 s	62 <sup>o</sup> C
	elongation	8 min	72 <sup>o</sup> C
1 x	final extension	10 min	72 <sup>o</sup> C

**Table 2.8: Master mix – short amplicon**

Reagent	Amount
5 x GC-buffer	8 $\mu$ L
dNTPs (NEB)	0.8 $\mu$ L
221 bp Primer Mix	8 $\mu$ L
Template	5 $\mu$ L
Phusion-HF Polymerase	0.5 $\mu$ L
H <sub>2</sub> O	17.7 $\mu$ L
<b>Total volume</b>	<b>40 <math>\mu</math>L</b>

**Table 2.9: Cycle parameters – short amplicon PCR**

Repeat		Duration	Temperature
1 x	initial denaturation	180 s	98 <sup>o</sup> C
19 x	denaturation	30 s	98 <sup>o</sup> C
	annealing	30 s	54 <sup>o</sup> C
	elongation	3 min	72 <sup>o</sup> C
1 x	final extension	5 min	72 <sup>o</sup> C

### 2.1.11 DNA sequencing

DNA sequencing was performed for the mutation load assay (section 2.1.12) and as an additional verification after subcloning and mutagenesis. All sequencing was performed by GATC Biotech AG. Sequences were subsequently analysed using either CLC Main Workbench (Qiagen) and or Serial Cloner (freeware).

### 2.1.12 Mitochondrial mutation load assay

The mutation load assay is performed as previously described (Spelbrink et al., 2000), with some minor alterations concerning the subcloning strategy. We use the NucleoSpin®Plasmid kit from Macherey-Nagel to extract mitochondrial DNA from mammalian cells (see section 2.1.3) and a 685 bp region of mitochondrial DNA with the addition of EcoRI and HindIII restriction sites is amplified (for primers see Appendix 5). The amplicons are purified and subcloned into pBluescript, one hundred clones from each cell line investigated are picked and sequenced, additionally we make an additional amplicon from one of the clones (Clone 2). Using this amplicon of one single clone we can estimate the error rate of the polymerase. All sequences are aligned to the mitochondrial reference genome and the errors/ mutations in a 500 bp sequence are manually counted. With these counts the implied mutation rate is then calculated with the following method:

$$\textit{Implied mutations} = \textit{total sequenced bp per cell line} / \textit{found mutations}$$

## 2.2 Protein and mammalian cell culture techniques

For the work in this thesis a number of different mammalian expression systems were used, depending on the experiment performed. The section in this subchapter will deal with specifications for the different cell lines and how they are generated (sections 2.2.1 to 2.2.3), general protocols we use when working with cells (sections 2.2.4 to 2.2.6), as well as the protocols used to analyse bioenergetic profiles (sections 2.2.9), proliferation (section 2.2.8), cell size (subsection 2.2.5) and protein content (sections 2.2.13, 2.2.16, 2.2.17, 2.2.18).

### 2.2.1 Stable, controlled, expression in HEK293 -T-Rex cells

Aurelio Nardoza generated stable inducible HEK293-T-REx cell lines, expressing controlled levels MacroD1-GFP, MacroD1<sup>TM</sup>-GFP MacroD1-G270E-GFP, mtTARG1-GFP, mtTARG1-K84A-G123E-GFP and mitochondrially targeted GFP by inserting the macrodomain mtGFP constructs into the pcDNA5/FRT/TO plasmid (Invitrogen) (Appendix 6). Subsequently the stable cell lines were generated using the the Flp-In T-REx-293 cell line and the FIT system (Invitrogen), in accordance with the manufacturer's instructions.

These were the cell lines used in the mass spectrometric experiments (sections 2.2.18 and 2.2.19).

### 2.2.2 CRISPR/Cas9 MacroD1 knockout in HEK293 -T-Rex & U2-OS cells

CRISPR/Cas9 MacroD1 knockout cell lines, characterised in this thesis in subsection 3.3.1, were generated using the approach described in Ran et al. (Ran et al., 2013). Rebecca Smith designed the primers for the MacroD1 targets using the guide design tool from the Zhang Lab (<http://crispr.mit.edu>) (Appendix 5) and generated MacroD1 knockout cells in the U2-OS and HEK293-T-Rex background with support from Giuliana K. Möller. In short, the designed primer sets were phosphorylated, annealed and cloned into pSpCas9(BB)-2A-GFP (Addgene plasmid

ID: 48138). Colonies were sequence verified and transfected into human U2-OS and HEK293-T-REx cancer cell lines. GFP-positive cells were monoclonalised using the FACsAria II flow cytometric cell sorter. We then verified the knockout of MacroD1 by immunofluorescence and western blotting using a monoclonal MacroD1-specific antibody (see antibodies in Appendix 4).

HEK293-T-REx cells were used to analyse the effect of MacroD1-deletion by CRISPR/Cas9-mediated knockout (Söllner et al., 2020b), both with proliferation assays and bioenergetic profiling. Cell lines are listed in table (Appendix 7).

U2-OS MacroD1 knockout and wild-type cells were used to assay MacroD1-dependent proliferation, bioenergetics, mitochondrial morphology as well as maintenance and regulation of the mitochondrial genome (Söllner et al., 2020b) (subchapter 3.3).

### 2.2.3 Stable expression of proteins in MD1KO & WT U2-OS cells

During the work for this thesis U2-OS cell lines, both wild-type and MD1KO, stably expressing mtGFP, MD1-GFP, G270E-GFP, POLG1-GFP, NGFP-POLG2, LIG3-GFP, and TWINKLE-GFP were generated.

We do this by subcloning the constructs into mammalian expression vectors containing the neo gene imbuing transformed cells with G418 resistance (plasmids in Appendix 6). Cells are transfected with the plasmids and maintained on selection media for three weeks prior to monoclonalisation. GFP-positive cells are chosen and monoclonalised using the FACsAria II flow cytometric cell sorter (section 2.2.17). After 2-3 weeks of additional continued culturing using selection media, transgene-expression levels are assayed by GFP-fluorescence and/ or antibody-staining using the BD LSRFortessa flow cytometric analyser (section 2.2.17).

U2-OS MacroD1 knockout and wild-type cells stably expressing MacroD1-GFP, MacroD1G270E-GFP and mtGFP were used to assay MacroD1-dependent proliferation, bioenergetics, mitochondrial morphology as well as maintenance and regulation of the mitochondrial genome (Söllner et al., 2020b) (subchapter 3.3).

#### 2.2.4 Routine maintenance of cells

To maintain cells, they have to be fed and passaged according to their specific needs. All cell lines discussed in this thesis are grown in T-75 culture flasks in DMEM (Gibco) supplemented with 10% FBS and 1% PS. While being cultured the cells are kept at 37°C in humidified air containing 5% CO<sub>2</sub>. Viability and cell size are routinely checked using the Vi-CELL XR Cell Viability Analyzer from Beckman Coulter. During this work we generated and used different cell lines, specific differences in cell line treatment are indicated in the cell line specific subsections.

##### 2.2.4.1 HEK293 -T-Rex

HEK293 -T-Rex cells are a HEK293 cell line (ATCC® CRL-1573™) derivative and exhibit a fibroblastic morphology. HEK293 -T-Rex cells cultured from the kidney tissue of a human female embryo and grow normally in a monolayer (Graham et al., 1977). Not passaging regularly and sparingly enough will however lead to clumping of the cells. We trypsinize HEK293 -T-Rex cells for 4 minutes at 37°C. This cell line adheres very loosely to the culturing dish – something that should be considered during washing with 1xPBS. We subsequently subcultivate at a ratio of 1:6 or 1:10 and incubate for 48 or 72 hours respectively.

To turn on the macrodomain or mtGFP expression in the modified HEK293-T-REx cells (section 2.2.1) we supplement the growth medium with 500 ng/ ml doxycycline for at least 20h.



#### 2.2.4.2 U2-OS

The human osteosarcoma cell line, U2-OS (ATCC® HTB-96™), was derived from a mesenchymal tumour of a young female (Ponten and Saksela, 1967). Cells in the U2-OS cell line are relatively large (with a median size of approximately 18 microns after trypsinization) and adhere tightly to cell culture flasks, making them an optimal model for imaging studies. We trypsinize U2-OS cells for 5-6 minutes at 37°C. We then sub cultivate at a ratio of 1:3 or 1:4 and incubate for 48 or 72 hours respectively.

The culture medium of the stably transfected U2-OS cell lines (section 2.2.3) is supplemented with 500 mg / mL G418 (Gibco) to ensure expression of the transgene.

#### 2.2.5 Counting cells

Cells are counted using an automated cell counter. The Vi-CELL XR Cell Viability Analyzer is a video imaging system with automated trypan blue exclusion protocol, thereby providing data on both percent viability and cell count.

Cells are washed in 1 x PBS, trypsinized and resuspended in fresh culture medium. Between 0.5 to 1 mL is aliquoted and put into the autosampler cue from where it is injected into the flow cell and measured. Average cell size, cell viability and cell count per mL are noted down. All further data analysis is performed using Microsoft Excel and GraphPad Prism6.

#### 2.2.6 Transfecting cells

Transfection is the hybrid term (transformation and infection) used for non-viral introduction of nucleic acids into eukaryotic cells.

In this work we use transfection to introduce plasmids containing transgenes for over-expression and small interfering RNAs to perform knock-down of specific endogenous genes. To perform a transfection, we first seed the cells 24 hours prior into the appropriate dishes at a concentration depending on experiment and cell line. The specifications, volumes, transfection reagents and other specifications are listed with the separate cell lines (see subsections 2.2.6.1, 2.2.6.2).

#### 2.2.6.1 HEK293 -T-Rex

Cells were seeded as specified in table 2.10 depending on the downstream application. Transient transfections to over-express exogenous transgenes were performed using Lipofectamine 2000 from ThermoFisher, in accordance with the manufacturer's instructions.

**Table 2.10: HEK293 -T-Rex cells are seeded in accordance with the specifications below:**

Culture dish	Volume per well	Cells seedet per well
15 cm dish	25 mL	$2.5 \times 10^6$
10 cm dish	10 mL	$1 \times 10^6$
6 well dish	2 mL	$2 \times 10^5$
8 well LabTekII	400 $\mu$ L	$0.4 \times 10^5$

#### 2.2.6.2 U2-OS

Cells were seeded as specified in table 2.11 depending on the downstream application. Both transient and stable transfections are performed using Xfect reagent from Clontech, for knock-down experiments we use siRNA SmartPools from Dharmacon and Lipofectamine RNAiMAX in accordance with the manufacturer's instructions.

**Table 2.11: U2-OS cells are seeded in accordance with the specifications below:**

Culture dish	Volume per well	Cells seedet per well
15 cm dish	25 mL	$6 \times 10^6$
10 cm dish	10 mL	$2 \times 10^6$
6 well dish	2 mL	$3 \times 10^5$
8 well LabTekII	400 $\mu$ L	$0.6 \times 10^5$

### 2.2.7 Cell-treatment with ddC

2',3'-dideoxycytidine (ddC) is a potent inhibitor of mtDNA replication (Brown and Clayton, 2002) and can be used in cell culture to deplete mtDNA in relation to nuclear DNA.

We perform ddC-treatment with 20  $\mu$ M ddC for either 0, 48 or 92 hours. During this treatment we change the medium, both for controls and for the treated samples, every 24 hours as ddC is not very stable and we want to ensure potent and prolonged mtDNA depletion.

### 2.2.8 Cell proliferation rate analysis

To determine proliferation rate, we seed cells at  $2 \times 10^5$  cells into 10cm cell culture dishes in normal growth medium. Cells are trypsinized, counted (subsection...) and replated every 48 hours for up to 8 consecutive days. After 8 days the dishes are too full and cell number will plateau out. We analyse proliferation in three biological replicates with two technical duplicates each.

### 2.2.9 Seahorse XF cell mito stress analysis

The Seahorse XF Cell Mito Stress analysis tests in this thesis are performed in a collaboration with Evangelia Tzika, who optimized seeding, culturing and injections for the U2-OS and HEK293-T-REx cell lines, performed measurements as well as analysis. Subsequent statistical significance tests on the data provided by her are performed using GraphPad Prism 6.

The experiments are carried out as follows, 40000 U2-OS cells and 60000 HEK293-T-REx cells are seeded into a XF96 cell culture microplate. On the day of the assay the cells are incubated for one hour in mito stress test assay medium (180  $\mu$ L/well) at 37°C in a CO<sub>2</sub> free incubator for 1h. The medium used for the assay is buffered at pH 7.4 at 37°C and contains 2mM L-glutamine, 1mM sodium pyruvate, 10 mM glucose and 5mM HEPES. Both the oxygen

consumption rate (**OCR**) and the extracellular acidification rate (**ECAR**) are measured in an XF96 extracellular flux analyser from Agilent Technologies. The basal measurements are performed prior to sequential injections of 10  $\mu$ M oligomycin A, 10  $\mu$ M FCCP and 5  $\mu$ M rotenone/antimycin A. After the addition of each ETC inhibitor, three measurements are performed to determine the ATP turnover, the spare respiratory capacity and the non-mitochondrial respiration effects. All measurements are performed in 6 biological replicates and 6 technical replicates per biological repeat.

#### 2.2.10 Subcellular localization - fractionation of cells

Fractionation of cells followed by SDS-PAGE and western blot detection (section 2.2.13) is one of the methods we use to determine subcellular localisation of proteins.

For this method we collect cells, wash in 1x PBS and pellet at 700 g for 3 minutes RT. Cell pellets are then resuspended in SEM buffer (70 mM sucrose, 220 mM mannitol, 20 mM HEPES, 5 mM EDTA, pH 7.4) containing complete EDTA-free protease-inhibitor cocktail (Roche). Cells are disrupted by suctioning through a 26G needle 5 times. The homogenate is centrifuged at 2000 g for 10 minutes at 4°C, the resulting pellet is the nuclear fraction. After two additional homogenization and centrifugation steps, the supernatant is centrifuged at 16500 g at 4°C for 10 min, the resulting pellet contains the mitochondria. The remaining supernatant is the cytosolic fraction.

The isolated mitochondria can subsequently be used for either immunoprecipitation of tagged-overexpressed proteins and downstream analysis by mass spectrometric analysis (sections 2.2.18 and 2.2.19), or to determine the subcellular localization of various proteins. To determine subcellular localization using these samples, the pellets are resuspended in 1x Sample buffer and analysed by SDS-PAGE and western blotting (section 2.2.13).

### 2.2.11 Submitochondrial localization: proteinase k treatment of mitochondria

To determine the submitochondrial localisation of proteins, subfractionation of isolated mitochondria by Proteinase K treatment is performed.

To perform Proteinase K digestion, we use 50 µg isolated mitochondria (section 2.2.10) per sample. The mitochondria are diluted in either 500 µL SEM or hypotonic buffer (20 mM HEPES pH 7.4) swelling and disrupting the mitochondrial outer membrane. Triton-X (TX-100) is added to a final concentration of 0.3% before addition of Proteinase K (100 µg/ mL). After 30 minutes on ice, all reactions are stopped for 5 minutes with 1 mM PMSF and 1x Complete Protease Inhibitor. The sample containing Triton-X is precipitated using TCA (final concentration 12%). The remaining samples are washed with SEM (1mM PMSF and 1x Complete Protease Inhibitor), subsequently pellets are TCA-treated with 12% TCA. All samples are stored for at least 30 minutes at 20°C prior to washing with ice-cold acetone. The pellets are resuspended in 1x Sample buffer and analysed by SDS-PAGE and western blotting (section 2.2.13).

### 2.2.12 Immunoprecipitation

In this thesis we performed immunoprecipitation of endogenous (e.g. PARP1) and over-expressed tagged-proteins (e.g. POLG1-GFP). When immunoprecipitating GFP-tagged proteins we use GFP-trap from Chromotek in accordance with the manufacturer's instructions. For Flag-tagged and endogenous proteins we use Dynabeads (Thermo Fisher) coupled to either anti-Flag or protein specific antibodies, according to manufacturer's instructions with slight modifications. In brief, per IP we couple 40 µL Dynabeads protein G or protein A (depending on species antibodies were raised in) to the specific antibody for at least 2 hours at 4°C. Antibody-bead complexes are washed thrice in 1x RIPA + PI Complete and incubate with cell lysates at 4°C, O/N, on a rotator. Antigen-antibody-bead complexes are washed 3 times in 1x RIPA + PI Complete. We incubate the complexes in Laemmli (- reducing agent) on ice for two minutes to elute the antigen from the antibody. Subsequently the

supernatant (antigen) is transferred to a new sample tube, reducing agent is added and sample is boiled and analysed by SDS-PAGE and western blotting (section 2.2.13).

### 2.2.13 Western blot analysis

We use sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) to separate proteins with different molecular mass. To do this we routinely lyse cells and/ or resuspend pellets in RIPA buffer and boil samples with reducing sample buffer. Samples are separated using either NuPAGE™ 4-12% Bis-Tris Protein Gels with NuPAGE™ MOPS SDS Running Buffer from ThermoFisher or self-cast SDS-PA gels ranging from 7.5% to 15%, depending on target protein size and spread (see Appendix 3 for buffers), we then transfer onto either polyvinylidene difluoride (PVDF) or a nitrocellulose membrane. After transfer the membrane is blocked for 40 – 60 minutes. For blocking we use either 3-5% Milk or BSA, depending on the antibody specifications. After blocking the membrane is incubated with primary antibody in either 3% milk or BSA, over-night in the cold-room. We wash 3x 20 min in 1xTBST on an oscillator prior to secondary-antibody incubation for 40-60 minutes in either 3% milk or BSA. We rinse the membrane 6 x with 1x TBST then wash for an additional 20 minutes before signal detection.

For detection, we use secondary antibodies conjugated to horseradish peroxidase (HRP) and SuperSignal West Dura Extend Duration Substrate from ThermoScientific. The different antibodies are noted in the results section and listed in the antibody table (Appendix 4).

To strip membranes, we use either mild or harsh stripping buffer depending on the strength of the signal we hope to strip and the antibodies we intent to use afterwards. When using the mild stripping buffer, we cover the membrane in stripping buffer and incubate 2 x using fresh buffer each time at room temperature for 5-10 minutes. We wash the membrane 2 x for 10 minutes with 1xPBS, followed by 2x with 1x TBST for 5 minutes each. After these washing steps the membrane is ready for re-blocking and re-probing. When performing a harsh strip, we warm the harsh stripping buffer to 50°C in a small plastic box with a lid then add the

membrane. The membrane is incubated at 50°C for up to 45 minutes with some agitation before the membrane is rinsed under running water, then washed 2x with 1x TBST for 5 minutes each. After these washing steps the membrane is ready for re-blocking and re-probing (for buffers see Appendix 3).

#### 2.2.14 Far western blotting

In far western blotting proteins are first separated by SDS-PAGE and transferred onto a membrane, much like in western blot analysis. However, instead of the initial primary antibody incubation, the membrane is first probed with a non-antibody tagged - protein. This method is widely used to study protein- protein interactions (Wu et al., 2007), but can also be employed to analyse whether a protein has specific modifications, such as MAR- and PARylation (Khadka et al., 2015). In our lab, this method was established by Aurelio P. Nardoza, using the purified V5-tagged MacroD1 and TARG1 macrodomains as probes.

We use this method to detect whether specific proteins are ADP-ribosylated. In this manner we can detect ADP-ribosylation of either endogenous or overexpressed bait proteins. To verify the modification status of our specific proteins we start by immunoprecipitating an either highly expressed endogenous (e.g. PARP1) or an overexpressed and tagged protein (e.g. TFAM-GFP, POLG1-GFP). After immunoprecipitation (section 2.2.12), SDS-PAGE and transfer onto a nitrocellulose membrane (section 2.2.13), the membrane is blocked for 1 hour in protein binding buffer and subsequently incubated O/N at 40C in 5 mL protein binding buffer with 30 µg purified V5-tagged macrodomain protein. On the next morning the membrane is washed and incubated with primary (anti-V5) and secondary (anti-mouseHRP) antibody following the standard western blot procedure. To ensure that the bands observed during the Far western approach are indeed the proteins of interest, the membrane is stripped (using the harsh stripping buffer method) and re-incubated with antibodies specific for the immunoprecipitated substrate.

### 2.2.15 Mitochondrial DNA immunoprecipitation (MIP) assay

Chromatin immunoprecipitation (**ChIP**) assays are used to investigate the interaction between proteins and DNA in the cell nucleus. In the method described here, Mitochondrial DNA Immunoprecipitation (**MIP**) assay, we use the same principles employed in standard ChIP but instead of looking for genes / DNA fragments embedded in the nuclear genome, we use primers specific for mtDNA. In this way we can determine relative mtDNA - protein interactions in different cell lines.

In the MIP assay,  $1 \times 10^7$  cells per sample/ cell line are crosslinked with 1% formaldehyde in 1xPBS at room temperature for 13 min. The reaction is stopped by incubation with 125 mM Glycine for 10 min. The cell pellets are washed twice by centrifugation in ice-cold 1x PBS at 4°C. The cell pellets are lysed in Lysis buffer for 5 minutes and sheared by sonication in a Covaris S220, thereby generating DNA-protein fragments with an average length of 200-500 bp. The supernatant is diluted 1:10 with RIPA Buffer, at this point the aliquots can be shock frozen and stored at -80°C or used directly in immunoprecipitation. We use 1 aliquoted sample ( $1 \times 10^6$  cells) per IP, the protein of interest with attached mtDNA is precipitated using specific antibodies coupled to Dynabeads (as outlined in section 2.2.12). After immunoprecipitation the samples are treated with RNase A for 30 minutes and incubated with Proteinase K O/N at 67°C at 1000 rpm on a thermomixer. The DNA is purified using the ChIP DNA Clean and Concentrator kit from Zymo Research. The qPCR is performed with the same primer sets and settings as indicated in the quantitative PCR section, the data is normalized to percent input.

### 2.2.16 Immunofluorescent staining and confocal imaging

Cells are routinely either fixed or imaged live 24 - 48 hours after seeding or 24 hours after transient transfection (section 2.2.6). To visualize DNA, mtDNA and mitochondria, growth medium is aspirated from the Lab-Tek slides and replaced with fresh medium containing 0.3 g/mL Hoechst 33342 (ThermoFisher) and 100 nM MitoTracker Red (M7512 - ThermoFisher)

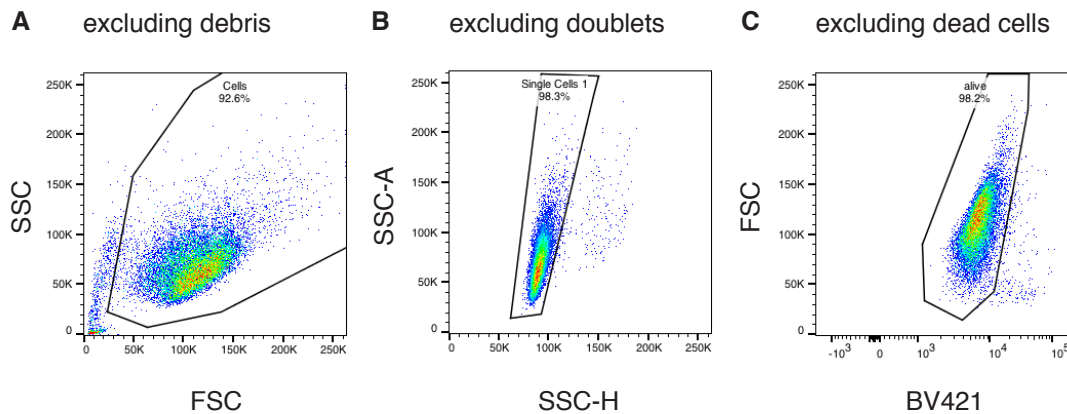


for 15 minutes at 37°C in humidified air containing 5% CO<sub>2</sub>. For live cell imaging, growth medium replaced with CO<sub>2</sub>-independent Leibovitz L-15 medium, supplemented with 10% fetal bovine serum, 2 mM glutamine, 50 U/ml penicillin, and 50 mg/ml streptomycin immediately prior to imaging. For fixation, the cells are washed twice in warm 1xPBS, then fixed in an ice-cold methanol: acetone (70:30) solution for 10 minutes at -20°C. The cells are washed twice in 1xPBS and blocked for 30 minutes in 3% BSA in 1xPBS prior to antibody incubation. Confocal images are obtained on a Zeiss AxioObserver Z1 spinning-disk microscope and Yokogawa CSU-X1 scan head using a Plan Apo 100× oil objective and a CMOS Orca Flash 4.0 camera by Hamamatsu.

#### 2.2.17 Immunofluorescent staining and flow cytometric analysis

Cells are either fixed, stained and analysed or live sorted and analysed 24 hours after seeding. To stain and quantify mitochondria, growth medium is replaced with fresh medium containing either 50 nM MitoTracker Red or MitoTracker Deep Red FM (M7512/ M22426) for 15 minutes at 37°C in humidified air containing 5% CO<sub>2</sub>. For cell sorting and analysis, cells are harvested and washed twice in warm 1xPBS. To analyse and sort live cells, these are incubated in FACS sorting buffer with cell viability stain SYTOX Blue (Thermo Fisher) and filtered through a 35 µm cell strainer cap for FACS tubes just prior to sorting/analysis. To fix cells we resuspend cell pellets in 30 µL FACS sorting buffer, then add 600 µL ice-cold methanol: acetone (70:30) solution for 10 minutes at -20°C. Subsequently, cells are washed twice in FACS sorting buffer and left to block in the same buffer for at least 30 minutes prior to antibody incubation.

During data acquisition and analysis, the cells were gated to exclude cell debris, doublets and dead cells (see figure 2.1).



**Figure 2.1: Gating example for flow cytometric analysis.** A) Gate identifying cells. B) Gating out doublets. C) Gate to isolate only living cells in population by viability dye.

In support of this thesis three different flow cytometric cell sorters and analysers have been used: i) the FACSaria II flow cytometric cell sorter for monoclonalisation and sorting of both stably transfected and CRISPR/Cas9-mediated MacroD1 knockout cell lines ii) the BD LSRFortessa to analyse monoclonalised cell lines, determine their relative size, mitochondrial load, as well as endogenous and transgene expression levels of either fixed or living cells iii) and the Amnis Image Stream XMark II for flow cytometric imaging analysis in the de novo mtDNA synthesis assay.

### 2.2.18 Immunoprecipitation, MS sample preparation and analysis

Fabian Hosp performed immunoprecipitation of proteins from isolated mitochondria, MS sample preparation and initial analysis as part of a collaboration with Aurelio P. Nardoza at the Max Planck Institute of Biochemistry in the group of Matthias Mann.

Each well of a GFP-multiTrap plate (ChromoTek) is washed three times with 200  $\mu$ L buffer 1 and incubated with the cleared mitochondrial lysate by shaking at 100 rpm for 60 min at 4°C. Next, each well is washed 2 x 200  $\mu$ L buffer 2 and 4 x with 200  $\mu$ L buffer 1, prior to incubation with 25  $\mu$ L elution buffer at room temperature for 90min. Next, the resulting peptides are

alkylated with 25  $\mu$ l alkylation buffer and subsequently washed with 50  $\mu$ l urea buffer for 10 min each. The supernatants from the elution, alkylation and washing steps are collected after each step and combined in a clean 96-well plate. This plate is incubated overnight at room temperature to ensure a complete digest. The next morning, the digestion is stopped by addition of 10  $\mu$ l 10% TFA per well. The acidified peptides are purified on StageTips (Rappsilber et al., 2003).

MS analysis is performed using Q Exactive Plus mass spectrometers (Thermo Fisher Scientific, Bremen, Germany) coupled online to a nanoflow ultra-high-performance liquid chromatography (UHPLC) instrument (Easy1000 nLC, Thermo Fisher Scientific). Peptides are separated on a 50-cm-long (75 mm inner diameter) column packed in house with ReproSil-Pur C18-AQ 1.9mm resin (Dr. Maisch). Column temperature is kept at 50°C by an in house-designed oven with a Peltier element. Peptides are loaded with buffer A and eluted with a nonlinear gradient of 5%–60% buffer B at a flow rate of 250 nL/min and eluted with a linear 100 min gradient. The survey scans (300–1,650 m/z, target value = 3E6, maximum ion injection times = 20 ms) are acquired at a resolution of 70,000, followed by higher-energy collisional dissociation (HCD)-based fragmentation (normalized collision energy = 25) of up to 10 dynamically chosen, most abundant precursor ions. The MS/MS scans were acquired at a resolution of 17,500 (target value = 1E5, maximum ion injection times = 120 ms). Repeated sequencing of peptides was minimized by excluding the selected peptide candidates for 20 s.

All MS data is further analysed using the MaxQuant software version 1.4.3.14 (Cox and Mann, 2008) with a false discovery rate (FDR) cut-off threshold of 1% for protein and peptide spectrum matches. Peptides are required to have a minimum length of 7 amino acids. MaxQuant is used to score fragmentation scans for identification based on a search with an initial allowed mass deviation of 0.5 Da, the allowed fragment mass deviation is 20 ppm. Fragmentation spectra are identified using the UniprotKB Homo sapiens database (UniProt, 2014), combined with 262 common contaminants by the integrated Andromeda search engine (Cox et al., 2011). Enzyme specificity is set as C-terminal to arginine and lysine, also allowing cleavage before proline, and a maximum of two missed cleavages.

Carbamidomethylation of cysteine is set as fixed modification and N-terminal protein acetylation as well as methionine oxidation as variable modifications. Both 'label-free quantification (MaxLFQ)' with a minimum ratio count of 2 and 'match between runs' with standard settings are enabled (Cox et al., 2014). Basic data handling, normalization, statistics and annotation enrichment analysis is performed with the Perseus software package (Tyanova et al., 2016). Missing values are imputed with values representing a normal distribution (generated at 1.8 standard deviations of the total intensity distribution, subtracted from the mean, and a width of 0.3 standard deviations).

### 2.2.19 Mass spectrometry data analyses

The data sets received from Aurelio P. Nardoza and Fabian Hosp from the Immunoprecipitation, MS Sample Preparation and Analysis are further filtered and analysed for this thesis with the python programming support from Benjamin E. Oliver. In brief, to filter proteins for further analysis and visualization Benjamin wrote a python program which generates a table mapping every protein to its associated gene. Using the UniProt API (UniProt, 2019), the associated FASTA files for each protein are retrieved and used to generate files suitable for upload to the batch endpoint of the TargetP webpage (<http://www.cbs.dtu.dk/services/TargetP-1.1/index.php>). If any of the protein identifiers of a gene indicate mitochondrial targeting, the corresponding protein in the table is marked as being a putative candidate for mitochondrial localization. Using this information in conjunction with extra constraints of fold change and p value the data set is filtered focusing only on proteins that may be of interest. The lists generated using this program can further be used in gene ontology analysis (Ge et al., 2020) or directly visualized in volcano plots.

Data files and python scripts both for MTS prediction analysis as well as volcano plot generation can be found on GitHub under <https://github.com/Flavinoid/mts>. The programmes and files published there are further explained in Appendix 9.

### 2.2.20 POLG2 assay for relative quantification of active mtDNA replication forks

The POLG2 assay is a novel replication assay for mtDNA and was conceived, invented, established and tested during the work for this thesis. It is described further in the first part of this thesis (subchapter 3.1), as well as in two manuscripts accompanying this thesis (appendices 1 and 2). In brief, POLG2 is the accessory subunit of POLG1 and has been shown to localize to active, DNA-synthesizing, mitochondrial replication forks (Lewis et al., 2016; Young et al., 2015). To generate the cell-lines which are the main tool in this assay we stably integrated NGFP-POLG2 (kind gift from W. Copeland) into U2-OS cells (see section 2.2.3). With these cell lines we are able to use the BD LSRFortessa flow cytometric analyser to comparatively quantify active mitochondrial replication forks. Samples are prepared as described in the section on Immunofluorescent staining and flow cytometric analysis and measured using the BD LSRFortessa (see section 2.2.17). Using this assay, we can measure relative active mtDNA replication forks and in parallel determine changes to both cell size and mitochondrial content.

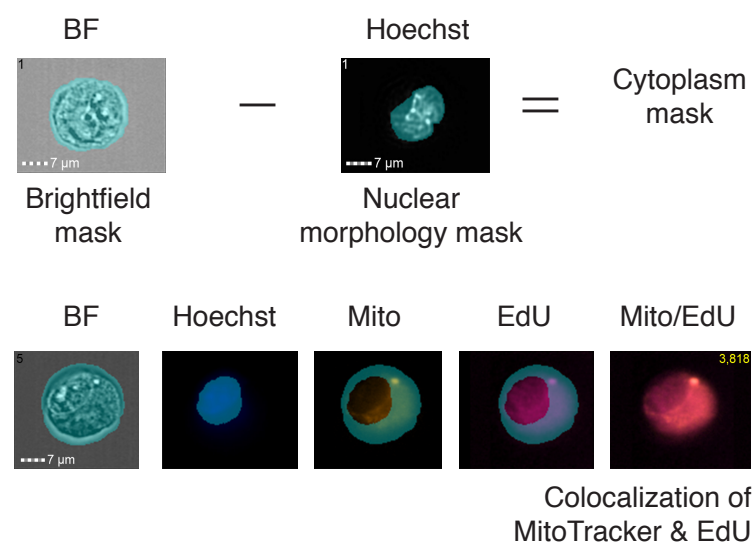
### 2.2.21 Flow-cytometric imaging to analyse *de novo* mtDNA synthesis

To analyse *de novo* mtDNA synthesis we use EdU-incorporation (Click-iT EdU AlexaFluor 647 labelling kit) coupled with flow cytometric imaging using the ImageStream MarkII imaging cytometer from Amnis.

EdU-labelling is performed in accordance with the manufacturer's instructions, with minor deviations. In brief,  $1 \times 10^7$  cells are seeded into 15 cm cell culture dishes in order to achieve 80% confluency upon labelling time. After 24 hours, cells are incubated in growth medium supplemented with 7 mM aphidicolin (A4487, Sigma) to stall nuclear DNA synthesis. After 4 hours the medium is aspirated, and cells are incubated in fresh medium containing 50 mM EdU-pulse reaction. After one-hour mitochondria are labelled for 15 minutes. Next, cells are harvested and fixed as described for general flow cytometric analysis (section 2.2.17).

Subsequently we perform the Click IT reaction using the fluorescent picolyl azide dye and stain the nuclear DNA by incubating the fixed cells with 0.3 g/mL Hoechst 33342 for 30 minutes.

All images are detected via the ImageStream MarkII imaging cytometer at 60X magnification using the IDEAS software. We assign Channel 01 to brightfield, channel 02 to GFP, channel 04 to MitoTracker, channel 07 to Hoechst and channel 11 to the EdU signal. The compensation of the data collected is performed by using single colour controls. To quantify the EdU signal and thereby de novo mtDNA synthesis, we initially assign gates for focused single cells. Then we create first a brightfield mask for the whole cell, a morphology mask for the nucleus in channel 07 and subtract the area of the nuclear mask from the cell mask –thereby generating the cytoplasm mask. We then make a colocalization between channel 04 and the cytoplasm mask and in that way determine the intensity of EdU-fluorescence within this colocalization, and in so doing within mitochondria (see Figure 2.2).



**Figure 2.2: Gating for flow cytometric analysis in imaging flow cytometric assay in the IDEAS software.** Subtraction of nucleus from whole cell is used to generate a cytoplasm mask. Co-localization between MitoTracker Red staining and EdU-staining, within the cytoplasm mask. The subsequently measured EdU-fluorescent intensity from the colocalization is the EdU-fluorescent intensity within mitochondria.

## 2.3 Fly work

For this thesis additional organismal analysis of a putative mitochondrial macrodomain protein was performed in the model organism *Drosophila melanogaster*.

Three different fly strains are used for the work presented in this thesis. Wild-type 2u flies, a MacroD1-like macrodomain RNAi-induced knockdown strain and a Mef2-GAL4 strain driving the knock-down in the RNAi-line from the myoblast stage. RNA-interference is performed in all muscle using the Mef2-GAL4 driver line. The wild-type 2u line crossed with the Mef2-GAL4 driver serves as a control. For further specifications of these fly lines see Appendix 8.

### 2.3.1 Routine maintenance of fly lines

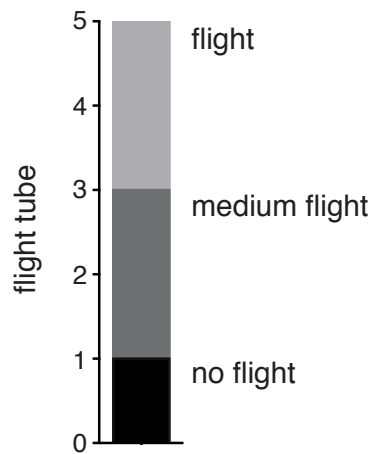
Fly stocks are kept at 25°C, in standard medium. All crosses are performed with virgin females from the driver and wild-type line and young males from the RNAi-line, with all required controls. Fly stocks used for this thesis can be found in Appendix 8.

### 2.3.2 Immunofluorescent staining and confocal imaging of *Drosophila melanogaster* flight muscle

To image the flight muscle thoraxes are fixed for 20 minutes using 4% PFA in 0.5% PBS-T and washed in PBS-T. Thoraxes are cut sagittally with a scalpel into hemithoraces prior to staining. Hemithoraces are incubated with rhodamine phalloidin (Molecular Probes, 1:500) in PBS-T for 3 h at RT. Hemithoraces are washed three times in 0.5% PBS-T and mounted in Vectashield containing DAPI. Hemithoraces are imaged on a Zeiss AxioObserver Z1 spinning-disk microscope and Yokogawa CSU-X1 scan head using a Plan Apo 100× oil objective and a CMOS Orca Flash 4.0 camera by Hamamatsu.

### 2.3.3 Flight test

Flight tests are performed using a 1 m high, plexiglass, flight chamber, subdivided into 5 equally large sections (see Figure 2.3). To test flight capacity young males of the desired genotype are selected 24 h prior to testing, while anesthetized. The next day, flies are released from the top of the cylinder. Their flight capacity is scored depending of where within the cylinder they land. The bottom zone (zone 1) is referred to as flightless, zones 2 and 3 are medium fliers, while flies landing in zones 4 and 5 are good fliers (see Figure 2.3). Flight experiments were performed on at least 100 flies per genotype in at least 3 separate experiments.



**Figure 2.3: Tube for flight test.** The 1 m long plexiglass tube is divided into 5 x 20 cm long sections 1 to 5. Flies falling into section 1 are classified as no flight, sections 2 and 3 are medial fliers, while flies that land in sections 4 and 5 are classified as good fliers.



## 2.4 Statistical & biostatistical analyses

### 2.4.1 Multiple sequence alignments

We use reference DNA and protein sequences downloaded from Uniprot and NCBI. As human mitochondrial genome sequence reference NC\_012920 is used. Multisequence alignments are performed with the *in silico*, cloning and alignment tool CLC Main Workbench from Qiagen.

### 2.4.2 Sequence searches

To predict the presence of a mitochondrial targeting sequence the protein amino acid sequences is retrieved from UniProt and imported into iPSORT (Bannai et al., 2002), Mitoprot II (Claros and Vincens, 1996) and Target P1.1 (Emanuelsson et al., 2000).

The sequence searches performed in and for the MTS predictions are further elucidated in Appendix 9.

### 2.4.3 Calculations

Data analysis is performed using Microsoft Excel and GraphPad Prism6, it is expressed as mean  $\pm$ SEM. Unless otherwise specified, normally distributed data is analysed using one-way ANOVA with Turkey's test or a two-tailed unpaired t-test if only two samples were compared. For non-parametric data we used Kruskal-Wallis with Dunn's test. P-values are displayed as: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ .

## 3. Results

In subchapter 3.1 the novel POLG2-based mtDNA replication assay is introduced and tested. Subchapter 3.2 centres around the mass spectrometric analysis of the MacroD1- interactome and ADP-ribosylome. While the impact MacroD1 and mitochondrial ADP-ribosylation have on a cellular and organismal level are discussed in subchapter 3.3.

### 3.1 Mitochondrial DNA: a novel fluorescence-based replication assay

Dysregulation of the mitochondrial genome is implicated in a variety of neurological diseases and mitochondrial myopathies. The aetiologies behind these diseases are often either mutations in the mitochondrial genome, or mutations of the nuclear genome sequences coding for factors involved in the maintenance of mtDNA. The regulation of replication, and thereby of copy number, can be modulated and controlled at various points; at mtDNA accessibility, at transcription initiation, and at the switch between transcription initiation and replication. How exactly mtDNA replication is modulated and regulated however, is largely unknown. With the POLG2-based mtDNA replication assay we have devised a novel method to facilitate the study of mtDNA replication *in vivo* using a flow-cytometric approach. Thereby, enabling an automatable approach using vast sampling sizes and drastically minimising sampling bias. The focal point of our new mtDNA replication assay is the requirement for the homodimeric POLG2 subunit during mitochondrial genome replication by POLG1 (Fan et al., 2006; Young et al., 2015). MtDNA nucleoids are distributed throughout the mitochondrial network, some nucleoids undergo replication and some transcription, how this is organized is not entirely clear. What is clear is that all nucleoids with active replication forks contain POLG2 and that non-replicative nucleoids do not (Lewis et al., 2016; Young et al., 2015). Furthermore, although POLG2 is essential in mtDNA replication, stable insertion of NGFP-POLG2 into mammalian cells does not alter their bioenergetic profile (Young et al., 2015), nor their mtDNA copy number (Lewis et al., 2016). Based on these premises, we decided to generate cell lines stably expressing NGFP-POLG2 (section 3.1.1), which we hypothesized could be used to monitor ongoing mtDNA replication changes in living and in fixed cells. We

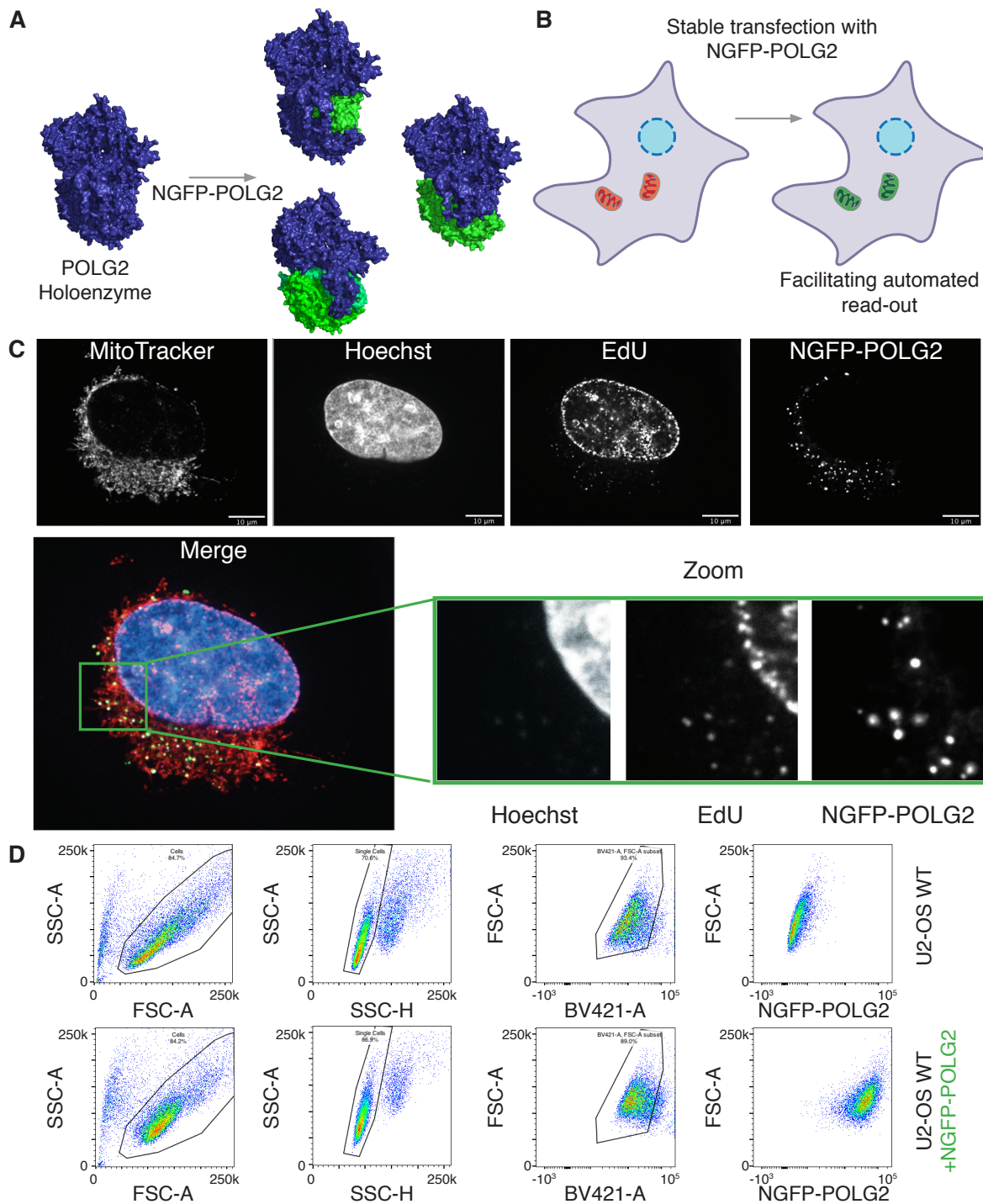
tested our tool using pharmacological interruption of replication (section 3.1.2), as well as by overexpression of various factors reported to either deplete (section 3.1.3) or increase (sections 3.1.4 and 3.1.5) mtDNA synthesis.

The results presented in this subchapter are also worked into a methods manuscript entitled *A Novel Fluorescence-based Assay to Monitor Mitochondrial DNA Replication in Vivo* (Söllner et al., 2020a). We used the method in the manuscript *Protein ADP-ribosylation Regulates Human Mitochondrial DNA Replication* (Söllner et al., 2020b), which is currently under revision at Cell Reports.

### 3.1.1 The synthesis and characterization of monoclonalised NGFP-POLG2 U2-OS cell lines

The first step towards the mtDNA replication POLG2-Assay has to be the cell line tool in which different mtDNA replication conditions can be monitored. For this we generated monoclonalised U2-OS cell lines stably expressing NGFP-POLG2 (Figure 3.1 A, B), as described in the methods sections 2.2.3 and 2.2.17.

NGFP-POLG2 has been reported to form distinct foci within mitochondria, which correlate nicely with EdU-incorporated newly-synthesized DNA foci (Lewis et al., 2016; Young et al., 2015). These distinct NGFP-POLG2 foci can also be observed in the stable NGFP-POLG2 expressing cell lines generated by us (Figure 3.1 C). The NGFP-POLG2 foci in the U2-OS WT + NGFP-POLG2 cl.1 cell line are situated in the mitochondrial network and correlate with the newly synthesized mtDNA as visualized by EdU-incorporation and Click-iT chemistry (Figure 1.3 A zoomed in area).



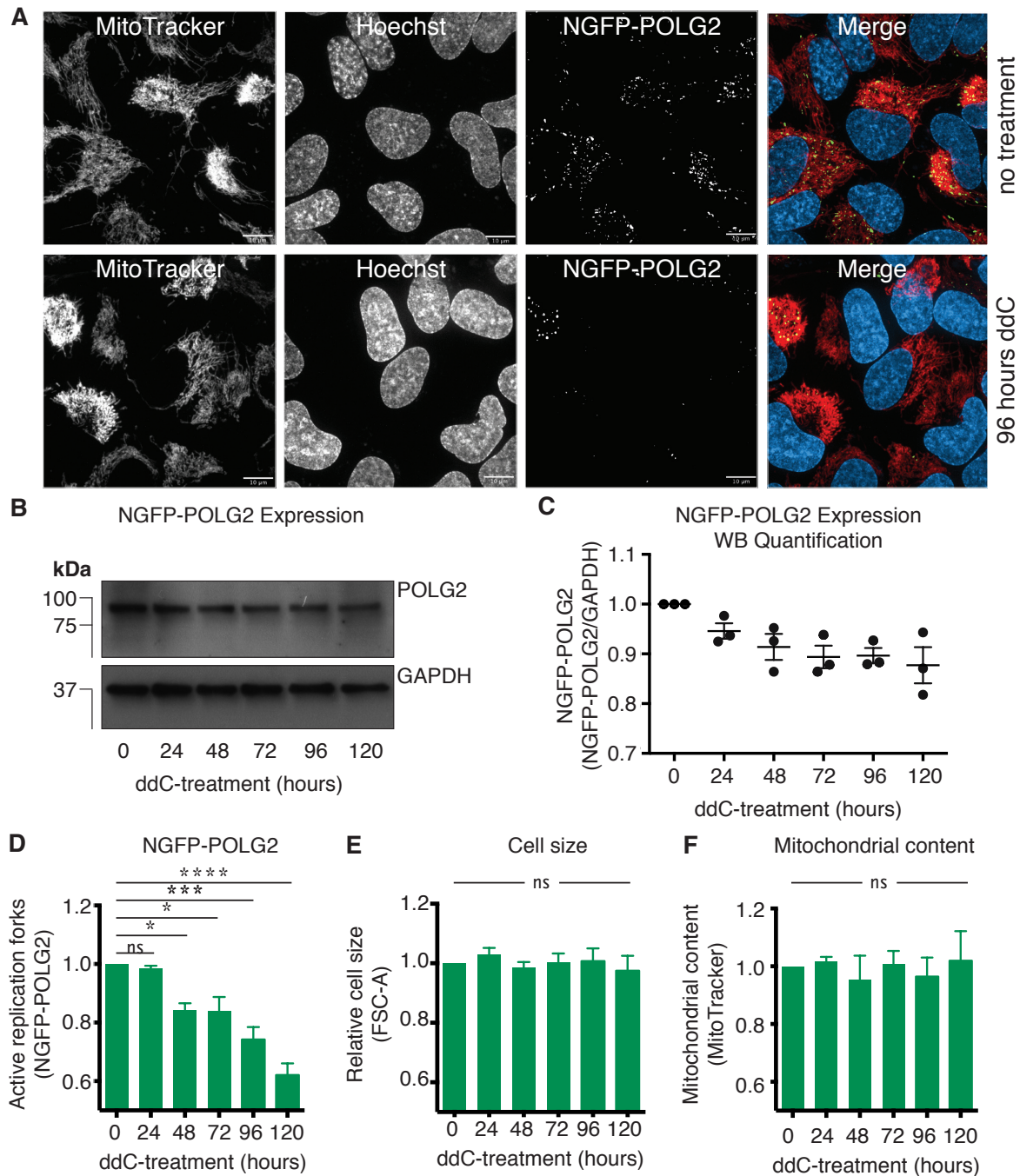
**Figure 3.1: Stable NGFP-POLG2 Expressing U2-OS Cell Lines for POLG2 Assay.** A) NGFP-POLG2 mutant can replace endogenous POLG2 without affecting catalytic activity or function. Shown on the 3IKM crystal structure of the POLG holoenzyme. B) Schematic representation of the POLG2 mtDNA replication assay tool. U2-OS cell line stably expressing NGFP-POLG2 facilitate automated read-out of changes in active mtDNA replication forks. C) Immunofluorescent staining depicting the U2-OS + NGFP-POLG2 cl.1 cell line. In the merged image mitochondria are red, DNA is in blue, newly synthesized DNA is in magenta and NGFP-POLG2 is green. The zoomed in area shows that mtDNA, EdU and NGFP-POLG2 foci in mitochondria correlate. D) Flow cytometric analysis showing the U2-OS WT and U2-OS WT + NGFP-POLG2 cl. 1 cell lines.

During the work for this thesis we tested 4 of the monoclonalised cell lines; clones 1,2,3 and 10. We observed no significant clonal effects. We performed most of the experiments using monoclonal cell lines 1 and 10. To further ensure that the green-fluorescing protein in IF is NGFP-POLG2, cell lysates were analysed by SDS-PAGE and western blot detection (not shown here). Using an anti-GFP antibody we were able to detect a band migrating relative to 80 – 85 kDa in the U2-OS + NGFP-POLG2 cell lines, which fits well with the predicted sizes for POLG2 (~55 kDa) plus GFP (~27 kDa).

Based on the IF and Flow cytometric analysis (Figure 3.1), we have been successful in generating stable cell lines which not only express NGFP-POLG2 but also localise the protein to the appropriate subcellular compartment, the mitochondria. Furthermore, we were able to demonstrate that the fusion protein, NGFP-POLG2, has not lost its functionality in our experiments and still colocalizes with active replication forks in our cell lines (Figure 3.1 C).

The premise behind our approach is that a reduction in mtDNA replication will lead to a reduction in NGFP-POLG2 foci and with that a reduction in green fluorescent intensity. To automatize our assay, we decided to measure green fluorescent intensity per cell using flow cytometric analysis. With this approach, we will be able to not only measure relative active replication forks per cell, but also monitor other parameters, such as relative cell size and mitochondrial load by forward scatter (FSC) and MitoTracker fluorescent intensity. Using the flow cytometric analyser, it will be possible to test many thousand cells, either fixed or living, simultaneously while excluding cell debris, doublets and dead cells. Here demonstrated on the U2-OS WT versus the U2-OS WT + NGFP-POLG2 cl.1 cell line (Figure 3.1 D).

We additionally generated stable NGFP-POLG2 -expressing cell lines in the MacroD1 knock out cells, discussed in subchapter 3.3 and in our Manuscript (Söllner et al., 2020b).



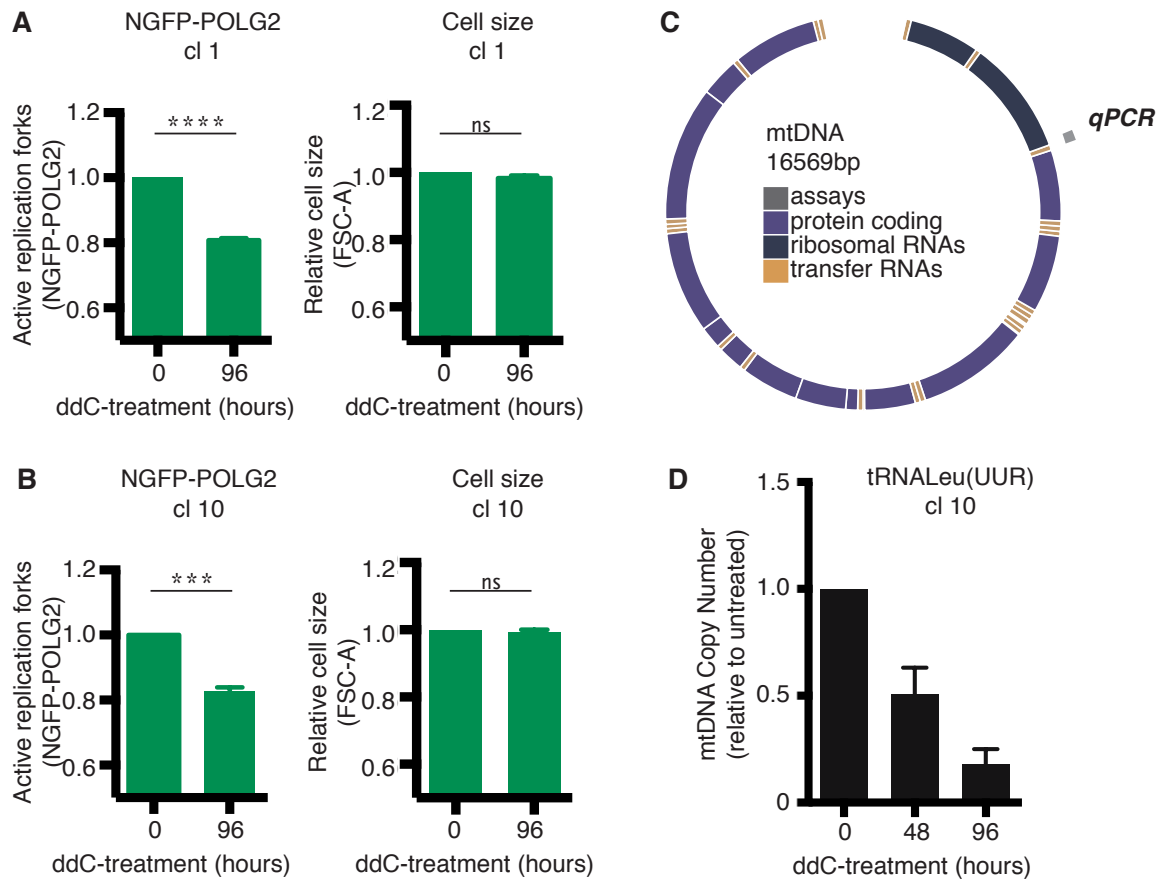
**Figure 3.2: Pharmacological inhibition of mtDNA replication.** A) IF using the U2-OS WT + NGFP-POLG2 cl. 1 cell line with and without 96 h of ddC-treatment. Observable reduction of NGFP-POLG2 foci in ddC-treatment. In the merged images: nuclei are stained in blue, NGFP-POLG2 is green and MitoTracker in red. B-F) are performed in the U2-OS WT + NGFP-POLG2 cl. 10 cell line. B) Western blot analysis of the cell line with 0, 24, 48, 96 and 120 hours of 20  $\mu$ M ddC. C) WB quantification of 3 biological replicates. NGFP-POLG2 signal is normalized to GAPDH, expression levels are shown in relation to untreated cells. D-F) Flow cytometry of fixed cells with increasing length of ddC-treatment. Experiment is performed in 3 biological replicates, data is displayed as mean  $\pm$  SEM. Statistical analysis is performed using ordinary one-way ANOVA with Dunnett's multiple comparisons test, \* $p < 0.05$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ .

### 3.1.2 ddC-treatment leads to a reduction of NGFP-POLG2, while cell size and mitochondrial content remain unchanged

To test our tool, we decided to expose our cell lines to conditions known to alter mtDNA replication. A potent inhibitor of mtDNA replication frequently used in control experiments is the POLG inhibitor 2',3'-dideoxycytidine (ddC) (Brown and Clayton, 2002).

After 96 hours of 20  $\mu$ M ddC-treatment we saw a clear reduction of NGFP-POLG2 foci in IF experiments (Figure 3.2 A). As an orthologous experiment and to ensure that there was in fact a depletion of the NGFP-POLG2 protein upon ddC-treatment, we also performed western blot analysis and saw a clear decrease in the NGFP-POLG2 signal, both in a single blot (Figure 3.2 B) and upon quantification of three separate biological replicates (Figure 3.2 C). Finally, we performed the test assay on the BD LSRFortessa flow cytometer, for this experiment we used the fixed U2-OS WT + NGFP-POLG2 cl. 10 cell line. As expected, we observed that increasing lengths of ddC-treatment leads to a significant decrease in NGFP-POLG2-signal per cell (Figure 3.2 D). Mitochondrial content and cell size, however, were unaffected (Figure 3.2 E and F).

We repeated the experiment, using unfixed – living cells, in both the clone 1 and 10 cell lines. Again, we observed a significant decrease in NGFP-POLG2 fluorescent intensity per cell with no changes to the overall cell size (Figure 3.3 A and B). To ensure that the decrease in fluorescent signal from the NGFP-POLG2 cell line (Figure 3.2 A and D and Figure 3.3 A and B), along with the decrease of the protein in the western blots (Figure 3.2 B and C) are in fact due to inhibited mtDNA replication we decided to also perform a mtDNA/nDNA ratio qPCR to assess relative mtDNA copy number in ddC-treated cells compared to untreated samples. For the qPCR we used a mitochondrial primer set for the tRNA<sup>Leu(UUR)</sup> area (Figure 3.3 C) and a nuclear primer set in the beta2-microglobulin gene. As anticipated, mtDNA content was steadily reduced upon ddC-treatment (Figure 3.3 D). Thereby showing that our assay is a robust readout for mtDNA replication, not subject to clonal effects and can be performed using either fixed or unfixed cell samples.



**Figure 3.3: Pharmacological inhibition of mtDNA replication measured by flow cytometry on living cells.** A and B) Flow cytometric analysis of living U2-OS WT + NGFP-POLG2 cl. 1 and U2-OS WT + NGFP-POLG2 cl. 10 cells, respectively. Cells are either not treated or treated for 96 hours with 20  $\mu$ M ddC. NGFP-POLG2 fluorescence diminishes significantly, while cell size remains unchanged. Experiment is performed in 3 biological replicates, data is displayed as mean  $\pm$  SEM. Statistical analysis is performed using a two-tailed, unpaired, t test, \*\*\* $p$ <0.001, \*\*\*\* $p$ <0.0001. C) Schematic of mitochondrial DNA molecule. In grey is the amplicon amplified during qPCR, inside the tRNA<sup>Leu</sup>(UUR) area. D) Comparative analysis of mtDNA copy number, measured by qPCR, of untreated, as well as 48 and 96 hours of ddC-treatment. Experiment is performed in 3 biological replicates. Data is displayed as mean  $\pm$  SEM.

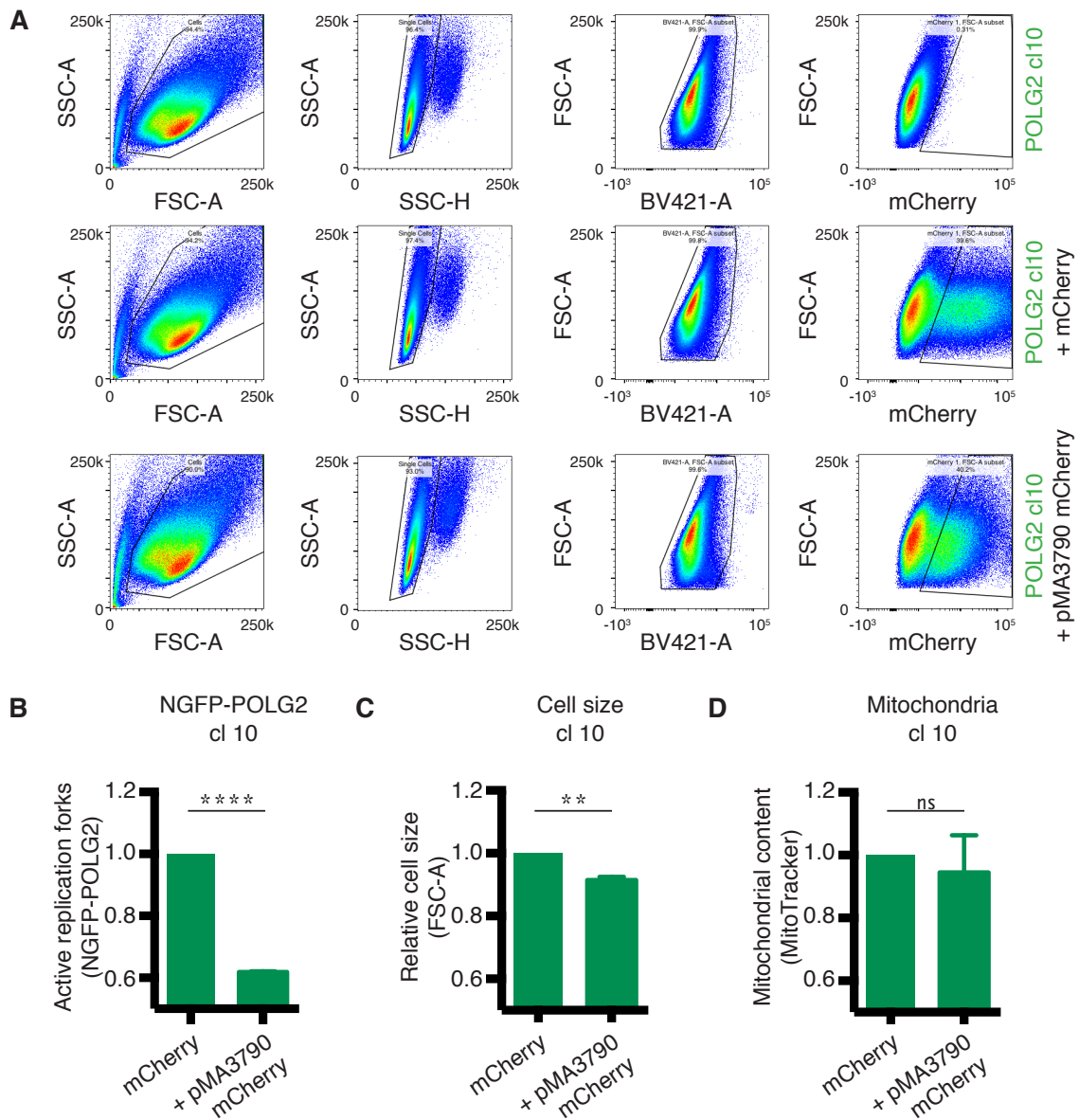


### 3.1.3 Overexpression of the exonuclease pMA3790-mCherry in mitochondria leads to a reduction in NGFP-POLG2 fluorescence

Exonucleases synthetically targeted to the mitochondrion have been shown to deplete mtDNA (Spadafora et al., 2016). We cloned the mutant Y147A human uracil-N-glycosylase (mUNG1) protein tool (Spadafora et al., 2016) into an open reading frame with mCherry.

Overexpression of pMA3790-mCherry (mUNG1) should deplete mtDNA content in the transfected cells, thereby depleting the green fluorescence from NGFP-POLG2. Due to the mCherry-tag, the transfected subpopulations can be separated from an untransfected cell pool – making it possible to assess green-fluorescence solely in cells containing the mitochondrially targeted exonuclease. As a control we used cells transfected with the empty mCherry vector. During analysis we first excluded cell debris and doublets, then we used viability dye to determine living cells. Subsequently we gated cells expressing mCherry and analyzed GFP, MitoTracker and FSC fluorescence in this subpopulation of cells (Figure 3.4 A).

As anticipated, we observed a significant decrease in green fluorescence from the subpopulation of cells expressing the mitochondrially targeted exonuclease, meaning we are able to observe the reduction of mtDNA by measuring a reduction in active replication forks, while simultaneously also measuring relative changes to cell size and mitochondrial content.



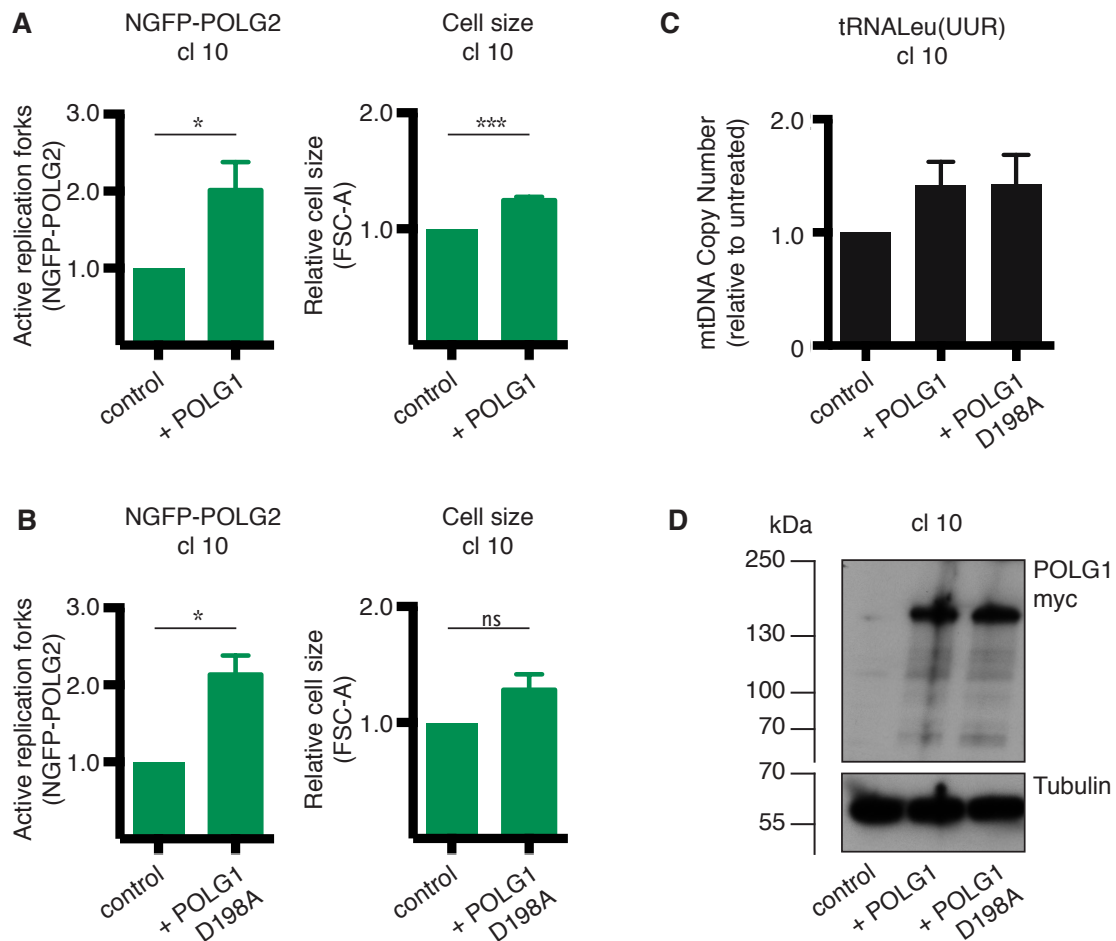
**Figure 3.4: Overexpression of the mitochondrially targeted exonuclease mUNG1 (pMA3790 mCherry) reduces active mitochondrial replication forks.** A) Gating for flow cytometric analysis of living U2-OS WT + NGFP-POLG2 cl. 10, untransfected, transfected with mCherry or transfected with pMA3790 mCherry. Cells are transfected for 72 hours. B) NGFP-POLG2 fluorescence diminishes significantly in the cells expressing the mitochondrial exonuclease, C) so does cell size. D) Mitochondrial content is not significantly altered. The experiment is performed in 3 biological replicates, data is displayed as mean  $\pm$  SEM. Statistical analysis is performed using a two-tailed, unpaired, t test, \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$ .

### 3.1.4 Overexpression of POLG1 leads to an increase in NGFP-POLG2 fluorescence

It has previously been shown that knockout of the POLG1 yeast analogue Mip1p, drastically reduces mtDNA copy number (Göke et al., 2019; Puddu et al., 2019) and that stable long-term expression of human wild-type POLG1 in cell culture had no deleterious effects on cell metabolism or copy number, in fact POLG1 is not believed to be the determinant factor in mtDNA copy number (Spelbrink et al., 2000). However, we wondered whether short-term overexpression of POLG1 in cell culture might lead to measurable changes in mtDNA replication and mtDNA copy number. Hence, we decided to use the same reporter gene fusions of full-length POLG1 wild-type and the D198A mutant, C-terminally fused to a Myc tag (kind gift from Johannes Spelbrink), for transfection in our POLG2-Assay. After 48 hours, we fixed the cells and used fluorescent antibodies against the myc-tag to separate transfected cells from untransfected ones during FACS analysis. The control cells are the untransfected subpopulation from the same pool of cells (Figure 3.5 A and B).

As anticipated, we see an increase in active mtDNA replication forks (green fluorescent signal) from both; cells overexpressing wild-type POLG1-Myc (Figure 3.5 A) and cells overexpressing the exonuclease-deficient POLG1-D198A-Myc (Figure 3.5 B). In both of these conditions we could also observe an overall increase in cell size, which was more pronounced in the cells overexpressing the wild-type POLG1 version (Figure 3.5 A and B).

We verified our results and ensured that the increase in green fluorescence is in fact due to an increase in mitochondrial replication and with that mtDNA content, by performing qPCR analysis on a parallel set of samples (Figure 3.5 C). We also performed western blot analysis to ensure antibody specificity as well as protein expression of both wild-type and mutant POLG1 (Figure 3.5 D).

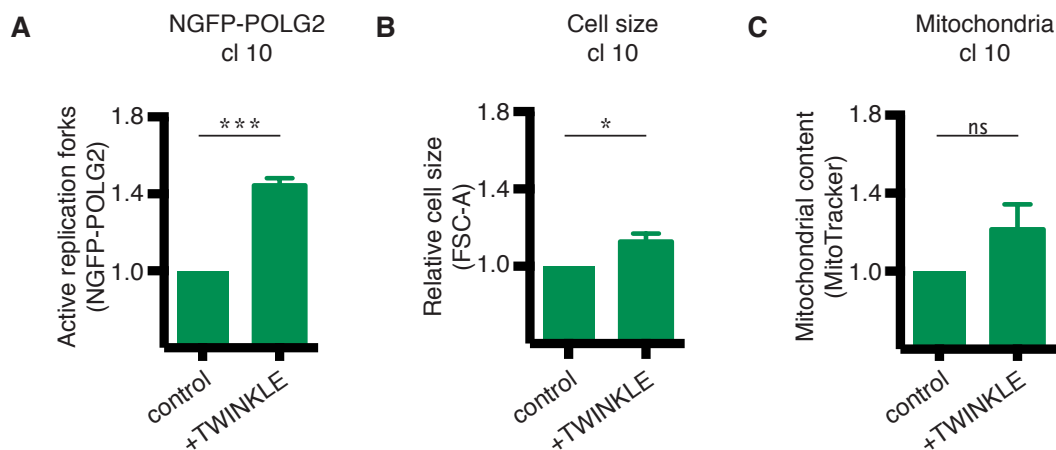


**Figure 3.5: Overexpression of wild-type POLG1 and exonuclease-deficient POLG1 increases active mitochondrial replication forks and mtDNA copy number.** A-B) The experiments are performed in 3 biological replicates, data is displayed as mean  $\pm$  SEM. Statistical analysis is performed using a two-tailed, unpaired, t test, \* $p < 0.05$ , \*\*\* $p < 0.001$ . A) NGFP-POLG2 fluorescence increases significantly in the cells overexpressing POLG1-Myc and B) exonuclease-deficient POLG1D198A-Myc. C) Comparative analysis of mtDNA copy number, measured by qPCR, of untransfected cells and cells transfected with POLG1-Myc and POLG1D198A-Myc reveals an increase in mtDNA copy number in the POLG1-overexpressing cells. Experiment is performed in 3 biological replicates, data is displayed as mean  $\pm$  SEM. D) Western blot verifying expression of POL1-Myc and POLG1D198A-Myc, as well as antibody specificity of the anti-myc antibody.

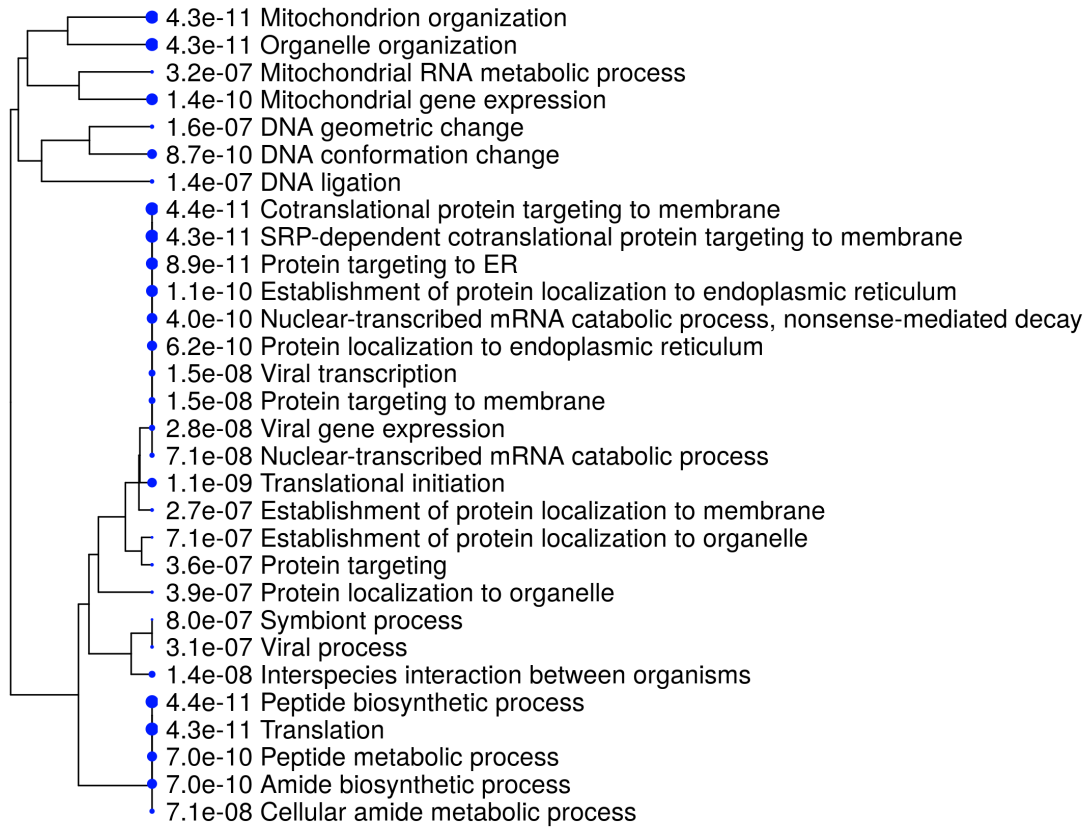
### 3.1.5 Overexpression of TWINKLE leads to an increase in NGFP-POLG2 fluorescence

There are a number of different lines of investigation suggesting that the mtDNA helicase TWINKLE might be a rate limiting step in mtDNA replication and that the overexpression of TWINKLE leads to an increase in mtDNA copy number (Ikeda et al., 2015; Peter and Falkenberg, 2020; Tynysmaa et al., 2004). Based on these reports we decided to overexpress Flag-tagged TWINKLE in our POLG2-Assay. To separate transfected from untransfected cells in the same population we used an anti-Flag antibody. The control cells in our analysis are cells from the untransfected subpopulation.

As expected, we see a clear increase in active replication forks in the TWINKLE transfected subpopulation of cells (Figure 3.6 A) which coincides with a significant increase in cell size (Figure 3.6 B). Additionally, TWINKLE transfected cells appear to have an all-over increase in mitochondrial content, this however is a trend and not statistically significant (Figure 3.6 C).



**Figure 3.6: Overexpression of TWINKLE increases active mitochondrial replication forks.** A) NGFP-POLG2 fluorescence increases significantly in the cells overexpressing TWINKLE-Flag, so does B) cell size. C) Although there is a trend towards increased mitochondrial content in the TWINKLE-overexpressing cells, this is not statistically significant. The experiment is performed in 3 biological replicates, data is displayed as mean  $\pm$  SEM. Statistical analysis is performed using a two-tailed, unpaired, t test, \* $p < 0.05$ , \*\*\* $p < 0.001$ .



**Figure 3.7: ShinyGO hierarchical clustering tree of the MacroD1-enriched interactors implicating MacroD1 in mtDNA homeostasis and mitochondrial gene expression.** Hierarchical clustering tree summarizing the correlation amongst significantly enriched pathways. Pathways with many shared genes are clustered together.

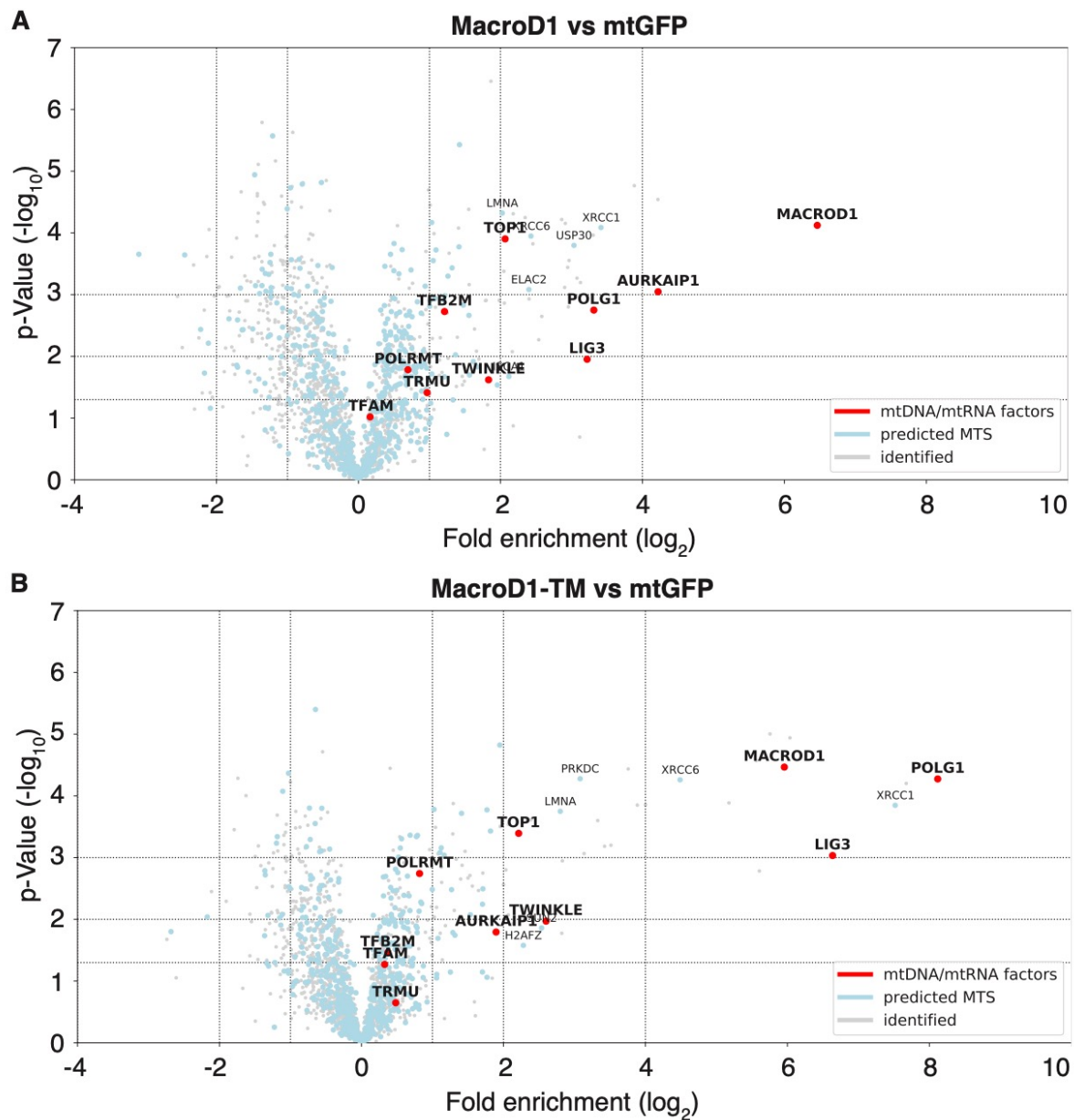
## 3.2 Analysis of MacroD1 interaction partners

Mono-ADP-ribosylated proteins of the mitochondrial matrix have been reported, notably glutamate dehydrogenase, whose ADP-ribosylation regulates metabolism (Haigis et al., 2006; Herrero-Yraola et al., 2001). However, a targeted, proteomic analysis of mitochondrial ADP-ribosylation has not been performed, nor do we have any insights into the mitochondrial matrix substrates of the ADP-ribosyl-binding MacroD1 protein. To gain insights into the ADP-ribosylation factors and ADP-ribosylation substrates in mitochondria; Aurelio P. Nardoza generated a MacroD1 and mtTARG1-specific interactome and ADP-ribosylome in collaboration with Fabian Hosp in the Matthias Mann Lab (see sections 2.2.1, 2.2.10 and 2.2.18 for methods and background). The original mass spectrometry Excel analysis sheet from Fabian and Aurelio can be found linked in Appendix 9. These original datasets are analysed in this thesis with informatics support from Benjamin E. Oliver. See Appendix 9 for MTS and Volcano plot analysis. Some of the mass spectrometry identified hits were validated in this subchapter (section 3.2.6) using far western blotting.

A number of results discussed in this subchapter are also vital parts of the manuscript *Protein ADP-ribosylation Regulates Human Mitochondrial DNA Replication*(Söllner et al., 2020b). In the manuscript we have some additional data and cell lines.

### 3.2.1 Analysis of MacroD1 interaction partners

Initially, to make sense of the mass spectrometric datasets, we performed an enrichment analysis to identify recurring molecular pathways and functional categories. To perform this analysis we used ShinyGO (Ge et al., 2020) (accessed in 2020 via: <http://bioinformatics.sdstate.edu/go/>) and entered the protein-identifiers of hits significantly enriched in the MacroD1-GFP versus mtGFP IPs with a  $\log_2$  fold enrichment of  $\geq 1$ . Thereby generating a ShinyGO hierarchical clustering tree (Figure 3.7).



**Figure 3.8: The MacroD1 Interactome.** A) Volcano plot of the mass spectrometric data, visualizing MacroD1-enrichments versus mitochondrial GFP control samples. Magnitude of enrichment is on the x-axis and significance of enrichment on the y-axis. The experiment was performed in three biological replicates. On the plot the grey dots are the total identified proteins, in light blue the identified proteins that have a putative MTS (according to TargetP predictions) and the red dots are proteins known to interact with mtDNA and mtRNA. B) Volcano plot of the mass spectrometric data, visualizing MacroD1-TM-enrichments versus mitochondrial GFP control samples. Magnitude of enrichment is on the x-axis and significance of enrichment on the y-axis. The experiment was performed in three biological replicates.

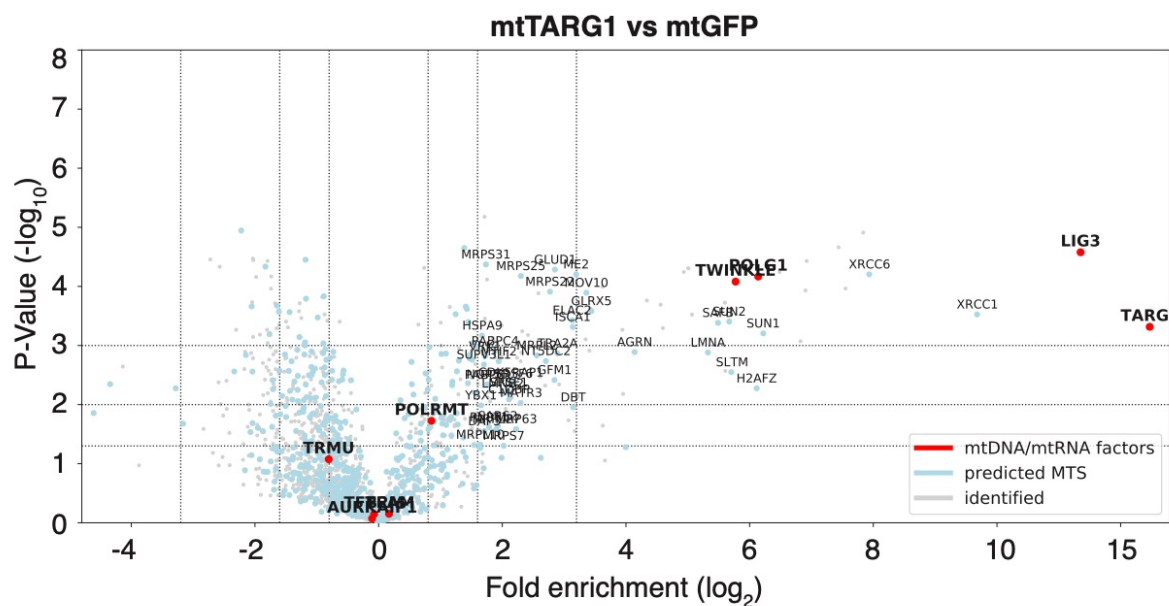


From the hierarchical clustering tree, we saw that a significant amount of the proteins identified in the MacroD1 interactome are involved not only in DNA processes such as DNA geometric and conformational changes, but also more specifically in mitochondrial gene expression and mitochondrial RNA metabolic processes (Figure 3.7).

Additionally, we identified both Huntington's disease (adj. Pval: 4.0e-03) and Parkinson's disease (adj. Pval: 2.1e-03) as relevant pathways in KEGG pathway analysis (accessed via: <http://bioinformatics.sdstate.edu/go/>), both of which have been implicated in relation to mtDNA maintenance and homeostasis (Askeland et al., 2018; Coxhead et al., 2016; Dolle et al., 2016; Luoma et al., 2004; Wang et al., 2013a).

### 3.2.2 MacroD1 interactome: MacroD1 interacts with the mitochondrial DNA replication and repair factors.

To further analyse and visualize the mass spectrometric data Benjamin wrote a python script generating a table which maps every protein to its associated gene. Using the UniProt API (UniProt, 2019), the associated FASTA files were retrieved for each protein and used to generate files suitable for upload to the batch endpoint of the TargetP webpage (<http://www.cbs.dtu.dk/services/TargetP/>). If any of the amino acid sequences of a protein indicate mitochondrial targeting, then the protein in our table is marked as being a putative candidate for mitochondrial localization (see Appendix 9 for links to lists). In our volcano plot visualization, we denoted the names of all proteins containing a putative signalling peptide, combined with a significant  $\log_2$  fold enrichment of  $\geq 2$ . Additionally, all proteins with a putative MTS according to the TargetP algorithms are indicated using the blue dot symbol (Figures 3.8 to 3.11). Based on the ShinyGo hierarchical clustering (Figure 3.7), we knew that a significant number of proteins identified in the MacroD1 interactome are involved in mitochondrial DNA homeostasis and mitochondrial RNA metabolism. Therefore we highlighted



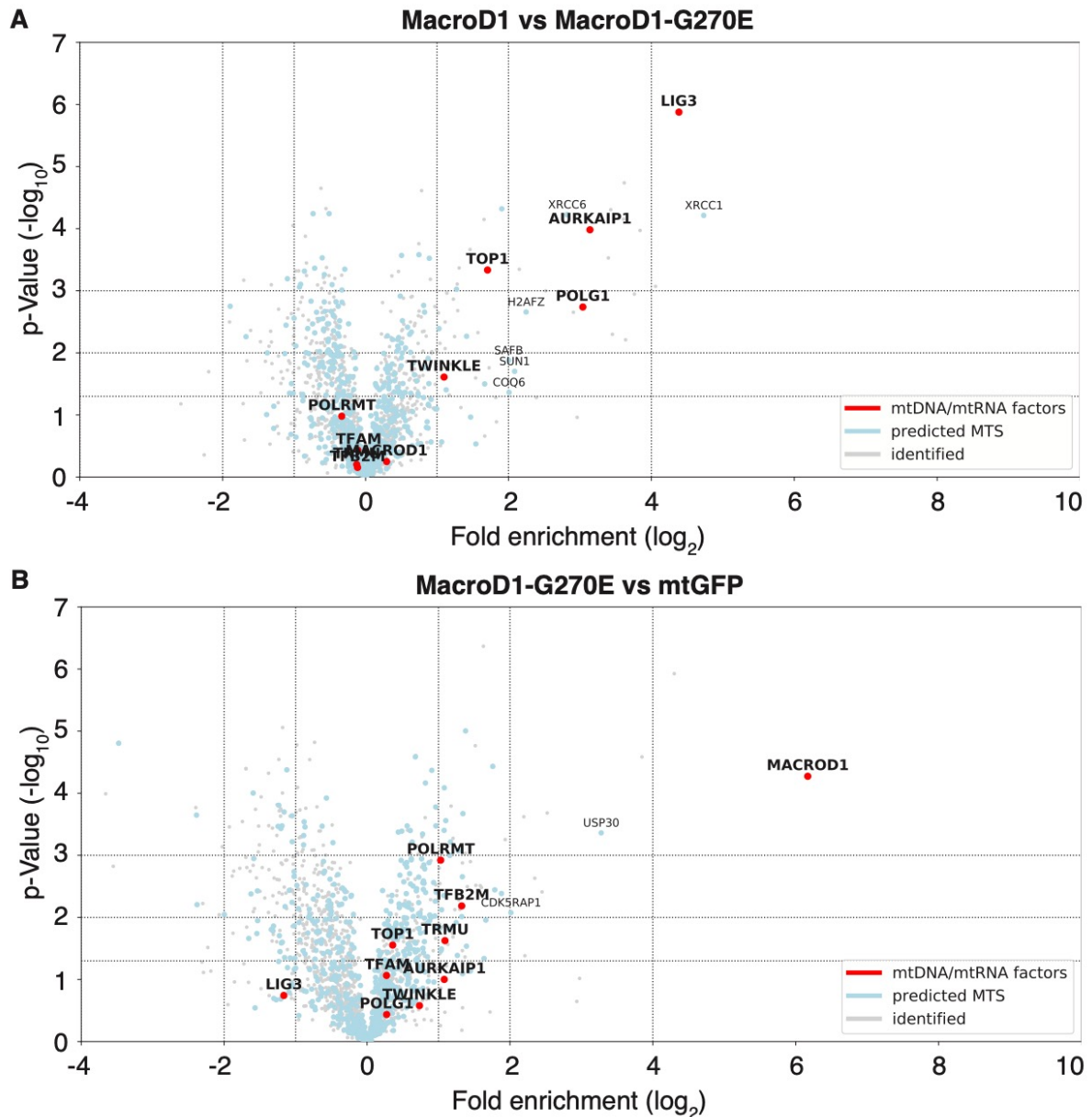
**Figure 3.9: The mtTARG1 Interactome.** Volcano plot of the mass spectrometric data, visualizing the mtTARG1 interactome by plotting data from the mtTARG1-enrichments versus mtGFP control samples. Experiments are performed in biological triplicate. On the plots the grey dots are the total identified proteins, in light blue the identified proteins that have a putative MTS (according to TargetP predictions) and the red dots are proteins known to interact with mtDNA and / or mtRNA.

known mtDNA and mtRNA factors (red dots) with a  $\log_2$  fold enrichment less than 2, as well as MacroD1, TARG1 and GLUD1, in order to facilitate observation of these proteins throughout the different volcano plots (Figures 3.8 to 3.10).

The MacroD1 interactome is visualized in the MacroD1-GFP versus mtGFP (Figure 3.8 A) and with the 'stronger binder' MacroD1-TM-GFP mutant versus mtGFP (Figure 3.8 B) IP plots. As suggested by the hierarchical clustering, we see a number of known mtDNA and mtRNA factors significantly enriched in the MacroD1 IPs, namely: TFB2M, TWINKLE, ATAD3B, AURKAIP1, TOP2A, TOP1, LIG3, PARP1 and POLG1. As well as some factors that contain a putative MTS and function in DNA metabolism, such as XRCC1 and XRCC6. This analysis is further strengthened by the mtTARG1 interactome (Figure 3.9).

TARG1 is a PAR- and MAR-interacting macrodomain capable of cleaving the acyl bond between a carboxylate and ADP-ribose, such as in MARYlated glutamate residues (Peterson et al., 2011; Sharifi et al., 2013). In these experiments the normally nuclear TARG1 is artificially targeted to mitochondria (mtTARG1) by subcloning in an open reading frame with the MacroD1 MTS.

Using the same constraints and parameters as in the MacroD1 interactome, we plotted the mtTARG1 interactome and found that POLRMT, TWINKLE, TOP2A, PARP1, LIG3 and POLG1 are all enriched in the mtTARG1 IP (Figure 3.9). Thereby showing that mtTARG1 and MacroD1 enrich a highly related set of mitochondrial proteins (Figures 3.8, 3.9 and 3.10 A), thus validating our MacroD1 interactome. Interestingly, GLUD1, the first mitochondrial enzyme shown to be modified and regulated by ADP-ribosylation (Haigis et al., 2006; Herrero-Yraola et al., 2001), was also enriched in the mtTARG1 IP - further strengthening our results.

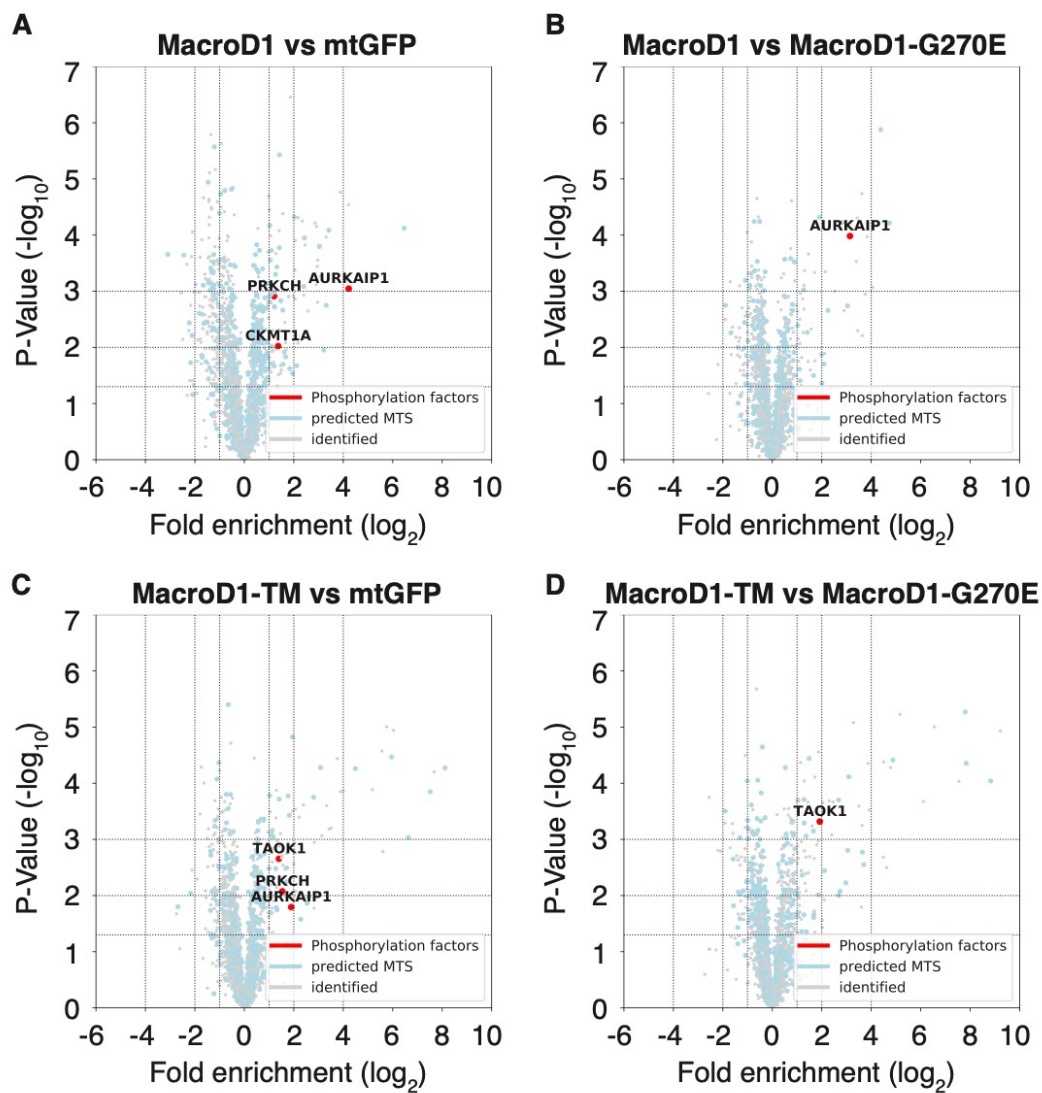


**Figure 3.10: The MacroD1-specific ADP-ribosylome & ADPr-binding-independent interaction partners of MacroD1.** A) Volcano plot of mass spectrometric data, visualizing interaction partners enriched in MacroD1-enrichment versus interaction partners enriched using MacroD1-G270E as bait. The experiment was performed in three biological replicates. The grey dots are the total identified proteins, in light blue the identified proteins that have a putative MTS (according to TargetP predictions) and the red dots are proteins known to interact with mitochondrial DNA and/or RNA. B) Volcano plot of mass spectrometric data, visualizing interaction partners enriched in MacroD1-G270E-enrichment versus mtGFP control samples. The experiment was performed in three biological replicates. The grey dots are the total identified proteins, in light blue the identified proteins that have a putative MTS (according to TargetP predictions) and the red dots are proteins known to interact with mitochondrial DNA and/or RNA.

### 3.2.3 MacroD1-based ADP-ribosylome: MacroD1 interaction with the mtDNA replication factors depends on ADP-ribosylation

In the mitochondrial macrodomain interactomes (Figures 3.8, 3.9 and 3.10 A) we identified a number of, core, mtDNA transcription and replication factors: TFB2M, TWINKLE, TOP2A, TOP1, LIG3, POLRMT and POLG1. Which led us to question whether the interactions between these proteins and our macrodomains is ADPr-binding dependent or not. To investigate further we generated a MacroD1-based ADP-ribosylome, by analysing the enrichment in the MacroD1-GFP IPs versus the MacroD1-G270E-GFP IPs (Figure 3.10 A). Which will indicate the factors requiring a functioning ADPr-binding macrodomain for the interaction. Further, we wanted to see whether there are interaction partners of MacroD1, which bind irrespective of the ADPr-binding site – so we additionally plotted the MacroD1-G270E-GFP IPs versus the mtGFP IPs (Figure 3.10 B). The core replication factors POLG1, TWINKLE, TOP1, TOP2B, and LIG3 interact with MacroD1 in an ADPr-dependent manner (Figure 3.10 A). Which suggests that these essential replication factors are ADP-ribosylated (further tested in section 3.2.6) and that their interaction with MacroD1 supports the canonical role of the ADP-ribose binding pocket of macrodomains as a PTM recognition module. TRMU, TF2BM and POLRMT, on the other hand, interact with MacroD1 independent of its ADPr binding capacity (Figure 3.10 B). TF2BM and POLRMT are essential in mitochondrial transcription, as well as in the initiation of mtDNA replication, while TRMU is a mitochondria-specific tRNA-modifying enzyme.

It is quite striking that the most enriched ADPr-binding dependent MacroD1 interacting proteins include the entire core machinery required for mtDNA replication: TWINKLE the mitochondrial DNA helicase, POLG the replicative mitochondrial DNA polymerase, the two topoisomerases TOP1 and TOP2A and LIG3, the only mitochondrial DNA ligase. While TF2BM and POLRMT, essential for both initiation of replication and transcription of mtDNA, are enriched as MacroD1-interactors irrespective of its ADPr -binding capacity. This leads to the question, whether ADP-ribosylation status and interaction is involved in the regulation between mitochondrial transcription and mtDNA replication initiation?



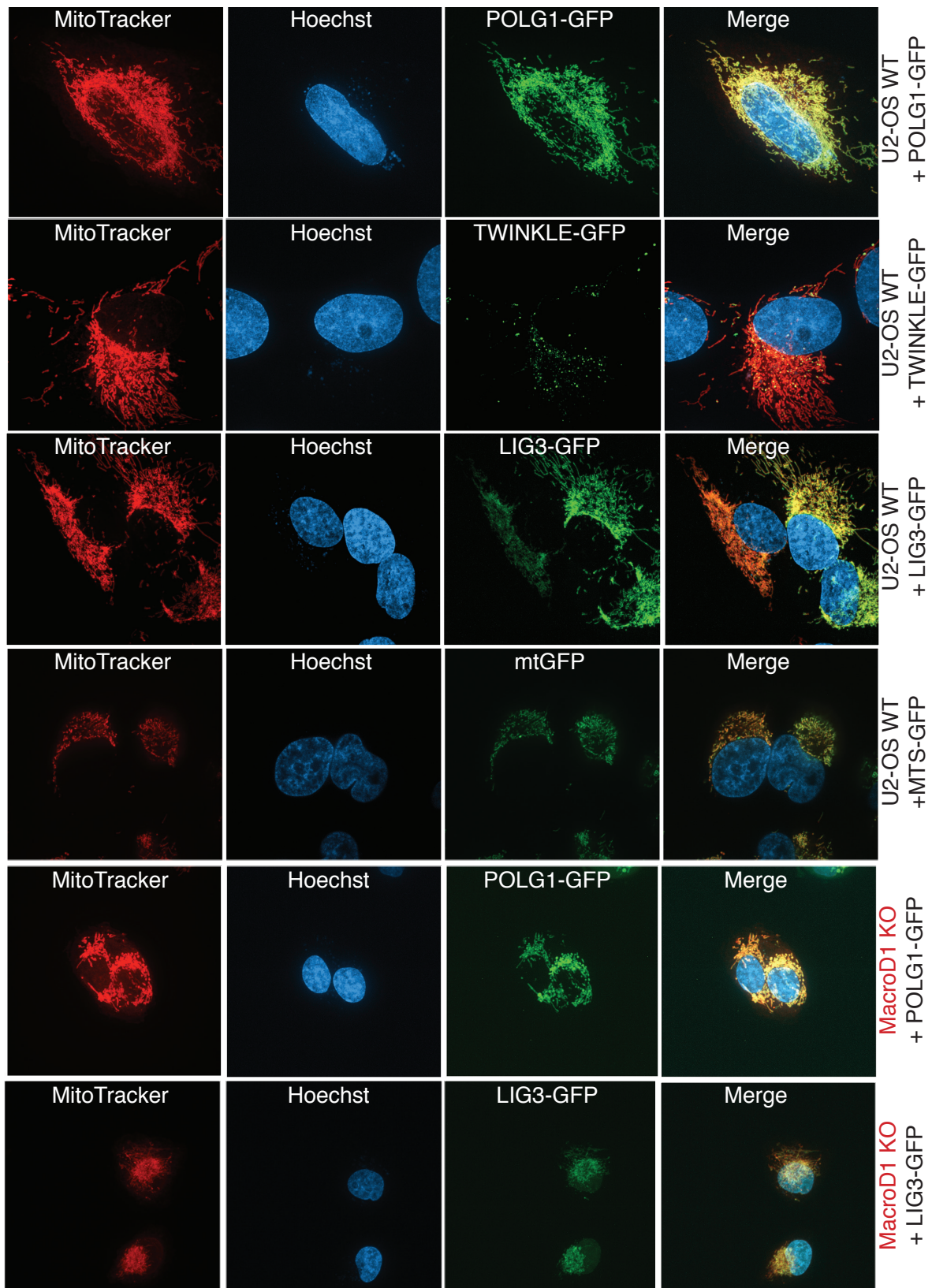
**Figure 3.11: Mitochondrial Phosphorylation Factors Interact with MacroD1.** A and C) Volcano plots depicting MacroD1-based interactomes. B and D) Volcano plots showing the ADP-ribosylome analysis. The grey dots are the total identified proteins, in light blue the identified proteins that have a putative MTS (according to TargetP predictions) and the red dots are proteins implicated in phosphorylation events.

### 3.2.4 MacroD1 interacts with factors involved in phosphorylation reactions

While analysing the various macrodomain protein interactomes and ADP-ribosylomes, we noticed that both the Serine/threonine-protein kinase TAO1 (TAOK1) and Aurora kinase A-interacting protein (AURKAIP1) are highly enriched in the MacroD1 interactomes and ADP-ribosylomes (see Figure 3.11 for an alternate representation).

AURKAIP1 is a known interactor and negative regulator of the serine/threonine kinase Aurora A (AURKA), while TAOK1 is itself a serine/threonine kinase. Both AURKAIP1 (Koc et al., 2013) and AURKA (Bertolin et al., 2018) have been implicated in mitochondrial homeostasis and metabolism. The function of AURKAIP1 is two-fold it i) functions as a negative regulator for AURKA and ii) it is directly involved in the expression of mitochondrially encoded proteins as a member of the mitochondrial ribosome proteins (Koc et al., 2013).

We don't know which residues of our identified proteins are ADP-ribosylated, but the interplay between Ser-phosphorylation and Ser-ADP-ribosylation has previously been highlighted (Larsen et al., 2018). Taken together with the reports that phosphorylation of TFAM Ser residues has a regulatory function in both TFAM-DNA interaction (King et al., 2018) and mitochondrial transcriptional abrogation (Lu et al., 2013), one wonders whether there also might exist a phosphorylation – ADP-ribosylation interplay in mitochondria?



**Figure 3.12: U2-OS cells stably expressing mtGFP and the mtDNA replication factors: POLG1-GFP, TWINKLE-GFP and LIG3-GFP.** Life cell imaging verifying expression and localisation of stably integrated transgenes. MitoTracker in red, nuclei are stained in blue, expressed GFP-fused transgenes are in green.

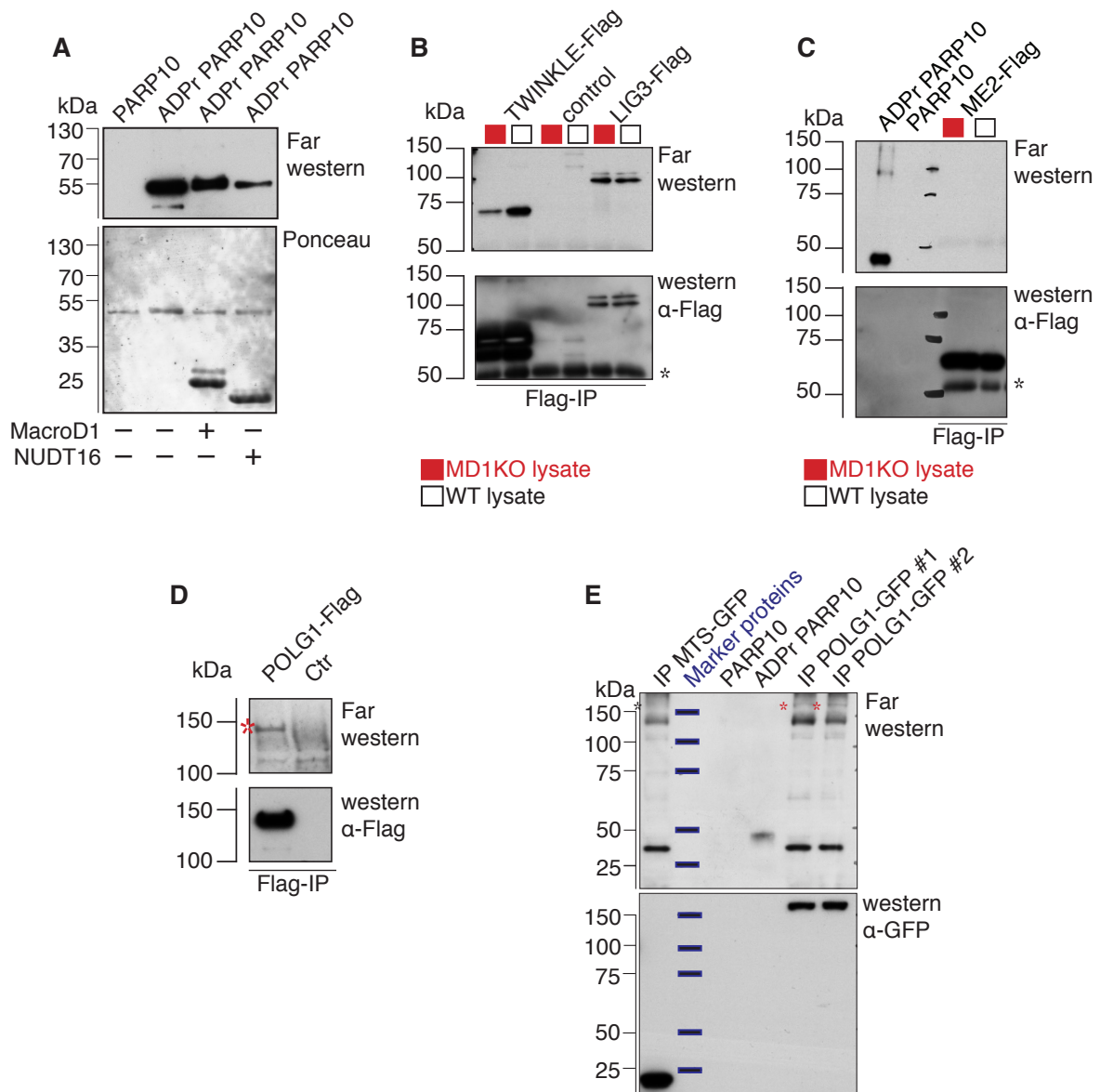


### 3.2.5 Generating and characterising monoclonalised U2-OS cell lines expressing core mtDNA replication factors

To further analyse the mtDNA replication factors, identified in the MacroD1-dependent ADP-ribosylome, we decided to generate monoclonalised cell lines stably expressing mtGFP and the mtDNA replication factors: POLG1-GFP, TWINKLE-GFP and LIG3-GFP in the U2-OS wild-type background (Figure 3.12). The stable transfection and monoclonalisation were performed as described in sections 2.2.3 and 2.2.17. The plasmids used for the transfections are listed in Appendix 6. POLG1-GFP, TWINKLE-GFP, LIG3-GFP and mtGFP are all stably expressed in the wild-type U2-OS background and localize to the mitochondrial compartment (Figure 3.12).

We decided to also stably integrate mtGFP, LIG3-GFP and POLG1-GFP into the MacroD1 knockout background. The CRISPR/Cas9 MacroD1 knockout cells stem from a collaboration with Rebecca Smith and are discussed in more detail in subchapter 3.3. POLG1-GFP, LIG3-GFP and mtGFP (not shown here) are all stably expressed in the MacroD1 knockout background and localize to the mitochondrial compartment (Figure 3.12).

These cell lines can be used for a number of things, such as interaction studies like IPs coupled with Far western blot analysis or Co-IPs.



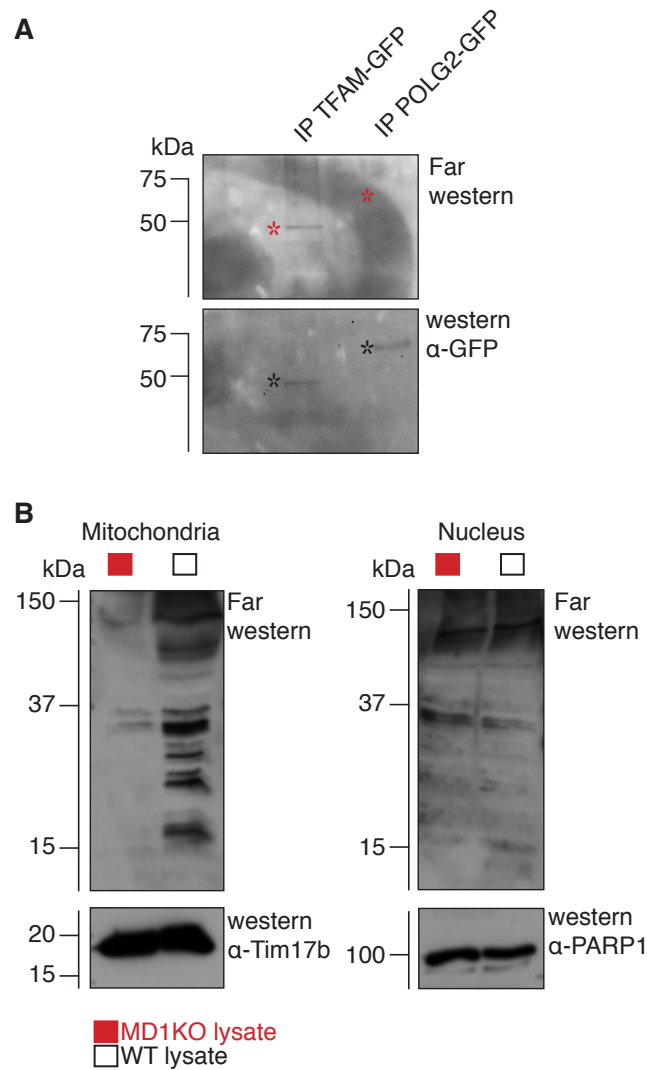
**Figure 3.13: ADP-ribosylation Verification.** A) Far western blot of purified PARP10, in both its unmodified (lane 1) and MARylated states (lanes 2-4). MARylated PARP10 was incubated with either active MacroD1 (lane 3) or NUDT16 (lane 4), hydrolyzing the MAR PTM. Ponceau staining is used as a loading control for the purified PARP10, MacroD1 and NUDT16 proteins (blot courtesy of Aurelio Nardoza). B) Far western blot of immunoprecipitated TWINKLE and LIG3 as well as C) ME2 from both wild-type (black) and MD1KO (red) lysates. The asterisk denotes an IgG heavy-band. D) Far western blots of immunoprecipitated POLG1. Red star indicates ADP-ribosylated POLG1 in the far western. E) Far western blot of immunoprecipitated POLG1-GFP from two stable POLG1-GFP expressing U2-OS cell lines. Red stars indicate the immunoprecipitated, MARylated POLG1 proteins. Modified and unmodified PARP10 serve as control for detection of MARylated protein. Immunoprecipitated mtGFP from a stably expressing U2-OS cell line was used as an additional control, the black star highlights the absence of the band seen in the POLG1 IPs. Anti-GFP western blot is used as an expression, immunoprecipitation, loading control.

### 3.2.6 Verifying ADP-ribosylation status of mitochondrial DNA replication and repair factors

The critical dependence of a functional ADPr-binding pocket in MacroD1 for the observed interaction with the human mtDNA replication machinery, raises the possibility that one or more mtDNA replication factors may be ADP-ribosylated. To test whether the identified hits carry this post-translational modification, we performed immunoprecipitation assays followed by far western detection of the blotted proteins with a recombinant, purified protein tool, developed and validated in our Lab by Aurelio P. Nardoza, which consists of three consecutive repeats of the MacroD1 MD1<sup>TM</sup> mutant fused to a V5-tag. **Aurelio's** validation can be seen in Figure 3.13 A. Here, he used unmodified (lane 1) and MARYlated PARP10 (lane 2), as well as MARYlated PARP10 incubated with either purified MacroD1 (lane 3) or NUDT16 (lane 4). Ponceau staining was used as a loading control (Figure 3.13 A). From the validation far western we see that the far western tool detects PARP10 only when PARP10 is MARYlated. If MacroD1 is an essential mono-ADP-ribosyl hydrolase enzyme, we would expect that incubating purified MARYlated proteins with MacroD1 would lead to a significant decrease in ADP-ribosylation signals. However, we find that the far western signal for mono-ADP-ribosylated PARP10 is only slightly depleted upon incubation with active MacroD1 enzyme (lane 3), but is robustly reduced by NUDT16 (lane 4), a Nudix hydrolase (Daniels et al., 2015; Luscher et al., 2018; Palazzo et al., 2015).

Having established that the far western tool can be utilized to detect MARYlation of substrate proteins, we decided to overexpress tagged mtDNA replication factors, POLG1, LIG3 and TWINKLE and to analyse them using far western analysis. To do this we utilized both transient transfection in HEK293-T-REx cells (Figure 3.13 B, C and D) and the stable monoclonalised cell lines described in section 3.2.5 (Figure 3.13 E).

Our far western blots reveal clear bands consistent with the ADP-ribosylation of TWINKLE, LIG3 and POLG1 (Figure 3.13 B, D and E).



**Figure 3.14: Essential mtDNA packaging factor TFAM is ADP-ribosylated & MacroD1 impacts mitochondrial, not nuclear, ADP-ribosylation.** A) Far western blot of immunoprecipitated TFAM-GFP and NGFP-POLG2 from transiently transfected U2-OS cells. B) Far western blot of wild-type U2-OS (WT; black) and MD1KO (red) mitochondrial and nuclear fractions. TIM17b is used as a mitochondrial loading control and PARP1 for nuclear fractions.

While we do not know which residues of the core factors required for mtDNA replication are modified, the assay verifies that they are in fact ADP-ribosylated, as implied by the mass spectrometric analysis (section 3.2.2). After validating the ADPr-status of essential mtDNA replication factors, POLG1, LIG3 and TWINKLE, we also investigated TFAM and POLG2 using the far western tool and saw that TFAM is ADP-ribosylated, POLG2 however is not (Figure 3.14 A).

If MacroD1 is an essential mono-ADP-ribosyl hydrolase enzyme of the mitochondrial matrix, we would expect that cells lacking MacroD1 function would exhibit increases in ADP-ribosylation signals. Interestingly, deletion of the MacroD1 gene using CRISPR/Cas9 (mentioned in section 3.2.5 and described in subchapter 3.3) fails to increase the far western signal for TWINKLE or LIG3 compared to wild-type cells (Figure 3.13 B). In fact, the far western ADP-ribosylation signal for TWINKLE is actually lower in MD1KO cells compared to wild-type. The main function of MacroD1 may thus be related to its ability to bind ADP-ribosylated proteins rather than its enzymatic, hydrolytic activity on MARylated proteins, which is thought to be relatively weak, something we have also seen in the *in vitro* far western validation (Figure 3.13 A). One of the core functions of mitochondrial MacroD1 may thus be to bind ADP-ribosylated proteins. The low hydrolase activity of MacroD1 and the fact that MacroD1 deletion can decrease the far western signal of mitochondrial ADP-ribosylation (rather than increasing it), raises the possibility that MacroD1's ability to bind ADP-ribosylated mitochondrial substrates might stabilize their ADP-ribosyl modification. To test this hypothesis, we performed far western analyses on both mitochondrial and nuclear extracts of wild-type and MD1KO U2-OS cells (Figure 3.14 B).

While MacroD1 deletion does not alter the far western blot for human nuclear extracts, the far western signal is almost completely abolished in MacroD1 knock-out mitochondria compared to wild-type cell extracts. This not only indicates that MacroD1 has a stabilizing effect on mitochondrial ADP-ribosylation, but also demonstrates that MacroD1 deletion strongly impacts mitochondrial ADP-ribosylation levels without directly affecting nuclear ADP-ribosylation.

### 3.3 The biological function of MacroD1 and ADP-ribosylation in mitochondria

In subchapter 3.2 we showed that MacroD1 interacts with core mtDNA replication, transcription and translation factors POLG1, TWINKLE, LIG3, TOP1, TOP2B, AURKAIP1, POLRMT and TFB2M. While, engineered mutant MacroD1 proteins reveal that the ADP-ribose binding pocket of the MacroD1 macrodomain module is key to the specific interaction with most of the enriched DNA replication components (Figures 3.8). Confirmed also by the use of an orthogonal approach with an engineered, mitochondrial-matrix targeted TARG1 (Figure 3.9).

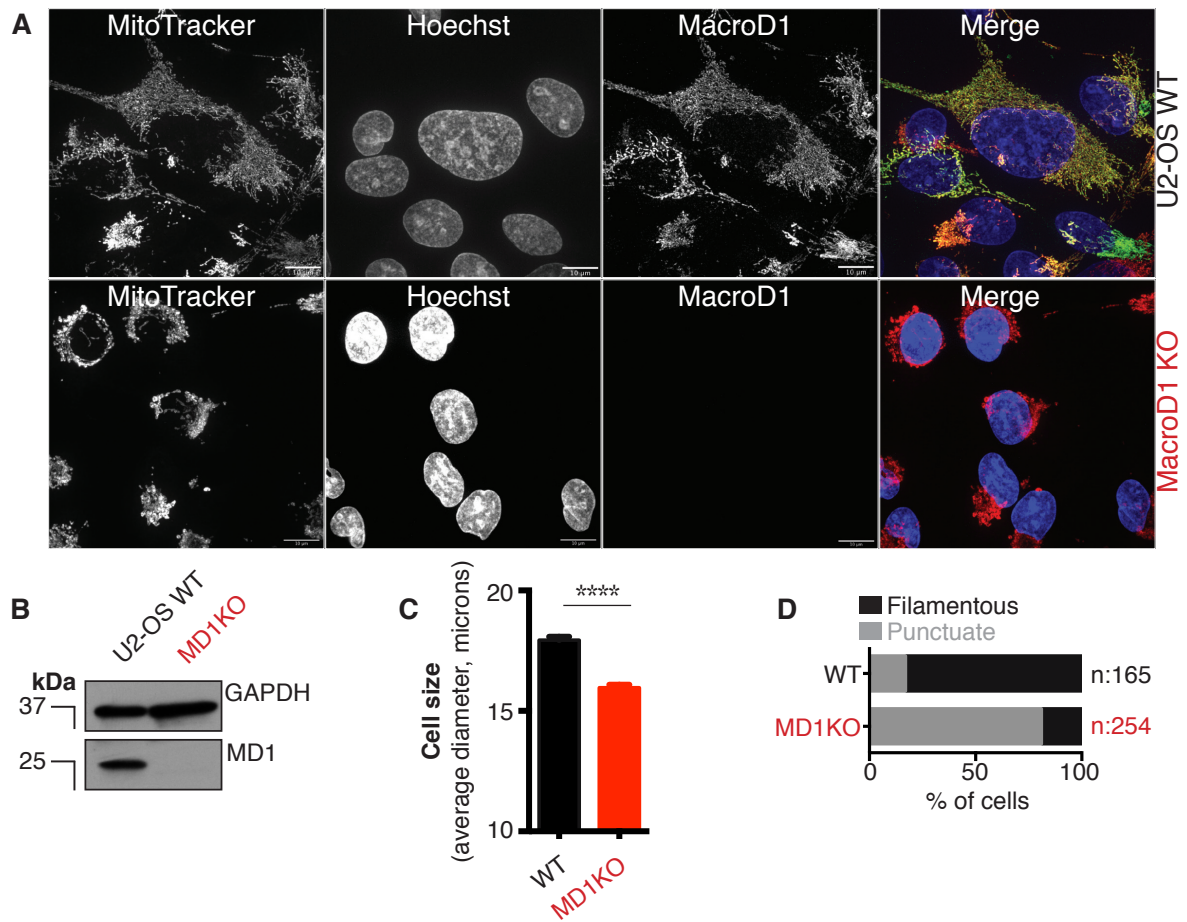
The strong biochemical association of mtDNA replication factors with MacroD1 establishes a mechanistic connection between mtDNA homeostasis and MacroD1 function. In support of the canonical role of the ADP-ribose binding pocket of macrodomains as a PTM recognition module, our results demonstrate that the core replication factors POLG1, TWINKLE, TOP1, TOP2B, and LIG3 interact with MacroD1 in an ADPr-dependent manner (Figure 3.10 A). In contrast, two factors required for mtDNA transcription, POLRMT and TFB2M, interact with MacroD1 irrespective of its ADPr binding capacity (Figure 3.10 B). Indicating the existence of distinct mechanisms of interaction between MacroD1 and mtDNA transcription/replication factors. Some factors require the ADP-ribosyl-binding function of the MacroD1 protein, while others likely interact with MacroD1 through a surface region.

Consistent with the ADP-ribosyl-binding function of MacroD1, we find that a far western method, related to that pioneered by others (Gibson et al., 2017; Khadka et al., 2015; Wu et al., 2007), indicates that TWINKLE, LIG3, POLG1 and TFAM may be ADP-ribosylated, while POLG2 is not (Figures 3.13 and 3.14). TFAM is the mtDNA packaging factor that is involved in mtDNA transcription initiation (Falkenberg, 2018; Ngo et al., 2011), its ADP-ribosylation underpins the close connections between mtDNA transcription and replication. We have not yet identified the ADP-ribosylation sites in mtDNA replication factors. However, our evidence suggests that mtDNA replication and transcription factors are direct targets of mitochondrial

ADP-ribosylation pathways. However, the question remains: what is the role of ADP-ribosylation of mtDNA replication factors?

To answer this question and to determine whether MacroD1 has regulatory functions in mitochondria, on a cellular level and on an organismal level, we generated and characterized MacroD1 knockout (MD1KO) cell lines using CRISPR/Cas9 in collaboration with Rebecca Smith (sections 3.3.1 to 3.3.7) and knocked down a putative mitochondrial macrodomain in the model organism *Drosophila melanogaster* (section 3.3.8), with expert advice from Carla Margulies and Maria Spletter.

Some of the results discussed in this subchapter are also vital parts of the manuscript *Protein ADP-ribosylation Regulates Human Mitochondrial DNA Replication* (Söllner et al., 2020b), in which we also show additional data using a HEK293-T-Rex MacroD1 knockout cell line.



**Figure 3.15: CRISPR/Cas9 deletion of the MacroD1 gene in U2-OS cells.** A) Immunofluorescent staining verifying the knock-out of MacroD1 in the U2-OS cell line, as well as the cell size difference. The MacroD1-deficient cells have punctuated, perinuclear, mitochondria. Wild-type U2-OS cells exhibit filamentous mitochondrial networks. In the merged images: nuclei are stained in blue, MD1 antibody staining in green and MitoTracker in red. B) Knock-out of MacroD1 in the U2-OS cell line as shown by western blot. C) MacroD1-deficient U2-OS cells are smaller than the parental control cell line, as measured by the Vi-CELL XR Cell Viability Analyzer. Measurements were taken on 40 replicates, cell size after trypsinisation is given as diameter in microns. Data is displayed as mean  $\pm$  SEM. Two-tailed, unpaired t-test, \*\*\*\* $p < 0.0001$ . D) Mitochondrial morphology classification into filamentous and punctuated subgroups. Classification was performed by 5 independent individuals and confirmed a shift of mitochondrial morphology from filamentous to punctuate in the MD1KO cells. The scoring was performed single-blinded.

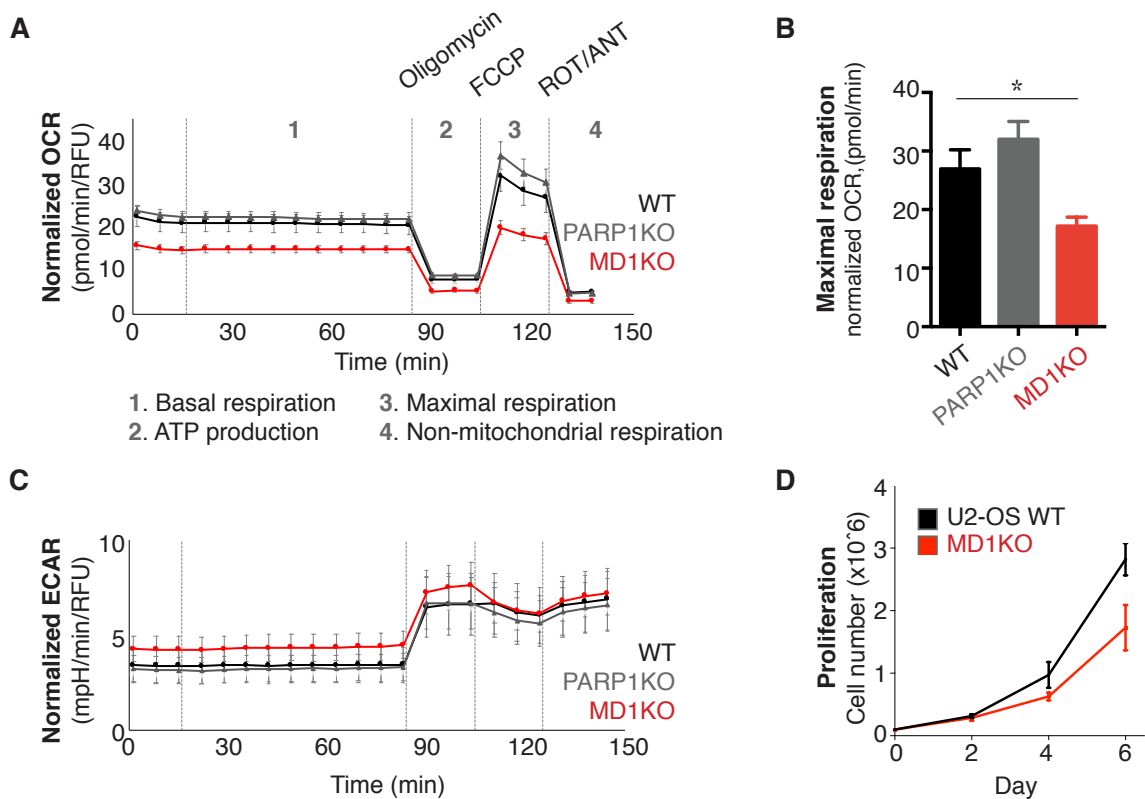


### 3.3.1 The characterization of MacroD1 knockout cells: MacroD1 impacts mitochondrial and cellular homeostasis

To validate the MacroD1 knockout cells we performed immunofluorescent staining of the U2-OS MacroD1 knockout cell line compared to the parental control cell line, using a MacroD1-specific antibody (Figure 3.15 A). The U2-OS wild-type cell line has a strong mitochondrial MacroD1-signal, while the MacroD1 knockout cells do not, as seen by merging MitoTracker and MacroD1-antibody staining (Figure 3.15 A).

To further ensure that the observed mitochondrial signal is in fact MacroD1, we also performed western blot analysis (Figure 3.15 B), the band that is seen in the wild-type cell line and gone in the MacroD1 knockout cell line migrates at just above 25 kDa, which is the size of the processed, mitochondrial, MacroD1 protein. Thereby, verifying both the specificity of the MacroD1-antibody, as well as the knockout of MacroD1 in the MacroD1 knockout cells. While analysing these cells, we observed clear cellular and mitochondrial morphology phenotypes. The wild-type U2-OS cells were significantly larger than their MacroD1 knockout counterparts and exhibited filamentous mitochondrial networks throughout the cytoplasm. In contrast, MD1KO cells had perinuclear clusters of fragmented punctuate mitochondria (Figure 3.15 A).

To quantify the size differences between the MD1KO cells and their parental controls, we used the Vi-CELL XR Cell Viability Analyzer, while classification of the mitochondrial morphology phenotype was performed blinded by 5 independent individuals into either punctuate or filamentous morphology. With these methods, we quantitatively verified that the knockout cells were significantly smaller than their parental control (Figure 3.15 C) and confirmed a shift of mitochondrial morphology from filamentous in the wild-type cells to punctuate in the MD1KO cells (Figure 3.15 D).



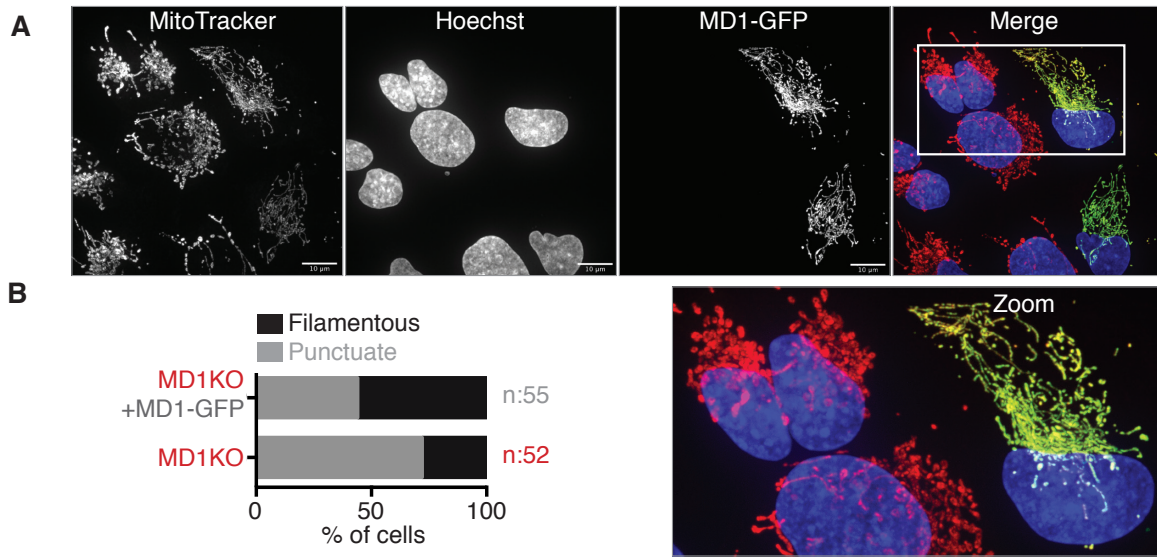
**Figure 3.16: MacroD1 knockout cells exhibit a reduction in respiration and a slower proliferation rate.** A) Oxidative consumption rate (OCR), as measured by seahorse XF, of wild-type U2-OS (black), MD1KO (red) and PARP1KO (grey) cells. Sequential injections of oligomycin A (10  $\mu$ M), FCCP (10  $\mu$ M), and rotenone/antimycin A (5  $\mu$ M). Measurements were performed in 6 biological replicates, with 6 technical repeats each. B) Maximal respiration measurements from the seahorse assay are displayed as mean  $\pm$  SEM. Ordinary one-way ANOVA, Turkey's multiple comparison test, \* $p < 0.05$ . C) Extracellular acidification rate (ECAR), measured in parallel with OCR. (Data A- C courtesy of Evangelia Tzika) D) MD1KO cells replicate slower than the wild-type U2-OS control cell line. Proliferation rate assay was performed on 3 biological replicates in technical duplicates, data is displayed as mean  $\pm$  SEM.

To test whether the observed morphological changes affect the cellular bioenergetic profile of the cells, we performed Seahorse XF Cell Mito Stress Test measurements in collaboration with Evangelia Tzika (3.16 A to C).

It has previously been reported that depletion of PARP1, a key cellular poly-(ADP-ribose) polymerase enzyme, leads to an increase in oxidative phosphorylation in whole cells and in isolated mitochondria (Szczesny et al., 2014). Using a CRISPR/Cas9-engineered U2-OS PARP1 knock-out cell line (Sellou et al., 2016) as a control for our assay, we observed the expected increase in oxidative phosphorylation in the PARP1 knock-out cells. In contrast, MacroD1KO cells exhibited a significant reduction in maximal respiration and a shift towards glycolysis (Figure 3.16 A-C). We were further able to show that this metabolic deficit is reflected in a diminished cellular proliferation rate (Figure 3.16 D). We observed the same bioenergetic shift and proliferation deficiency using a HEK293-T-REx MD1KO cell line, further demonstrating that MacroD1 regulates cellular respiration, metabolism and proliferation.

To determine whether the observed shift in mitochondrial morphology does in fact depend on the presence of the MacroD1 protein, we transiently re-introduced MacroD1 fused to GFP (MD1GFP) into the MD1KO cells and found that MD1-GFP-expressing cells did indeed partially rescue the mitochondrial morphology phenotype observed in MD1KO cells (Figures 3.17 A and B).

In summary, we found that deletion of the mitochondrial matrix macrodomain protein, MacroD1, leads to a number of mitochondrial and cellular phenotypes including; fragmented mitochondrial morphology, decrease in cell size, bioenergetic profile changes from oxidative phosphorylation to glycolysis, along with a reduction of proliferation rate (Figures 3.15 and 3.16). Further, we were able to show that the mitochondrial morphology phenotype is partially rescuable by transient gene supplementation (Figure 3.17).



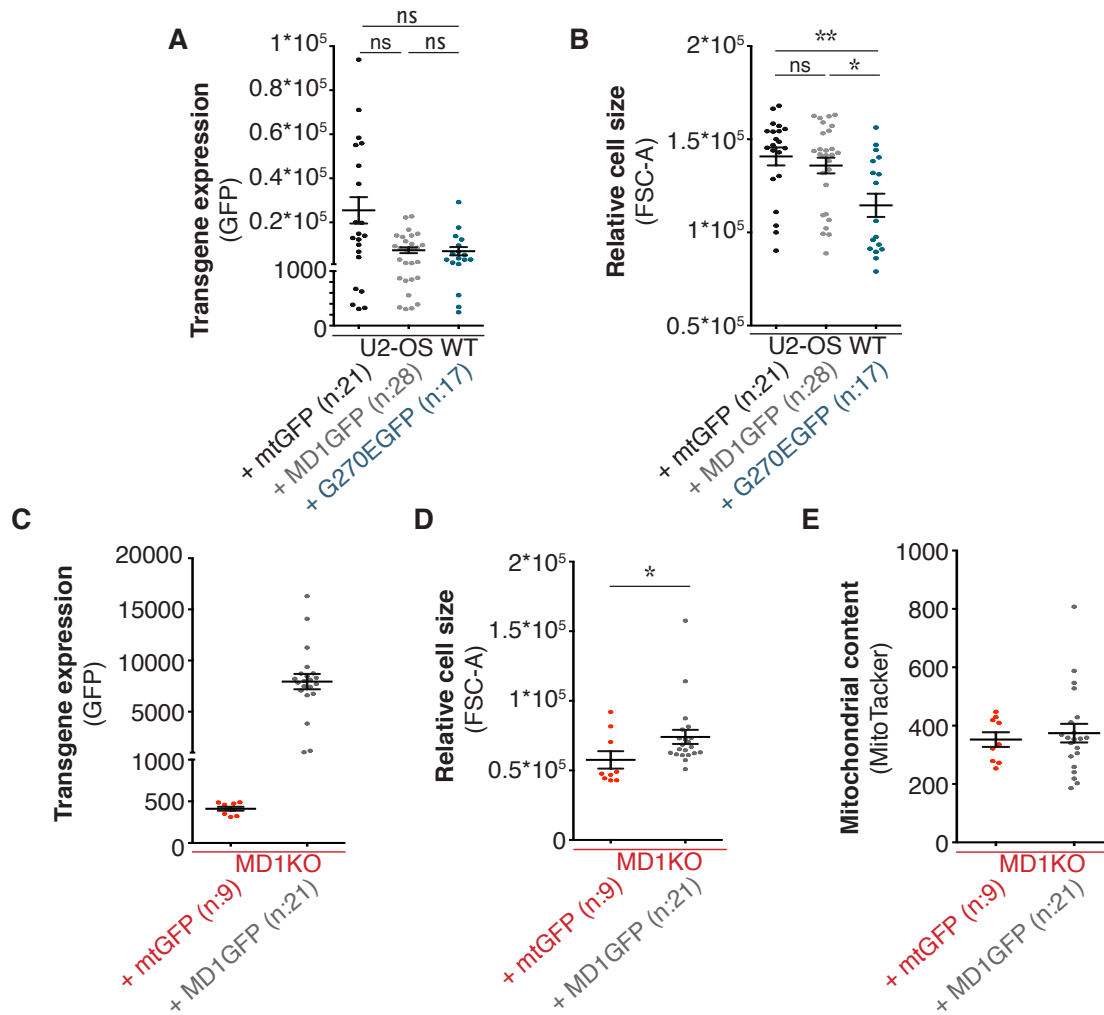
**Figure 3.17: The phenotypic changes are MacroD1-dependent and rescuable by gene-supplementation.**

A) Transient transfection of MacroD1 fused to GFP for 24-48 hours is sufficient to partially rescue mitochondrial morphology. In the merged images: nuclei are in blue, MD1-GFP is in green and MitoTracker in red. B) Classification by 5 independent individuals confirmed a shift of mitochondrial morphology from punctuate to filamentous in the MD1KO cells upon MacroD1-supplementation. Classification was performed blinded on 52 MD1KO and 55 transfected MD1KO + MD1-GFP cells.

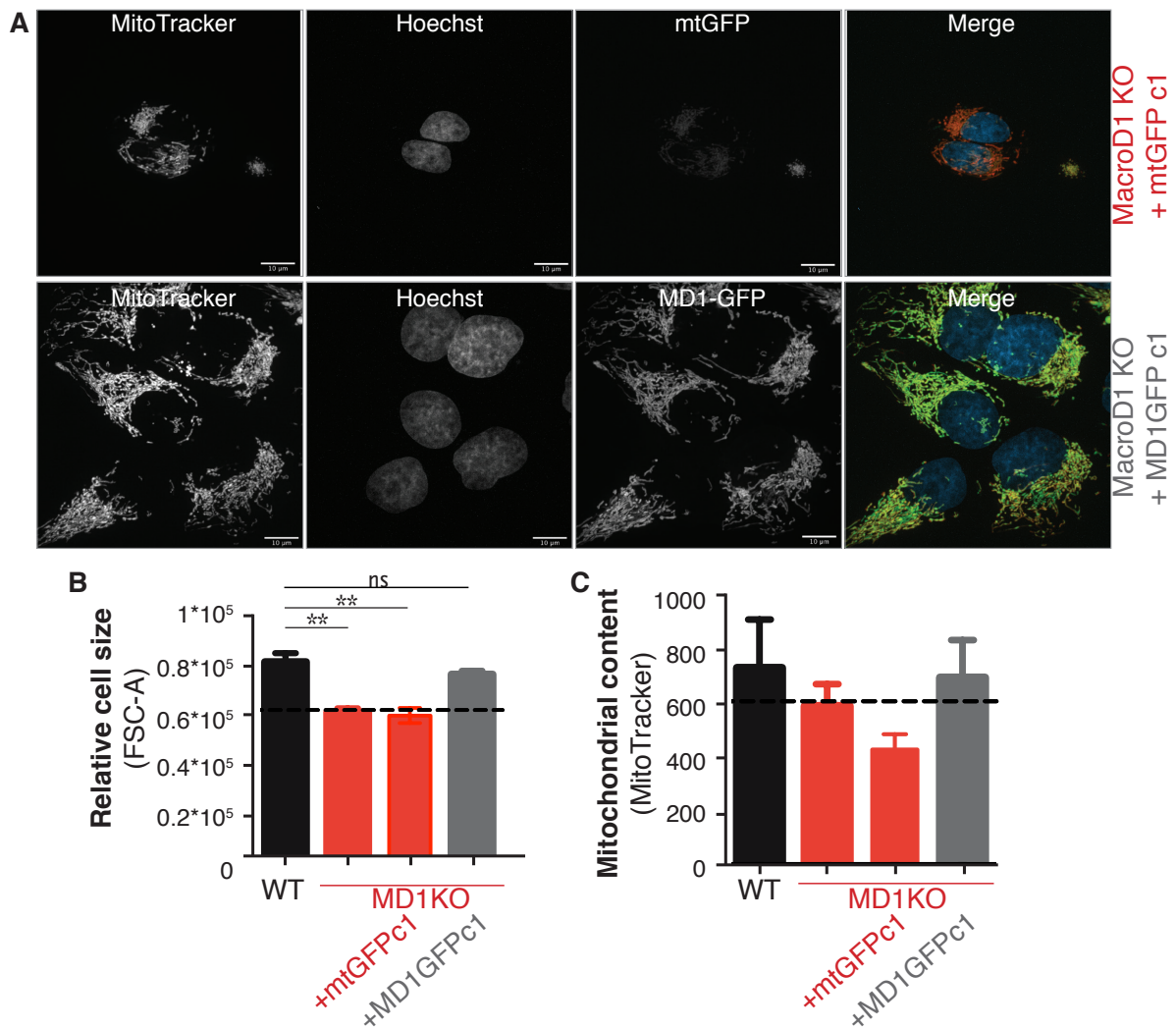
### 3.3.2 Rescue by gene supplementation depends on ADPr-binding capacity of MacroD1. Over-expression of ADPr-binding mutant exhibits dominant negative effect

Since MacroD1 binds the NAD<sup>+</sup>-metabolite ADP ribose and can hydrolyse mono-ADP-ribosylated proteins *in vitro*, we next tested whether a MacroD1 site-directed mutant that disrupts the interaction with ADP ribose (ADPr) and ADP-ribosyl-hydrolysis exhibits mitochondrial phenotypes, mitochondrial content and cell size. In order to do this, we sought to generate stable cell lines expressing mitochondrial GFP (mtGFP), wild-type MacroD1 (MD1GFP) or the ADPr-binding-deficient MacroD1 mutant (G270EGFP) (Chen et al., 2011) in both wild-type and MD1KO U2-OS cell lines (Figure 3.18).

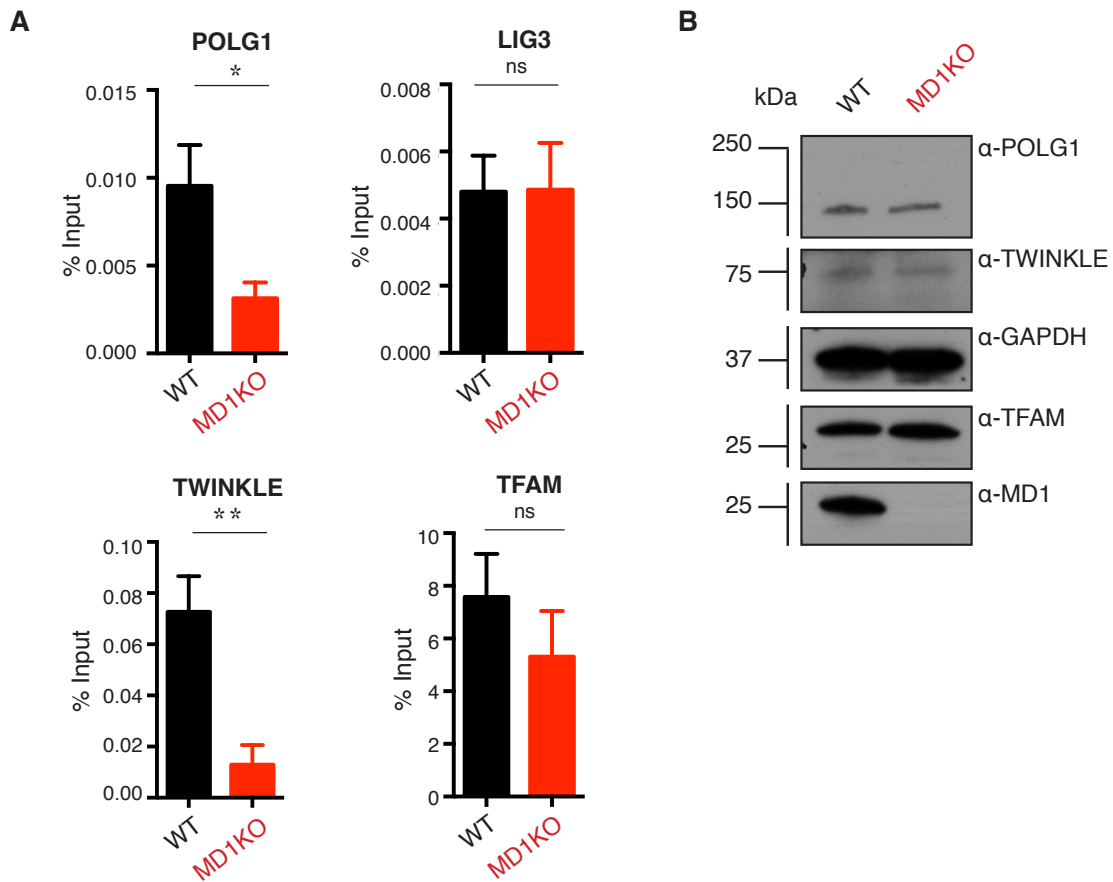
Surprisingly, we were unable to obtain stable cell lines expressing the ADPr-binding deficient MacroD1 mutant in the MD1KO background. In contrast, transfection of the G270EGFP construct was successful in the wild-type U2-OS background (Figures 3.18), as well as under transient conditions in the MD1KO cell line (data not shown). Our inability to obtain stable clones expressing the mutant MacroD1 transgene was thus not due to an inherent problem in cells expressing the variant protein, but rather indicates that MD1KO cells do not tolerate the ADPr-binding deficient MacroD1 mutant. Consistently, stable overexpression of the G270E mutant in the wild-type U2-OS cell line leads to a significant decrease in cell size (Figures 3.18 B), while overexpression of the wild-type protein does not lead to a shift in cell size. This indicates that the ADPr-binding deficient MacroD1 mutant has a dominant-negative effect over endogenous MacroD1. To test whether wild-type MacroD1 rescues the cell size phenotype of the MD1KO cell line, we stably re-introduced MacroD1. When fused to GFP, the MacroD1 protein increased the size of the MD1KO cells compared to GFP control, as measured both across all generated cell lines (Figures 3.18 D) and in a representative MD1KO clone rescued with the MD1GFP transgene compared to MD1KO and mtGFP controls (Figures 3.19 A and B). Importantly, the MacroD1 transgene fully rescues mitochondrial content in the MD1KO cells, while the mtGFP control does not (Figures 3.19 C). Targeted interference of the mitochondrial matrix protein MacroD1 thus disrupts mitochondrial homeostasis and reveals an important function for the protein's ADP-ribose-binding pocket in regulating cell size.



**Figure 3.18: Rescue by gene supplementation depends on ADPr-binding capacity of MacroD1.** Flow cytometric analysis. The individual measurements (dots) are the median of one cell population, the mean and  $\pm$  SEM are calculated based on measurements across all of the generated stable cell line clones. A) Over-expression of ADPr binding mutant exhibits dominant negative effect. Transgene expression of all wild-type expressing distinct transgenes (WT + mtGFP, WT + MD1-GFP and WT + MD1G270E-GFP), as measured by fluorescent intensity of GFP. Statistical analysis is performed using Kruskal-Wallis with Dunn's multiple comparison test. B) Comparative cell size of wild-type cells expressing distinct transgenes (WT + mtGFP, WT + MD1-GFP and WT + MD1G270E-GFP). Statistical analysis is performed using Kruskal-Wallis with Dunn's multiple comparison test, \* $p < 0.05$ , \*\* $p < 0.01$ . C) Transgene expression of all MD1KO + MD1-GFP and MD1KO + mtGFP cell lines, as measured by fluorescent intensity of GFP. D) Comparative cell size of all MD1KO + MD1-GFP and MD1KO + mtGFP cell lines. Statistical analysis is performed using a two-tailed Mann-Whitney test, \* $p < 0.05$ . E) Comparative mitochondrial content, measured by fluorescent intensity of MitoTracker Red.



**Figure 3.19: Characterization of MacroD1-rescue cell line** A) Confocal images of the MD1KO + mtGFP cl.1 and MD1KO + MD1-GFP cl.1 cell lines, confirming both expression and localization of the transgenes. In the merged images: nuclei are in blue, mtGFP and MD1-GFP are in green and MitoTracker is red. B - C) Flow cytometric measurements, performed in biological triplicate, on the BD LSRFortessa flow cytometer. Measurements calculated from median of a cell population and displayed as mean  $\pm$  SEM. B) Comparative cell size, as measured by forward scatter. Statistical analysis is performed using ordinary one-way ANOVA with Turkey's multiple comparisons test, \* $p < 0.05$ , \*\* $p < 0.01$ . C) Comparative mitochondrial content, measured by fluorescent intensity of MitoTracker Red.



**Figure 3.20: MacroD1 affects complex formation of core mitochondrial DNA replication factors on mitochondrial DNA.** A) MIP assay demonstrating occupancy of core replication factors on mtDNA in the wild-type U2-OS (WT; black) versus the MD1KO (red) cell lines. POLG1 and LIG3 IPs are performed in 6 biological replicates. TWINKLE and TFAM IPs are performed in 5 biological replicates. Data is normalized to percent input, biological replicates are displayed as mean  $\pm$  SEM. Statistical analysis is performed using a two-tailed, unpaired t-test, \* $p < 0.05$ , \*\* $p < 0.01$ . B) Changes in complex formation of replication factors on mtDNA is not due to decreased expression of factors, as shown by western blot of endogenous core mitochondrial replication factors comparing expression levels between the MD1KO and wild-type U2-OS cell lines.



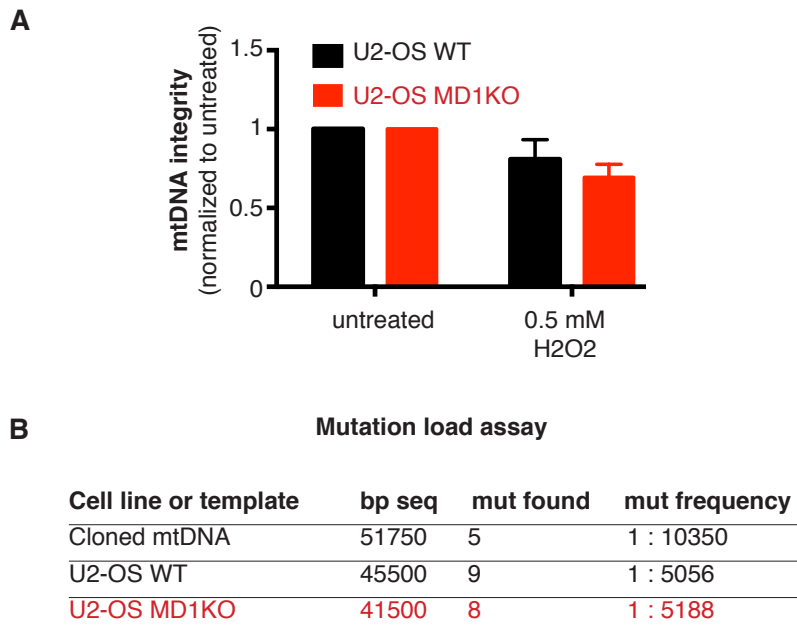
### 3.3.3 MacroD1 affects complex formation of core mitochondrial DNA replication factors on mitochondrial DNA

As MacroD1 interacts with core mtDNA replication factors in an ADP-ribose dependent manner and MacroD1 knockout cells exhibit mitochondrial phenotypes, we tested whether MacroD1 deletion impacts mtDNA homeostasis.

We first sought to determine whether the deletion of MacroD1 impacts the function of the ADP-ribosylated MacroD1 interactors, the core components of the mtDNA replication machinery, with the mtDNA genome. Using a mitochondrial version of chromatin immunoprecipitation (MIP) to determine the interaction of different DNA replication and repair factors to mtDNA in U2-OS wild-type compared to MD1KO cells and found that POLG1 and TWINKLE associate significantly less with mtDNA in the MD1KO cell lines compared to wild-type control. In contrast, LIG3 – and TFAM – mtDNA interaction remains largely unchanged (Figures 3.20 A).

To determine whether the decreased association of POLG1 and TWINKLE may be due to altered protein levels, we compared the expression of POLG1, TWINKLE and TFAM by western blot analysis in wild-type and MD1KO cell lines (Figures 3.20 B). MacroD1 deletion does not alter the steady-state levels of the nuclearly encoded mtDNA replication factors (Figures 3.20 B). We repeated the expression analysis in the HEK293-T-REx WT versus MD1KO cell lines with identical results (Söllner et al., 2020b).

Furthermore, in a collaboration with Francesca Sacco, we were able to show by mass spectrometry-based analysis of HEK293-T-REx WT versus MD1KO mitochondria that MacroD1 is deleted, but that there is no significant change in any of the components of mtDNA replication (data not shown). Thereby, revealing that the decreases in protein–mtDNA interactions are not caused by a depletion of the core mtDNA replication machinery.



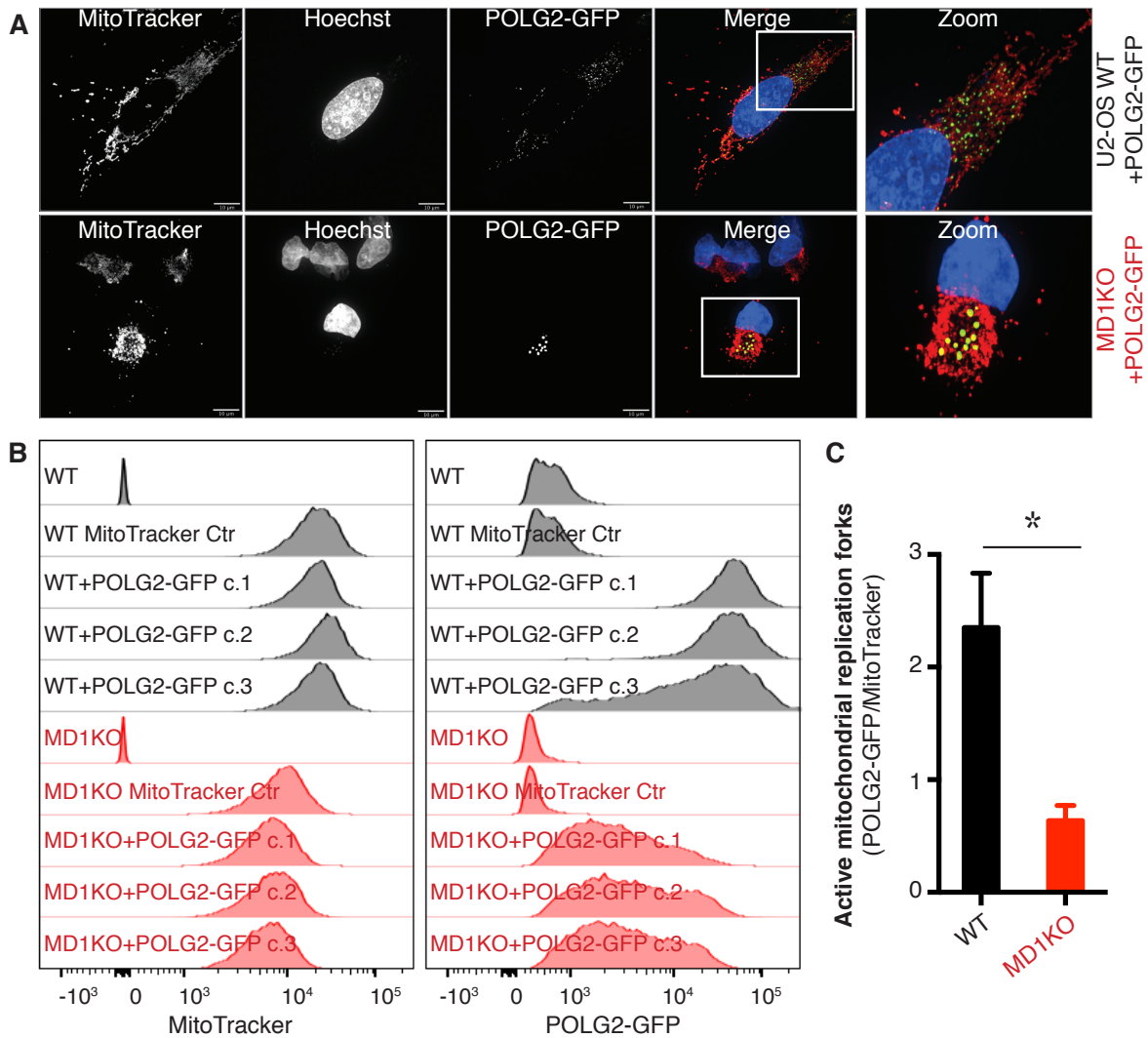
**Figure 3.21: Change in DNA and DNA factor complex formation does not lead to a significant change in mtDNA integrity or increased mutation load.** A) LA-PCR, measuring mtDNA integrity, before and after 1 hour of H<sub>2</sub>O<sub>2</sub>-treatment in the WT (black) and MD1KO (red) U2-OS cell lines. Experiment is performed in six biological replicates, and technical duplicate per replicate. Data is displayed as mean  $\pm$  SEM. B) Mutation load as measured by post-PCR cloning and Sanger sequencing. The analysed sequenced region spans 500bp, approximately 100 clones per condition were sequenced.

### 3.3.4 MacroD1 deletion in U2-OS cells does not affect mtDNA mutation load

As POLG, TOP1 and LIG3 are not only mtDNA replication factors but also implicated in mtDNA repair (Copeland, 2010), we decided to investigate mtDNA integrity and repair in the MD1KO cells. We wanted to see whether mtDNA in the MD1KO cell line is more susceptible to DNA stress than the parental control cell line, using long-amplicon PCR (LA-PCR). Upon hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>)-treatment, we observed only a slight decrease of mtDNA integrity in the MD1KO versus the wild-type cell-lines (Figure 3.21 A), which we repeated in the HEK293-T-REx WT versus MD1KO cell lines with identical outcome (Söllner et al., 2020b).

Next, the apparent tendency for decreased mtDNA integrity and reduced replication factor–mtDNA complex formation prompted us to test whether we would detect an increase in mutations in the assayed mtDNA sequence. We thus performed mtDNA mutation load analysis and found that the amplified area for the MIP assay and surrounding ~ 500bp do not contain more mutations in the MD1KO cells compared to wild-type controls (Figure 3.21 B).

Overall, our results reveal that both POLG1 and TWINKLE associate significantly less with mtDNA in the MD1KO cell line compared to the wild-type control and that these changes are independent of changes in expression levels, mtDNA integrity or mutation load.



**Figure 3.22: Deletion of MacroD1 decreases active mtDNA replication forks.** A) Transient transfection of NGFP-POLG2 into MD1KO and wild-type (WT) U2-OS cells. There are visibly less, yet brighter, NGFP-POLG2 dots (active replication forks) in the MD1KO cells. In the merged images: nuclei are in blue, NGFP-POLG2 is in green and MitoTracker in red. B) Representative histogram, from one biological replicate, depicting fluorescent intensity of MitoTracker Red and NGFP-POLG2 from U2-OS WT and MD1KO controls as well as 3 WT and 3 MD1KO clones with stable NGFP-POLG2 expression, as measured by flow cytometry. On the x-axis is fluorescent intensity of either MitoTracker Red or NGFP-POLG2, the y-axis is displayed as a percentage of the maximum count. C) Active mtDNA replication forks in MD1KO and WT + NGFP-POLG2 clones. NGFP-POLG2 fluorescence is normalized to mitochondrial content, three biological replicates are displayed as mean  $\pm$  SEM. Two-tailed unpaired t-test. \* $p < 0.05$ .

### 3.3.5 Deletion of MacroD1 decreases active mtDNA replication forks and mtDNA synthesis

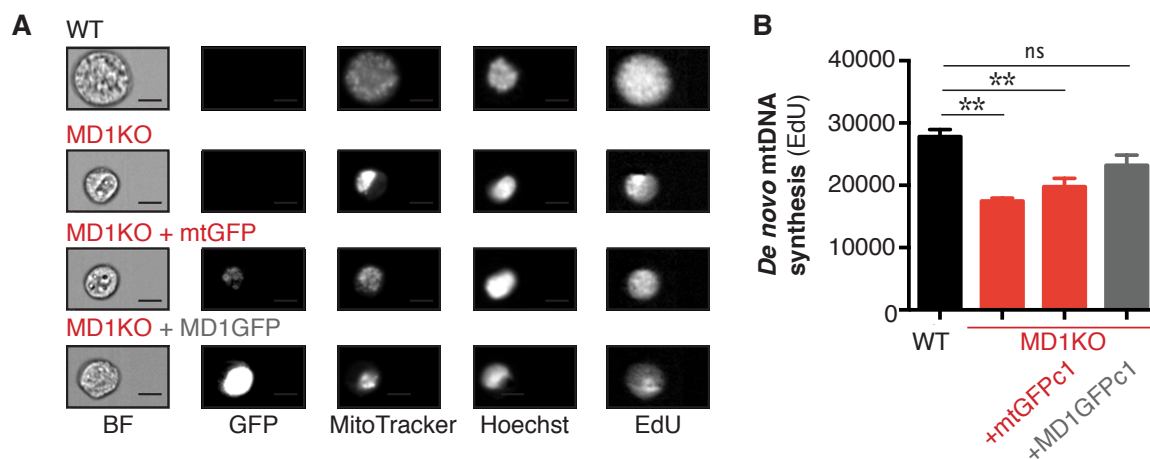
Since POLG1 and TWINKLE are less prevalent on mtDNA in cells with MacroD1 deletion, we investigated mtDNA replication in the MacroD1 deletion context. To do this we decided to utilize our novel POLG2-Assay in which we directly monitor active replication forks (subchapter 3.1).

Initially, we transiently transfected NGFP-POLG2 into both the U2-OS WT and MD1KO cells. There were visibly less, yet brighter/bigger NGFP-POLG2 spots in the MD1KO cells (Figure 3.22 A), indicative of a replication or mtDNA molecule segregation deficit subsequent to replication. This would make foci counting and thereby quantifying mtDNA replication unfeasible.

To further explore the importance of MacroD1 in active mitochondrial replication, we studied three wild-type clones with stably inserted NGFP-POLG2 and three distinct MD1KO clones (Figure 3.22 and generated as described in subchapter 3.1).

To evaluate the relative number of active replication forks per mitochondria, we measured the fluorescence of GFP from the NGFP-POLG2 and of MitoTracker, normalizing GFP fluorescence to the MitoTracker signal. This allowed us to comparatively quantitate active mtDNA replication forks between cell-lines, independent of mitochondrial content. Our analysis showed that MD1KO mitochondria contain significantly less active mtDNA replication forks compared to wild-type cells (Figure 3.22).

Using the NGFP-POLG2 assay, we thus demonstrated that the mitochondria of MD1KO U2-OS cells have significantly less active replication forks compared to wild-type U2-OS cells.

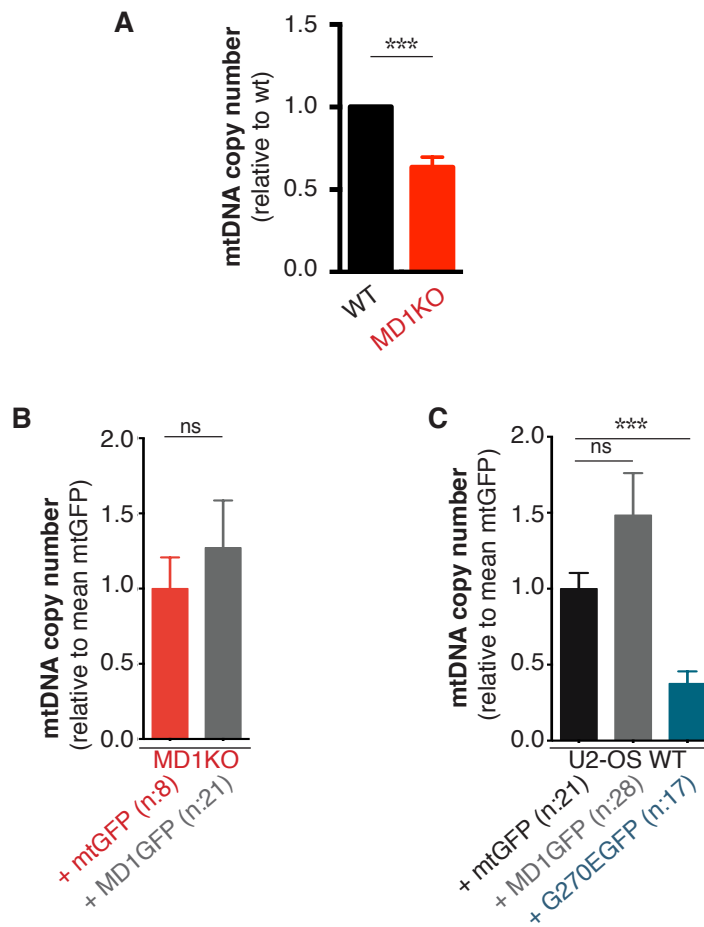


**Figure 3.23: Deletion of MacroD1 lowers de novo mtDNA synthesis.** Newly synthesized mtDNA is quantified by measuring median EdU fluorescent intensity within mitochondria of a given population (1 population = 60000 cells). A) Representative images of the fixed and stained wild-type U2-OS (WT), MD1KO, MD1KO + mtGFP and MD1KO + MD1-GFP cell lines. B) Measurements are taken in biological triplicate and displayed as mean  $\pm$  SEM. Statistical analysis is performed using a 2way ANOVA with Turkey's multiple comparison test. \*\* $p < 0.01$ .

Next, we wanted to address whether the reduction in active mtDNA replication forks correlates with a reduction in de novo mtDNA synthesis. A frequently-used method to quantify mtDNA synthesis is EdU-incorporation (Haines et al., 2010) coupled with foci counting. We implemented an assay based on this principle by coupling EdU-labelling with flow-cytometric imaging (Figure 3.23).

As explained in the methods, we applied masks for the whole cell, for the nucleus, and colocalized mitochondrial and EdU-staining (see section 2.2.21). This allowed us to measure the intensity and relative amount of EdU signal from mitochondria. Using this method, we observed that MD1KO cells have significantly less newly synthesized mtDNA compared to the parental control cell line. Importantly, reintroduction of MD1-GFP into the MD1KO cell line rescues the deficit in de novo mtDNA synthesis (Figure 3.23).

Taken together, the POLG2-Assay along with EdU-incorporation and flow cytometric imaging reveal a role for MacroD1 in mitochondrial DNA replication.



**Figure 3.24: MacroD1 affects mtDNA copy number in an ADPr-dependent manner.** A) Comparing mtDNA copy number between wild-type (WT) and MD1KO cell U2-OS lines. Data is normalized to U2-OS WT. Measurements taken in seven biological replicates and displayed as mean  $\pm$  SEM. Statistical analysis is performed using unpaired two-tailed t-test. \*\*\* $p < 0.0001$ . B) To compare mtDNA copy number between the MD1KO cell lines stably expressing mtGFP and those expressing MD1-GFP, we normalized all qPCR data to the mean of the MD1KO + mtGFP cell lines. Data is displayed as mean  $\pm$  SEM. Statistical analysis is performed using the two-tailed Mann-Whitney test. C) To compare mtDNA copy number between the WT cell lines stably expressing mtGFP, MD1-GFP and MD1G270E-GFP we normalized all qPCR data to the mean of the WT + mtGFP cell lines. Data is displayed as mean  $\pm$  SEM. Statistical analysis is performed using the Kruskal-Wallis test with Dunn's multiple comparison test. \*\*\* $p < 0.001$ .



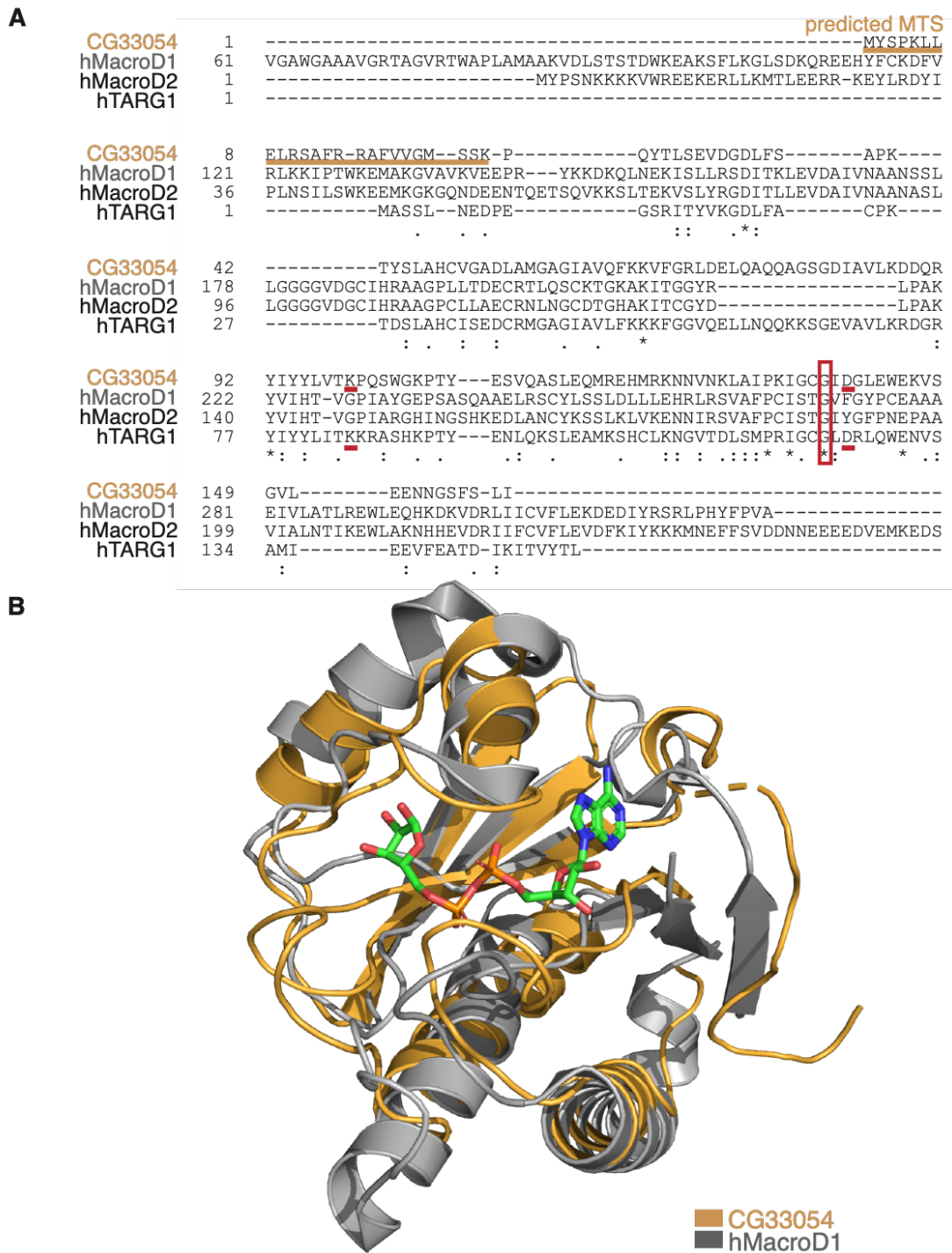
### 3.3.6 MacroD1 affects mtDNA copy number in an ADPr-dependent manner

Since MacroD1 impact the number of active mtDNA replication forks and the synthesis of new mtDNA, we sought to determine whether this would translate into a depletion of mtDNA copy number. To measure relative mtDNA levels, we performed quantitative PCR using the mtDNA primers, as well as a primer set for a nuclear housekeeping gene in order to normalize for cell number. Compared to U2-OS WT the MD1KO cells contained significantly less mtDNA (Figure 3.24 A). This is consistent with our previous results, which show reduced active replication and mtDNA synthesis (Figures 3.22 and 3.23).

Next, we checked whether this change in mtDNA content can be observed across multiple stable MacroD1-rescue and -overexpression cell lines (Figure 3.24 B and C).

Although a trend towards increased mtDNA copy number was discernible in the MD1KO + MD1-GFP cell line compared to the MD1KO + mtGFP control, the difference is not statistically significant (Figure 3.24 B). However, overexpression of the ADPr binding-deficient mutant in wild-type U2-OS cells, drastically depleted cells of mtDNA (Figure 3.24 C).

These data show that the depletion or out-competition of MacroD1 capable of binding ADP-ribosylated substrates dysregulates mtDNA replication.



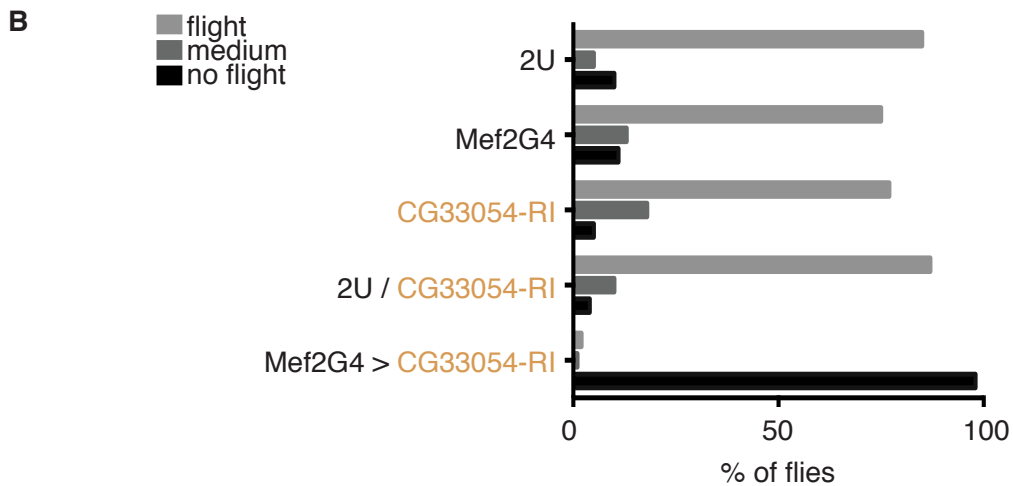
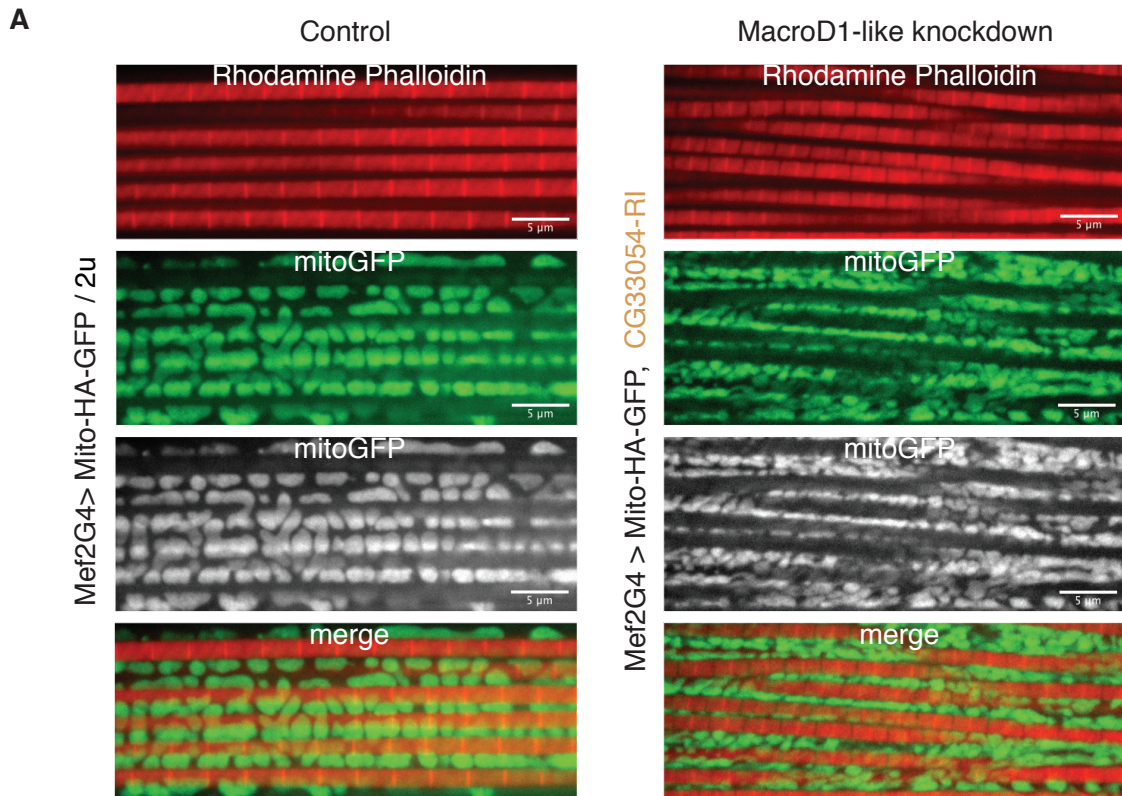
**Figure 3.25: CG33054, mitochondrial macrodomain protein in *Drosophila melanogaster*.** A) Protein alignment of human macrodomain ADP-ribosylhydrolases MacroD1 (grey), MacroD2 (black) and TARG1 (black) with *Drosophila* macrodomain protein CG33054 (yellow). The MitoProtII predicted MTS of CG33054 is underlined in yellow. The TARG1 active sites K84 and D125 are underlined in red and the ADPr-binding site (G270 in MacroD1) has a red box around it. B) Homology model of MacroD1 (grey - shown from residue 253-307) and CG33054 (yellow). ADPr is modelled into the MacroD1 structure based on the highly homologous MacroD2 structure (*homology model - courtesy of Gunnar Knobloch*).

### 3.3.7 Depletion of a mitochondrial macrodomain protein impacts on the organismal level

In *in vitro* and *in cellulo* studies we revealed that MacroD1 interacts with essential mtDNA replication and transcription factors (subchapter 3.2) and that depletion of MacroD1 leads to the dysregulation of mitochondrial DNA, affecting both cellular and mitochondrial morphology and bioenergetics (subchapter 3.3). Our subsequent question was whether we would see something similar on an organismal level. POLG1, POLG2, LIG3, TWINKLE, TOP1, TOP2A and TFAM are all conserved in *Drosophila melanogaster*, making fruit flies optimal model organisms to study mtDNA replication.

In addition to the core mtDNA replication factors, we identified a MacroD1-like macrodomain protein, CG33054, which according to MitoProtII, TargetP and iPSORT contains a strong putative mitochondrial targeting sequence (Figure 3.25 A) and has previously been experimentally identified in the mitochondria of *Drosophila* by mass spectrometric analysis (Lotz et al., 2014). Gunnar Knobloch generated a homology model for us, using the core macrodomain of human MacroD1 (residues 253 to 307) and the MacroD1-like CG33054 *Drosophila* amino acid sequence. Gunnar modelled ADP-ribose into the MacroD1 structure based on the highly homologous MacroD2 structure. The RMSD between the two shown structures is 4.819 Å (Figure 3.25 B).

With the help of both conservation analysis and homology model (Figure 3.25) it appears very likely that the function of the CG33054 protein product in fly mitochondria is orthologous to that of MacroD1 in human mitochondria. As expression data in both mouse and man indicate increased MacroD1 expression levels in mitochondria-rich tissue such as muscle (Agnew et al., 2018; Žaja et al., 2020), as well as the knowledge that the indirect flight muscle of adult thorax are loaded with mitochondria due to their high metabolic activity (Wang et al., 2016), we decided to perturb the *Drosophila* MacroD1-like protein in developing muscle in order to test whether this perturbation would affect mitochondrial morphology and possibly flight capacity.



**Figure 3.26: Mitochondrial Flight Muscle and Flight phenotypes upon MacroD1-like mitochondrial macrodomain protein knockdown in *Drosophila melanogaster*.** A) The structure of mitochondria is compromised in CG33054-depleted fibrillar indirect flight muscle. Mitochondria seen by mitoGFP expression (green and grey view for better contrasts), Actin is stained using rhodamine phalloidin (red). In the merged image mitochondria are in green and actin is red, the scale bar measures 5  $\mu$ m. B) Male offspring with depleted CG33054 in all muscle from the myoblast stage are not able to fly, while controls are all good fliers.

To knockdown the MacroD1-like protein we used a CG33054-specific shRNA knockdown line (KK108302 from VDR) and crossed male offspring with female virgins from a Mef2-GAL4 driver line expressing mitochondrial GFP (Mef2-G4 UAS-Mito-HA-GFP). Using this driver-line we depleted the MacroD1-like CG33054 protein from the myoblast stage onwards while also making mitochondria visible due to GFP expression. As control we used young male offspring from the wild-type 2u line crossed with the Mef2-G4 UAS-Mito-HA-GFP driver as control (Figure 3.26 A).

Comparing mitochondrial morphology in the wild type fibrillar indirect flight muscle with the mitochondria in the CG33054-depleted flies, we see that there is a shift of mitochondrial morphology towards smaller, more fragmented and less tubular mitochondria when the mitochondrial macrodomain is knocked down. As flight muscle is a high energy demand tissue, mitochondrial perturbations can lead to dire metabolic consequences. Such as reduced flight capacity.

Next, we wanted to test whether MacroD1-like CG33054-depletion would affect *Drosophila* flight capacity. We again performed knockdown by crossing male offspring from the CG33054-specific shRNA knockdown line with female virgins from a Mef2-GAL4 driver, this time without the mitochondrial GFP expression. In the flight experiments we used young male offspring from the wild-type 2u line crossed with the Mef2-GAL4 driver as control, as well as male offspring from the uncrossed 2u, Mef2-GAL4 and KK108302 lines. From the flight test we saw that all controls performed in the flight test as expected from strong fliers. However, when the MacroD1-like CG33054 protein was depleted in muscle cells, the flies were incapable of flight (Figure 3.26 B).

We have observed disturbed mitochondrial morphology, as well as a no flight phenotype (Figure 3.26) upon MacroD1-like CG33054-depletion. Reflecting the mitochondrial and cellular phenotypes observed in human cancer cell lines with MacroD1-deletion (section 3.3.1).

## 4. Discussion

To sustain mitochondrial homeostasis and with that the equilibrium of not only metabolism but also the maintenance of all other cellular processes, the integrity and controlled replication of mtDNA is vital. Dysregulation in the fidelity of replication is the leading cause behind increases in mtDNA mutation load and changes of mtDNA copy number (Gustafsson et al., 2016), both of which are underlying causes of numerous pathologies (Ciafaloni et al., 1992; Coxhead et al., 2016; Dolle et al., 2016; Hoekstra et al., 2016; Ikeda et al., 2018; Pfeffer and Chinnery, 2013). In this thesis we developed a novel fluorescent flow cytometry-based mtDNA replication assay (subchapter 3.1), discussed in 4.1. Additionally, we identified a novel and important role for the ADP-ribose-binding macrodomain protein MacroD1 in regulating mitochondrial DNA replication (subchapters 3.2 and 3.3), discussed in 4.2 and 4.3, respectively. Supporting the hypothesis that dynamic mitochondrial ADP-ribosylation events are essential for mitochondrial and cellular homeostasis (subchapter 3.3).

### 4.1 The novel POLG2-assay: monitoring mtDNA replication

In this thesis we have developed and tested an innovative POLG2-assay, which can be used to monitor mtDNA replication *in cellulo* (subchapter 3.1). To our knowledge we thus generated the first automatable mtDNA replication assay, measured using a flow cytometric analyser. In this assay we can in parallel to mtDNA replication changes, also monitor cell size, mitochondrial load, and protein expression levels in living and fixed cells. Measurements and conditions not easily quantifiable simultaneously by conventional methods, such as foci counting. At the same time our sample size is much greater than in conventional methods. While EdU-incorporation with subsequent *foci* counting is performed on maybe 10 cells in three replicates, we can assess NGFP-POLG2 expression for tens of thousands of cells per biological replicate.

We tested our assay by using a number of different methodologies. We reduced mtDNA replication by pharmacological inhibition (Figure 3.2 and 3.3) and by overexpression of an

exonuclease artificially targeted to the mitochondrion (Figure 3.4). To increase mtDNA replication we overexpressed both POLG1 (Figure 3.5) and TWINKLE (Figure 3.6), the mtDNA helicase which has been attributed as a limiting factor in mtDNA replication. As anticipated, NGFP-POLG2 protein levels and fluorescent intensity decreased with both ddC-treatment and as a result of mitochondrial exonuclease overexpression and increased with POLG1 and TWINKLE overexpression.

Subsequently, we used a slightly altered version of our novel assay, in which we used 6 different monoclonalised cell lines overexpressing NGFP-POLG2, three in the wild-type background and three in MacroD1KO background, to compare mtDNA replication activity between these conditions. We found that the MacroD1-deletion cells overexpressing NGFP-POLG2 exhibited significantly less mtDNA replication per mitochondrion. Using these cell lines, we also saw that foci counting would not have been a viable option, as the foci in the MacroD1KO cells were much larger and brighter, but with less spread. Thus, not only making our assay the first *in cellulo*, flow-cytometry based mtDNA replication assay, but also the only viable option for our specific MacroD1-deletion conditions.

## 4.2 MacroD1 interacts with the core mtDNA replication machinery

As there is little insight into both mitochondrial ADP-ribosylome and the biological role of proteins capable of recognizing this dynamic post-translational modification, we took advantage of the mitochondrial matrix localization of the ADP-ribosyl-binding MacroD1 protein (Söllner et al., 2020b; Žaja et al., 2020) and identified a number mitochondrial matrix proteins interacting with MacroD1 in both an ADPr-dependent and independent manner (subchapter 3.2). We show a robust link between MacroD1 and mitochondrial DNA replication and transcription factors at the biochemical level. Importantly, we find that the core machinery required for mtDNA replication emerges as one of the most prominent sets of mitochondrial MacroD1-interacting proteins. POLG1, TWINKLE, TOP1, TOP2B, LIG3, POLRMT and TFB2M, core factors involved in both transcription and replication of mtDNA, are novel interaction partners of the mitochondrial matrix macrodomain protein MacroD1.

Engineered mutant MacroD1 proteins further reveal that the ADP-ribose binding pocket of the MacroD1 macrodomain module is key to the specific interaction with most of the enriched DNA replication components. This is further confirmed by the use of an orthogonal approach with engineered mtTARG1. A protein here artificially located to mitochondria, that recognizes ADP-ribosylated proteins through its ADP-ribose binding macrodomain.

In support of the canonical role of the ADP-ribose binding pocket of macrodomains as a PTM recognition module, our results demonstrate that the core replication factors POLG1, TWINKLE, TOP1, TOP2B, and LIG3 interact with MacroD1 in an ADPr-dependent manner. In contrast, two factors required for mtDNA transcription, POLRMT and TFB2M, interact with MacroD1 irrespective of its ADPr binding capacity, but do not enrich with the orthogonal, mtTARG1. Thereby indicating the existence of distinct mechanisms of interaction between MacroD1 and mtDNA transcription/replication factors.

Consistent with the ADP-ribosyl-binding function of MacroD1, we find that a far western method, related to that pioneered by others (Gibson et al., 2017; Khadka et al., 2015; Wu et al., 2007), verifies ADP-ribosylation status of TWINKLE, LIG3, POLG1 and TFAM, while also indicating that POLG2 is not modified by this PTM. TFAM is the mtDNA packaging factor that is also involved in mtDNA transcription initiation (Falkenberg, 2018; Ngo et al., 2011). Hence, its ADP-ribosylation underpins the close connections between mtDNA transcription and replication.

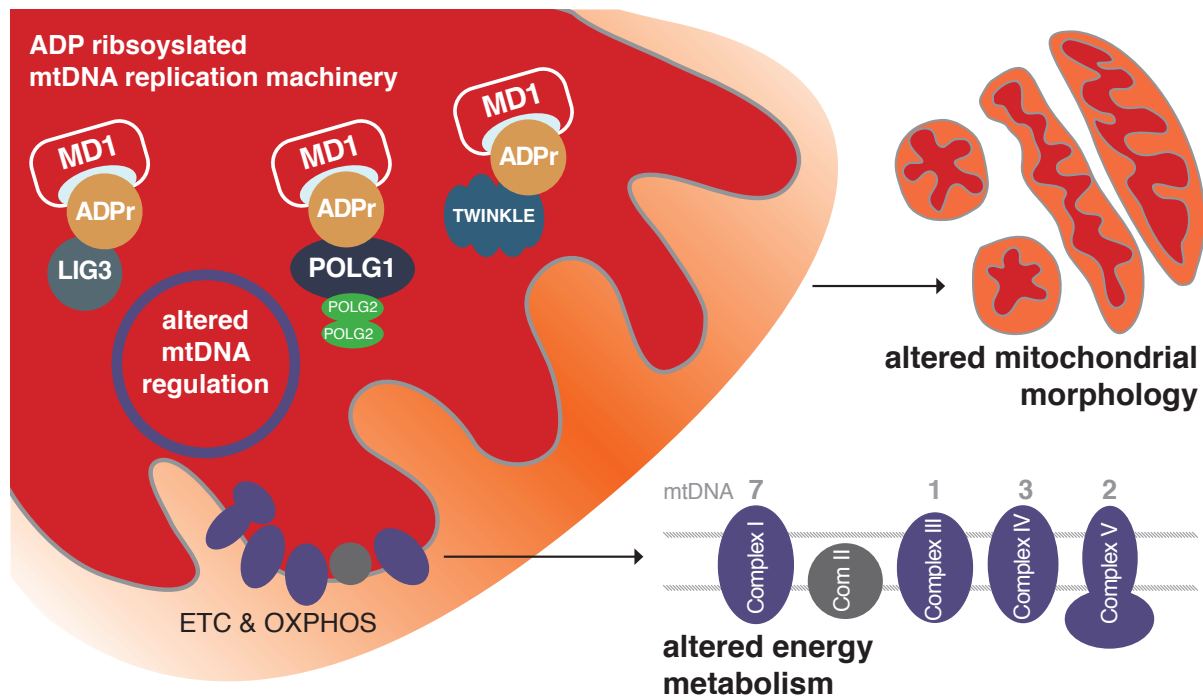
We have not yet identified the ADP-ribosylation sites in mtDNA replication factors. However, we identified two factors which might link mitochondrial ADP-ribosylation to Serine-phosphorylation, namely; TAOK1 and AURKAIP1. TAOK1 is a Serine/threonine-protein kinase. AURKAIP1 is a known interactor and negative regulator of the serine/threonine kinase AURKA, while also being directly involved in the expression of mitochondrially encoded proteins as a member of the mitochondrial ribosome proteins (Koc et al., 2013). Both of these factors interact with MacroD1 in an ADPr-dependent manner and according to PhosphoSitePlus (Hornbeck et al., 2004) (accessed in 2020 via:



<https://www.phosphosite.org/homeAction>) AURKAIP1, TAOK1, POLG1, LIG3, TOP1, TOP2A, TWINKLE and TFAM all have Ser-phosphorylation sites. As the interplay between Ser-phosphorylation and Ser-ADP-ribosylation has previously been highlighted (Larsen et al., 2018), it would be very interesting to see which residues in our Mitochondrial ADP-ribosylome are modified by the PTM and whether there is a similar interplay between phosphorylation and ADP-ribosylation within mitochondria? Maybe the ADP-ribosylation status of Serine/threonine phosphorylation factors is an additional level of regulation. Altering activity or specificity of these factors – which then in turn could lead to either phosphorylation or ADP-ribosylation of and MacroD1 interaction with the mtDNA replication factors.

### 4.3 MacroD1 is a vital factor in the regulation of mtDNA replication

From the MacroD1-interactome and the MacroD1-specific ADP-ribosylome we learned that some factors require the ADP-ribosyl-binding function of the MacroD1 protein, while others likely interact with MacroD1 through a surface that is not conserved in the TARG1 homolog of MacroD1. The strong biochemical association of mtDNA replication factors with MacroD1 made us question whether MacroD1 and mitochondrial ADP-ribosylation are somehow involved in mtDNA homeostasis. To elucidate the physiological function of MacroD1, we utilized different MacroD1 knockout cell lines and gleaned a number of MacroD1-specific phenotypes. These phenotypes include; fragmented mitochondrial morphology, decrease in cell size, bioenergetic profile changes from oxidative phosphorylation to glycolysis, along with a reduction of proliferation rate. These phenotypes are all consistent with, and have frequently been described in, the context of pathophysiological changes in mtDNA regulation.



**Figure 4.1: MacroD1 & ADP-ribosylation regulate mtDNA replication, altering mitochondrial morphology and metabolism.** Deletion or dysregulation of the mitochondrial matrix protein MacroD1 results in a dysregulation of mtDNA replication and homeostasis. The mitochondrial genome codes for 13 essential components of the electron transport chain, as well as for tRNAs and rRNAs needed to translate these. In our model, we hypothesize that the dysregulation of mtDNA precedes the morphological and bioenergetic changes seen upon MacroD1 dysregulation.

The relationship between mtDNA replication fidelity and cellular homeostasis of high energy demand tissues has for example been discussed in relation to the POLG mutator mouse model (a mouse model expressing a proofreading-deficient POLG1 enzyme) and its accumulation of mtDNA mutations upon cardiac progenitor cell differentiation (Orogo et al., 2015). In this study, the mutator mouse exhibits the following phenotypes upon cellular differentiation: reduced proliferation, increased reliance on glycolysis instead of oxidative phosphorylation, along with fragmented perinuclear mitochondrial clusters compared to wild-type controls. These reported mtDNA-dependent phenotypes are identical to the phenotypes we observe upon manipulation of normal MacroD1 function in our MacroD1 knock-out cell lines. Additionally, we also observed that knockdown of the *Drosophila melanogaster* mitochondrial MacroD1-like macrodomain lead to a flight inability phenotype and a shift in mitochondrial morphology – flight muscle mitochondria appeared thinner and smaller in the

MacroD1-like knockdown muscle. These functional observations reinforce the importance of the function of MacroD1 and ADP-ribosylation in regulating mtDNA replication.

Additional support for the requirement of a functional MacroD1 enzyme in mtDNA replication came with the observation that it was impossible to obtain stable cell clones over-expressing an ADPr binding- deficient MacroD1 mutant in the MacroD1 knockout cells. The over-expression in the wild-type U2-OS background was viable – but with a dominant-negative effect. Reminiscent of the mtDNA phenotypes observed upon overexpression of catalytic-dead POLG1 (Spelbrink et al., 2000), the catalytic subunit of the replicative mitochondrial DNA polymerase. Seminal research with human POLG in mammalian cell culture has shown that it is not possible to select cell clones stably expressing polymerase-deficient POLG mutant proteins, which implies negative effects of this mutant on mtDNA replication and/or cell viability (Spelbrink et al., 2000). This mirrors our observations with the ADP-ribosyl binding-deficient MacroD1 mutant protein, which we could not stably over-express in the MD1KO cell line and had deleterious effects on mtDNA content and cell size when stably inserted into wild-type U2-OS cells. MacroD1 and POLG1 thus appear to be functionally related with regard to their roles in maintaining mitochondrial genome homeostasis, establishing a mechanistic connection between mtDNA homeostasis and MacroD1 function.

Complementing our phenotypic findings, we find that manipulating MacroD1 leads to defects in mtDNA homeostasis. Our MIP assay shows robust changes in the occupancy of POLG1 and TWINKLE on mtDNA in MacroD1 knockout cells, which are not caused by the observed depletion of mtDNA (since we normalize to percent input), nor by a reduction of the nuclear-encoded replication and repair factors. Rather, our data suggest that ADP-ribosylation of these mtDNA replication factors is important for the association of POLG1 and TWINKLE with the replicating genome.

In recent years, the increase in mutation load over the lifespan of mtDNA has been attributed more to faulty replication mechanisms, rather than to increases in oxidative species (Kauppila et al., 2018; Zheng et al., 2006). However, by comparing the mutation load between wild-type

and MacroD1-knockout cells we were able to demonstrate that the reduction of POLG1 and TWINKLE on MD1KO mtDNA does not lead to increases in mutation load in our model. Although mitochondria contain a number of factors involved in mtDNA repair, they are believed to lack efficient double strand repair systems (Alexeyev et al., 2013). Furthermore, it has been shown that double strand breaks lead to a degradation of mtDNA by the replisome (Peeva et al., 2018). This would explain why we observe a significant depletion of mtDNA yet only a slight decrease in mtDNA integrity in the MacroD1 knock-out cell lines.

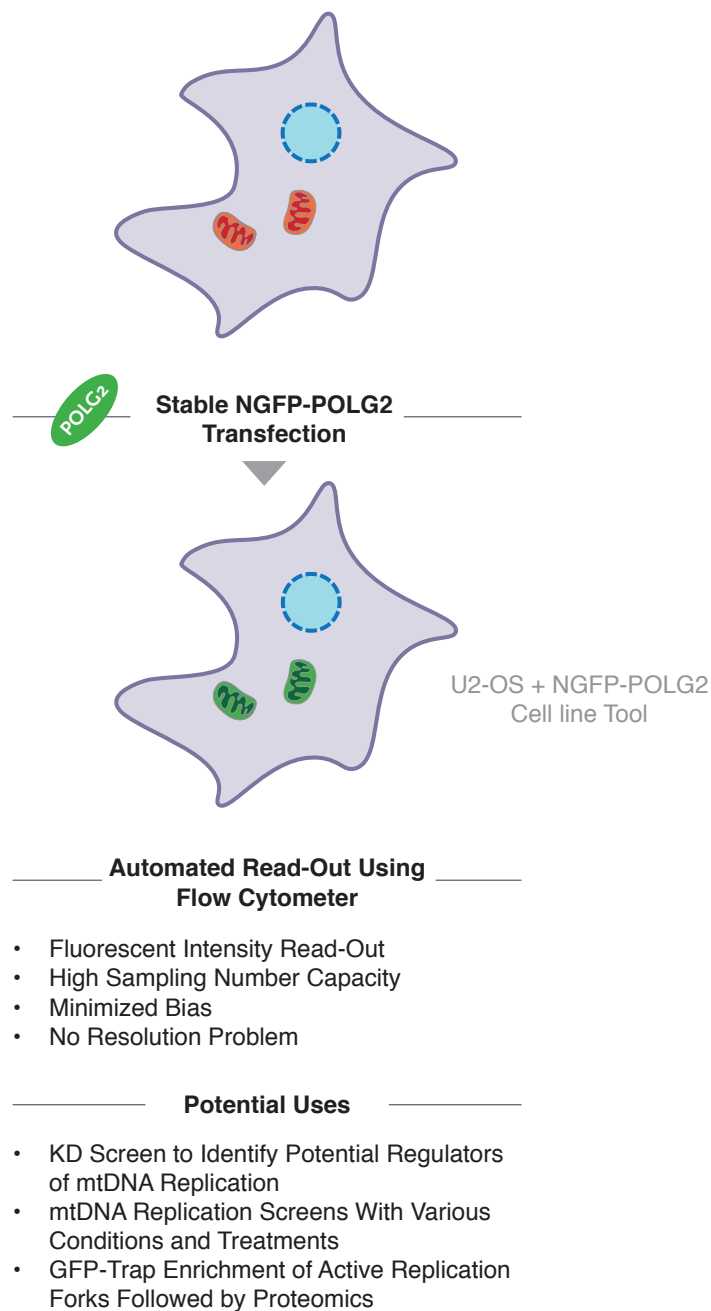
In order to further refine the role of mitochondrial ADP-ribosylation of mtDNA factors, we used our novel tool: the NGFP-POLG2 assay to investigate mtDNA replication. Using this tool, we show that there is less active mtDNA replication in the mitochondria of MD1KO+ NGFP-POLG2 cells compared to the wild-type WT+ NGFP-POLG2 control cell lines. To check whether the difference in the observed active replication forks is mirrored by a difference in *de novo* mtDNA synthesis, we validated our results using an established EdU-incorporation assay. By coupling EdU-incorporation with imaging flow cytometry we were able to determine total EdU-incorporation in mitochondria. Using this method, we observed a MacroD1-dependent decrease of *de novo* mtDNA synthesis in the MacroD1-depleted cells, which we rescued by expressing the MacroD1 gene. Finally, to quantify whether MacroD1 impacts steady-state mtDNA levels, we quantified mtDNA copy number and found that MacroD1 knockout cells have significantly less mtDNA compared to wild-type controls. While we were not able to convince ourselves that the wild-type protein rescues this phenotype, the over-expression of the ADPr binding-deficient MacroD1 mutant in wild-type cells greatly reduces mtDNA content. This strongly implicates the normal ADP-ribosyl-binding function of MacroD1 in regulating the copy number of mtDNA, revealing a novel role for mitochondrial ADP-ribosylation. We further hypothesize that the ADPr binding-deficient MacroD1 protein's dominant-negative phenotype may reflect competition with the endogenous protein, with the functional consequence that a mechanism essential for mtDNA replication is severely disrupted. In summary, the data within this thesis shows that MacroD1 function, regulates mtDNA replication.

## 5. Summary of findings

In the two parts of this thesis I i) conceptualized, generated, verified and utilized a novel tool, the POLG2 Assay to analyse mtDNA replication *in vivo* (subchapters 3.1, 4.1 and Figure 5.1) and ii) demonstrated the importance of the ADPr-binding, macrodomain protein, MacroD1 in the regulation of mitochondrial DNA replication and cellular homeostasis (subchapters 3.2, 3.3, 4.2, 4.3 and Figure 5.1).

The POLG2 Assay enables us to quantify active replication forks normalized to overall mitochondrial content, over tens of thousands of cells per biological repeat, in a semi-automated approach. Previously, this type of experiment was carried out using a confocal microscope followed by *foci* counting. Thus, with this assay there is not only a reduction in the amount of time it takes to perform the relative investigation of mitochondrial replication rate, but also increased reproducibility and accuracy, by increasing sampling size and decreasing experimental bias (see Figure 5.1), future applications of the POLG2 Assay are discussed in subchapter 6.1.

In the second and third part of this thesis we showed that MacroD1 interacts with core mtDNA replication factors POLG1, TWINKLE, LIG3, TOP1 and TOP2A in an ADPr-binding dependent manner. Further, we showed that POLG1, TWINKLE, LIG3 and TFAM are ADP-ribosylated. In *in vivo* studies we demonstrated that deletion of MacroD1 as well as overexpression of an ADPr-binding deficient MacroD1 mutant result in dysregulated mtDNA replication. Thereby, establishing an important homeostatic and potentially regulatory role for mitochondrial ADP-ribosylation and DNA replication mediated by the mitochondrial matrix macrodomain protein MacroD1. Our data shows that the ADP-ribosylation of core mtDNA factors and the subsequent interaction with MacroD1 affect the maintenance and regulation of the mitochondrial genome, which, in turn, impacts cellular homeostasis and metabolism. Thereby, revealing an essential macrodomain function in mitochondria, adding to the understanding of ADP-ribosylation as a biologically important PTM.



**Figure 5.1: The POLG2, mtDNA Replication Assay.** The POLG2 cell line tool consists of monoclonalized U2-OS cell lines stably expressing NGFP-POLG2. A fully functional, fluorescent, POLG2 mutant. NGFP-POLG2 colocalizes with active replication forks but does not alter bioenergetic profiles or mtDNA copy number. NGFP-POLG2 cell lines can be used to assay relative mtDNA replication changes (further explained in text).

## 6. Future work

### 6.1 Future applications for the POLG2-assay

With the POLG2 fluorescence-based assay to monitor mitochondrial DNA replication *in vivo*, I established and tested a novel method which can be used to monitor mtDNA replication changes, along with cell size, mitochondrial content and expression levels on a great number of cells/samples *in vivo* (subchapter 3.1). The POLG2 assay might also be utilized to screen for as of yet unknown factors involved in the regulation of mtDNA (see Figure 5.1). We propose a combination of knock-down screens along with transient overexpression studies in the NGFP-POLG2 U2-OS cells. In this manner it should be possible to determine factors involved in the regulation of mtDNA on a larger scale.

Using a CRISPR/Cas9-based knock-out screen (Joung et al., 2017; Shalem et al., 2014) it might be possible to identify potential regulators of mtDNA replication by targeting proteins on a genome-wide scale in the U2-OS + NGFP-POLG2 tool, followed by comparative green-fluorescence read-out. However, we believe that for our specific goal there would be a couple of drawbacks. During the weeks it would take to generate the full knock-out library, small clonal changes might take place within the cell populations, leading to phenotypic divergence not directly coupled to mtDNA replication regulation. Additionally, the U2-OS + NGFP-POLG2 cells should be maintained on their G418 selection media, while also requiring the antibiotic used in knock-out selection, to ensure the NGFP-POLG2 expression isn't lost during the process.

We propose to employ a targeted approach to knock-down mitochondrial proteins, as well as proteins with predicted mitochondrial localization, specifically. The first step would be to amass a library of likely candidates by employing databases such as MitoMiner v4.0 and MitoCarta (see data links in appendix 9), these lists could be supplemented by lists generated using the complete human proteome and screening for canonical MTS-sequences (described under MTS predictions in appendix 9). Compiled these lists are expected to encompass

approximately 1400 proteins (Calvo and Mootha, 2010; Morgenstern et al., 2017; Vogtle et al., 2017) from which one might want to subtract proteins that are well characterized as functioning outside the mitochondrial matrix (Vogtle et al., 2017). Based on these reduced lists it would be possible to perform knock-down using transient transfection of siRNAs. To perform an initial screen per knock-down, cells grown in one well of a 9-well dish would be sufficient for both – the fluorescent read-out of MitoTracker, NGFP-POLG2 and forward scatter as well as, if necessary, subsequent pPCR and/or western blot analysis. Promising hits from this screening might be followed up by knock-out and over-expression studies.

Recently, a very similar POLG2-based approach has been hypothesized as a possibility by David Prole, Patrick Chinnery and Nick Jones (Prole et al., 2020). In their review the authors propose that the fluorescently tagged POLG2 might be used to specifically label replicating mtDNA within cells, while simultaneous labelling of all mtDNA nucleoids could reveal subcellular heterogeneity in the replicative status of mtDNA. In our POLG2 mtDNA replication assay we have done exactly that, although instead of normalizing to mtDNA foci we assess total mitochondrial content. Additionally, we coupled the read-out to a flow-cytometer instead of a confocal microscope and foci counting, thereby increasing sample size and minimising sampling bias.

## 6.2 What are the effects of stable overexpression of mtDNA replication factors?

During the work for this thesis, we established a number of monoclonalised U2-OS cell lines, stably expressing fluorophore-tagged mtDNA replication factors (section 3.2.5). These cell lines could be used to investigate whether long-term over-expression of POLG1-GFP, LIG3-GFP and TWINKLE-GFP has effects on mtDNA copy number, mitochondrial content, cell size and proliferation rate compared to cell lines overexpressing mtGFP. Do any of these factors clearly increase mtDNA copy number or cell fitness across multiple monoclonalised cell lines?



Further, it would be possible to use the GFP-tagged mtDNA replication factors to perform GFP-trap immunoprecipitation followed by mass-spectrometric analysis (also see subchapter 6.5) and western-blot analysis.

### 6.3 Which stage of mtDNA replication is affected by ADP-ribosylation?

By immunoprecipitating either tagged or endogenous mtDNA replication factors followed by mass-spectrometric analysis in both the U2-OS WT *versus* MD1KO cell lines, we would be able to address some of our remaining open questions concerning for example the stage of mtDNA replication affected by MacroD1 interaction and ADP-ribosylation. For example, does the absence of MacroD1 in U2-OS cells not only affect the interaction between mtDNA replication factors and mtDNA but also interactions between different mtDNA replication factors in the mtDNA replisome, or with mtRNA and factors required for transcription?

### 6.4 Which enzymes MAR- and /or PARYlate mitochondrial proteins?

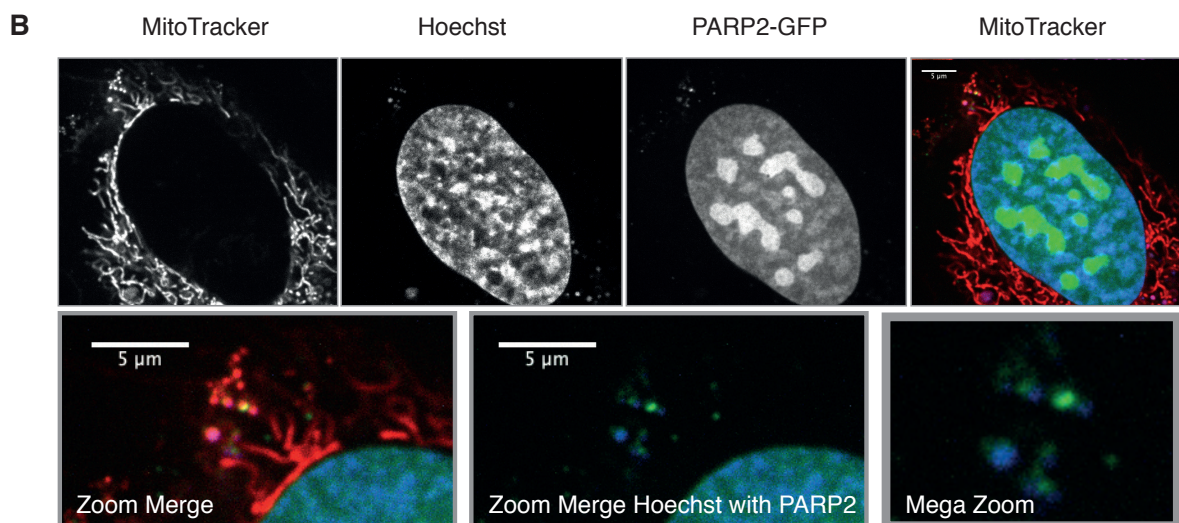
By demonstrating that the macrodomain protein, MacroD1, is localized exclusively to the mitochondrial matrix and by identifying MacroD1's interaction partners and function, in not only binding ADP-ribosylated target proteins, but also in protecting the ADPr-modification, we have identified MacroD1 as a powerful interaction partner of mitochondrial ADP-ribosylated proteins. However, as discussed in the introduction (section 1.2.3) it is as yet unclear which mitochondrial enzyme or enzymes perform mitochondrial ADP-ribosylation. So far, two ADP-ribosyltransferases have been identified in mitochondria SIRT4 and PARP1 (Haigis et al., 2006; Rossi et al., 2009), with the mitochondrial localization of PARP1 being highly contentious.

Mitochondrial localization of a specific protein is routinely tested by: i) analysing the amino acid sequence for the presence of a canonical MTS, ii) fusing the protein in question to a fluorophore and analysing its subcellular localization, or iii) using specific antibodies, iv)

determining subcellular and submitochondrial localization by subcellular fractionation, proteinase K treatment and western blot analysis.

**A**

Gene name	Alt. name	Activity	MitoProt II	TargetP	iPSORT
PARP1	ARTD1	poly-			
PARP2	ARTD2	poly-	yes	yes	yes
PARP5a	Tankyrase1/ARTD5	poly-		yes	
PARP16	ARTD15	mono-	yes		
SIRT4			yes	yes	yes



**Table 6.1: Which Enzyme MAR- and /or PARylate Mitochondrial Proteins?** A) ADP-Ribosyltransferase MTS prediction. Some PARPs are able to catalyze MARYlation, some PARylation. The Sirtuins SIRT4 and SIRT6 have also been shown to possess MARYlation capacity. Which of the identified PARPs and SIRTs have predicted mitochondrial targeting sequences according to three different prediction algorithms. B) Some PARP2 Colocalizes with Mitochondria. U2-OS Cell Line Transiently Expressing PARP2-GFP. In the merged image mitochondria are red, DNA is in blue, PARP2-GFP is green. The zoomed in area depicts PARP2 and mtDNA nucleoids in close proximity. Further colocalization studies would require further optimization and super resolution microscopy.

To determine whether any of the known ADP-ribosyltransferases (table 1.1) have canonical MTS sequences, I performed sequence analysis using three mitochondrial targeting prediction software tools: MitoProtII, TargetP and iPSORT (Appendix 9: Data Links). Based on sequence analysis PARP2, PARP5a, PARP16 and the already published SIRT4 have putative MTS regions. The only ADP-ribosyltransferase predicted to have an MTS by all three

prediction software tools is PARP2 (see table 6.1 and Appendix 9). To further determine whether PARP2 might localize to mitochondria, as well as to the nucleus, I transiently transfected U2-OS cells with a PARP2-isf1-GFP construct (kind gift from Charlotte Blessing) and co-stained using Hoechst and MitoTracker to visualize both DNA and mitochondria (Figure 6.1).

Based on these preliminary results, it does appear like a subfraction of PARP2 might localize to mitochondria. So far, the submitochondrial localization and function of PARP2 therein are unclear. To further verify and analyse it would be necessary to perform submitochondrial fractionation and IF experiments of endogenous PARP2 using PARP2-specific antibodies. Additionally, it would be interesting to perform any further IF experiments by super resolution microscopy. For, although it appears like there is PARP2-signal in the close vicinity of mtDNA nucleoid staining (Figure 6.1 B), the images are not clearly resolved.

## 6.5 Which residues are targeted by mitochondrial ADP-ribosylation?

We have previously stipulated that both PARP1 and/or PARP2 might have mitochondrially located subfractions. Further, we discussed that in a nuclear context both, PARP1 and PARP2, require interaction with HPF1 to shift from Glu/Asp-residues to Serine-residues. Alas, there are no reports of HPF1 in mitochondria, additionally we did not identify an MTS using the known software tools available. The question remains: which residues are targeted by mitochondrial ADP-ribosylation?

We are currently addressing this question, as well as the questions posed in sections 6.3 and 6.4, in collaboration with Sara Buch-Larsen, Ivo Alexander Hendriks and Michael Lund Nielsen from the University of Copenhagen and Novo Nordisk Foundation Center for Protein Research (Hendriks et al., 2019; Hendriks et al., 2018; Larsen et al., 2018; Larsen et al., 2017; Martello et al., 2016).

In this collaboration we are immuno-precipitating endogenous POLG1, TWINKLE and Tfam from both U2-OS WT and MD1KO cell lines, as well as over-expressed, Flag-tagged POLG2 from HekTRex WT cells and preparing them for mass-spec by on-bead Trypsin-digestion. The samples are then sent to Copenhagen where they will be reduced/alkylated and StageTipped for mass spectrometric analysis.

Using this method, we aim to identify which specific residues are modified by ADP-ribosylation on these factors, the ratio between modified and unmodified protein populations, as well as the ratio of modified and unmodified proteins comparing the wild-type versus the MD1KO samples. Furthermore, by following this path, we might also be able to delve deeper into the mechanistic understanding of how exactly ADP-ribosylation alters mtDNA replication, beyond the interaction between mtDNA and mtDNA replication factors.

## References

- Agnew, T., Munnur, D., Crawford, K., Palazzo, L., Mikoc, A., and Ahel, I. (2018). MacroD1 is a promiscuous ADP-ribosyl hydrolase localized to mitochondria. *Front Microbiol* *9*, 20.
- Ahel, I., Ahel, D., Matsusaka, T., Clark, A.J., Pines, J., Boulton, S.J., and West, S.C. (2008). Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. *Nature* *451*, 81-85.
- Akbari, M., Keijzers, G., Maynard, S., Scheibye-Knudsen, M., Desler, C., Hickson, I.D., and Bohr, V.A. (2014). Overexpression of DNA ligase III in mitochondria protects cells against oxidative stress and improves mitochondrial DNA base excision repair. *DNA Repair (Amst)* *16*, 44-53.
- Alano, C.C., Tran, A., Tao, R., Ying, W., Karliner, J.S., and Swanson, R.A. (2007). Differences among cell types in NAD(+) compartmentalization: a comparison of neurons, astrocytes, and cardiac myocytes. *J Neurosci Res* *85*, 3378-3385.
- Alexeyev, M., Shokolenko, I., Wilson, G., and LeDoux, S. (2013). The maintenance of mitochondrial DNA integrity--critical analysis and update. *Cold Spring Harb Perspect Biol* *5*, a012641.
- Altmeyer, M., Messner, S., Hassa, P.O., Fey, M., and Hottiger, M.O. (2009). Molecular mechanism of poly(ADP-ribosyl)ation by PARP1 and identification of lysine residues as ADP-ribose acceptor sites. *Nucleic acids research* *37*, 3723-3738.
- Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., *et al.* (1981). Sequence and organization of the human mitochondrial genome. *Nature* *290*, 457-465.
- Andersson, S.G., Karlberg, O., Canback, B., and Kurland, C.G. (2003). On the origin of mitochondria: a genomics perspective. *Philos Trans R Soc Lond B Biol Sci* *358*, 165-177; discussion 177-169.
- Askeland, G., Dosoudilova, Z., Rodinova, M., Klempir, J., Liskova, I., Kusnierczyk, A., Bjoras, M., Nesse, G., Klungland, A., Hansikova, H., *et al.* (2018). Increased nuclear DNA damage precedes mitochondrial dysfunction in peripheral blood mononuclear cells from Huntington's disease patients. *Scientific reports* *8*, 9817.
- Bacman, S.R., Kauppila, J.H.K., Pereira, C.V., Nissanka, N., Miranda, M., Pinto, M., Williams, S.L., Larsson, N.G., Stewart, J.B., and Moraes, C.T. (2018). MitoTALEN reduces mutant mtDNA load and restores tRNA(Ala) levels in a mouse model of heteroplasmic mtDNA mutation. *Nature medicine* *24*, 1696-1700.
- Bannai, H., Tamada, Y., Maruyama, O., Nakai, K., and Miyano, S. (2002). Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* *18*, 298-305.
- Barkauskaite, E., Jankevicius, G., Ladurner, A.G., Ahel, I., and Timinszky, G. (2013). The recognition and removal of cellular poly(ADP-ribose) signals. *FEBS J* *280*, 3491-3507.
- Bertolin, G., Bulteau, A.L., Alves-Guerra, M.C., Burel, A., Lavault, M.T., Gavard, O., Le Bras, S., Gagne, J.P., Poirier, G.G., Le Borgne, R., *et al.* (2018). Aurora kinase A localises to mitochondria to control organelle dynamics and energy production. *eLife* *7*.

- Bonekamp, N.A., and Larsson, N.G. (2018). SnapShot: Mitochondrial Nucleoid. *Cell* *172*, 388-388 e381.
- Bonfiglio, J.J., Colby, T., and Matic, I. (2017a). Mass spectrometry for serine ADP-ribosylation? Think o-glycosylation! *Nucleic acids research* *45*, 6259-6264.
- Bonfiglio, J.J., Fontana, P., Zhang, Q., Colby, T., Gibbs-Seymour, I., Atanassov, I., Bartlett, E., Zaja, R., Ahel, I., and Matic, I. (2017b). Serine ADP-Ribosylation Depends on HPF1. *Mol Cell* *65*, 932-940 e936.
- Boynton, S., and Tully, T. (1992). *latheo*, a new gene involved in associative learning and memory in *Drosophila melanogaster*, identified from P element mutagenesis. *Genetics* *131*, 655-672.
- Brown, T.A., and Clayton, D.A. (2002). Release of replication termination controls mitochondrial DNA copy number after depletion with 2',3'-dideoxycytidine. *Nucleic Acids Res* *30*, 2004-2010.
- Calvo, S.E., Clauser, K.R., and Mootha, V.K. (2016). MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic acids research* *44*, D1251-1257.
- Calvo, S.E., and Mootha, V.K. (2010). The mitochondrial proteome and human disease. *Annu Rev Genomics Hum Genet* *11*, 25-44.
- Cervantes-Laurean, D., Loflin, P.T., Minter, D.E., Jacobson, E.L., and Jacobson, M.K. (1995). Protein modification by ADP-ribose via acid-labile linkages. *The Journal of biological chemistry* *270*, 7929-7936.
- Chambon, P., Weill, J.D., and Mandel, P. (1963). Nicotinamide mononucleotide activation of new DNA-dependent polyadenylic acid synthesizing nuclear enzyme. *Biochem Biophys Res Commun* *11*, 39-43.
- Chen, D., Vollmar, M., Rossi, M.N., Phillips, C., Kraehenbuehl, R., Slade, D., Mehrotra, P.V., von Delft, F., Crosthwaite, S.K., Gileadi, O., *et al.* (2011). Identification of macrodomain proteins as novel O-acetyl-ADP-ribose deacetylases. *J Biol Chem* *286*, 13261-13271.
- Ciafaloni, E., Ricci, E., Shanske, S., Moraes, C.T., Silvestri, G., Hirano, M., Simonetti, S., Angelini, C., Donati, M.A., Garcia, C., *et al.* (1992). MELAS: clinical features, biochemistry, and molecular genetics. *Ann Neurol* *31*, 391-398.
- Claros, M.G., and Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem* *241*, 779-786.
- Copeland, W.C. (2010). The mitochondrial DNA polymerase in health and disease. *Subcell Biochem* *50*, 211-222.
- Copeland, W.C. (2012). Defects in mitochondrial DNA replication and human disease. *Crit Rev Biochem Mol Biol* *47*, 64-74.
- Cox, J., Hein, M.Y., Lubner, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* *13*, 2513-2526.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* *26*, 1367-1372.

- Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 10, 1794-1805.
- Coxhead, J., Kurzawa-Akanbi, M., Hussain, R., Pyle, A., Chinnery, P., and Hudson, G. (2016). Somatic mtDNA variation is an important component of Parkinson's disease. *Neurobiol Aging* 38, 217 e211-217 e216.
- D'Amours, D., Desnoyers, S., D'Silva, I., and Poirier, G.G. (1999). Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem J* 342 ( Pt 2), 249-268.
- Dalla Rosa, I., Huang, S.Y., Agama, K., Khiati, S., Zhang, H., and Pommier, Y. (2014). Mapping topoisomerase sites in mitochondrial DNA with a poisonous mitochondrial topoisomerase I (Top1mt). *The Journal of biological chemistry* 289, 18595-18602.
- Daniels, C.M., Thirawatananond, P., Ong, S.E., Gabelli, S.B., and Leung, A.K. (2015). Nudix hydrolases degrade protein-conjugated ADP-ribose. *Sci Rep* 5, 18271.
- Dinardo, M.M., Musicco, C., Fracasso, F., Milella, F., Gadaleta, M.N., Gadaleta, G., and Cantatore, P. (2003). Acetylation and level of mitochondrial transcription factor A in several organs of young and old rats. *Biochem Biophys Res Commun* 301, 187-191.
- Dolle, C., Flones, I., Nido, G.S., Miletic, H., Osuagwu, N., Kristoffersen, S., Lilleng, P.K., Larsen, J.P., Tysnes, O.B., Haugarvoll, K., *et al.* (2016). Defective mitochondrial DNA homeostasis in the substantia nigra in Parkinson disease. *Nat Commun* 7, 13548.
- Eisner, V., Picard, M., and Hajnoczky, G. (2018). Mitochondrial dynamics in adaptive and maladaptive cellular stress responses. *Nature cell biology* 20, 755-765.
- Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of molecular biology* 300, 1005-1016.
- Falkenberg, M. (2018). Mitochondrial DNA replication in mammalian cells: overview of the pathway. *Essays Biochem* 62, 287-296.
- Falkenberg, M., Gaspari, M., Rantanen, A., Trifunovic, A., Larsson, N.G., and Gustafsson, C.M. (2002). Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nature genetics* 31, 289-294.
- Falkenberg, M., and Gustafsson, C.M. (2020). Mammalian mitochondrial DNA replication and mechanisms of deletion formation. *Crit Rev Biochem Mol Biol*, 1-16.
- Fan, L., Kim, S., Farr, C.L., Schaefer, K.T., Randolph, K.M., Tainer, J.A., and Kaguni, L.S. (2006). A novel processive mechanism for DNA synthesis revealed by structure, modeling and mutagenesis of the accessory subunit of human mitochondrial DNA polymerase. *Mol Biol* 358, 1229-1243.
- Feijs, K.L.H., Cooper, C.D.O., and Zaja, R. (2020). The Controversial Roles of ADP-Ribosyl Hydrolases MACROD1, MACROD2 and TARG1 in Carcinogenesis. *Cancers (Basel)* 12.
- Fontana, P., Bonfiglio, J.J., Palazzo, L., Bartlett, E., Matic, I., and Ahel, I. (2017). Serine ADP-ribosylation reversal by the hydrolase ARH3. *eLife* 2017 6, 6:e28533
- Friedman, J.R., and Nunnari, J. (2014). Mitochondrial form and function. *Nature* 505, 335-343.

Fuste, J.M., Wanrooij, S., Jemt, E., Granycome, C.E., Cluett, T.J., Shi, Y., Atanassova, N., Holt, I.J., Gustafsson, C.M., and Falkenberg, M. (2010). Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. *Mol Cell* 37, 67-78.

Gagne, J.P., Isabelle, M., Lo, K.S., Bourassa, S., Hendzel, M.J., Dawson, V.L., Dawson, T.M., and Poirier, G.G. (2008). Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes. *Nucleic Acids Res* 36, 6959-6976.

Gagne, J.P., Pic, E., Isabelle, M., Krietsch, J., Ethier, C., Paquet, E., Kelly, I., Boutin, M., Moon, K.M., Foster, L.J., *et al.* (2012). Quantitative proteomics profiling of the poly(ADP-ribose)-related response to genotoxic stress. *Nucleic Acids Res* 40, 7788-7805.

Gammage, P.A., and Frezza, C. (2019). Mitochondrial DNA: the overlooked oncogenome? *BMC Biol* 17, 53.

Gammage, P.A., Viscomi, C., Simard, M.L., Costa, A.S.H., Gaude, E., Powell, C.A., Van Haute, L., McCann, B.J., Rebelo-Guiomar, P., Cerutti, R., *et al.* (2018). Genome editing in mitochondria corrects a pathogenic mtDNA mutation in vivo. *Nature medicine* 24, 1691-1695.

Ge, S.X., Jung, D., and Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36, 2628-2629.

Gibson, B.A., Conrad, L.B., Huang, D., and Kraus, W.L. (2017). Generation and Characterization of Recombinant Antibody-Like ADP-Ribose Binding Proteins. *Biochemistry*.

Göke, A., Schrott, S., Mizrak, A., Belyy, V., Osman, C., and, and Walter, P. (2019). Mrx6 regulates mitochondrial DNA copy number in *S. cerevisiae* by engaging the evolutionarily conserved Lon protease Pim1. *Mol Biol Cell*.

Graham, F.L., Smiley, J., Russell, W.C., and Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol* 36, 59-74.

Gray, M.W. (1992). The Endosymbiont Hypothesis Revisited. In, pp. 233-357.

Greenfield, A., Braude, P., Flinter, F., Lovell-Badge, R., Ogilvie, C., and Perry, A.C.F. (2017). Assisted reproductive technologies to prevent human mitochondrial disease transmission. *Nat Biotechnol* 35, 1059-1068.

Gustafsson, C.M., Falkenberg, M., and Larsson, N.G. (2016). Maintenance and expression of mammalian mitochondrial DNA. *Annu Rev Biochem* 85, 133-160.

Haigis, M.C., Mostoslavsky, R., Haigis, K.M., Fahie, K., Christodoulou, D.C., Murphy, A.J., Valenzuela, D.M., Yancopoulos, G.D., Karow, M., Blander, G., *et al.* (2006). SIRT4 inhibits glutamate dehydrogenase and opposes the effects of calorie restriction in pancreatic beta cells. *Cell* 126, 941-954.

Haines, K.M., Feldman, E.L., and Lentz, S.I. (2010). Visualization of mitochondrial DNA replication in individual cells by EdU signal amplification. *J Vis Exp*, e2147.

Hance, N., Ekstrand, M.I., and Trifunovic, A. (2005). Mitochondrial DNA polymerase gamma is essential for mammalian embryogenesis. *Hum Mol Genet* 14, 1775-1783.

Heddi, A., Lestienne, P., Wallace, D.C., and Stepien, G. (1993). Mitochondrial DNA expression in mitochondrial myopathies and coordinated expression of nuclear genes involved in ATP production. *The Journal of biological chemistry* 268, 12156-12163.



- Hendriks, I.A., Larsen, S.C., and Nielsen, M.L. (2019). An Advanced Strategy for Comprehensive Profiling of ADP-ribosylation Sites Using Mass Spectrometry-based Proteomics. *Mol Cell Proteomics* 18, 1010-1026.
- Hendriks, I.A., Lyon, D., Su, D., Skotte, N.H., Daniel, J.A., Jensen, L.J., and Nielsen, M.L. (2018). Site-specific characterization of endogenous SUMOylation across species and organs. *Nature communications* 9, 2456.
- Herrero-Yraola, A., Bakhit, S.M., Franke, P., Weise, C., Schweiger, M., Jorcke, D., and Ziegler, M. (2001). Regulation of glutamate dehydrogenase by reversible ADP-ribosylation in mitochondria. *EMBO J* 20, 2404-2412.
- Hillen, H.S., Morozov, Y.I., Sarfallah, A., Temiakov, D., and Cramer, P. (2017). Structural Basis of Mitochondrial Transcription Initiation. *Cell* 171, 1072-1081.e1010.
- Hoekstra, J.G., Hipp, M.J., Montine, T.J., and Kennedy, S.R. (2016). Mitochondrial DNA mutations increase in early stage Alzheimer disease and are inconsistent with oxidative damage. *Ann Neurol* 80, 301-306.
- Holt, I.J., and Reyes, A. (2012). Human mitochondrial DNA replication. *Cold Spring Harb Perspect Biol* 4.
- Hornbeck, P.V., Chabra, I., Kornhauser, J.M., Skrzypek, E., and Zhang, B. (2004). PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics* 4, 1551-1561.
- Hudson, G., and Chinnery, P.F. (2006). Mitochondrial DNA polymerase-gamma and human disease. *Am J Hum Genet* 15 Spec No 2, R244-252.
- Hudson, G., Takeda, Y., and Herbert, M. (2019). Reversion after replacement of mitochondrial DNA. *Nature* 574, E8-E11.
- Humble, M.M., Young, M.J., Foley, J.F., Pandiri, A.R., Travlos, G.S., and Copeland, W.C. (2013). Polg2 is essential for mammalian embryogenesis and is required for mtDNA maintenance. *Hum Mol Genet* 22, 1017-1025.
- Ikeda, M., Ide, T., Fujino, T., Arai, S., Saku, K., Kakino, T., Tynnismaa, H., Yamasaki, T., Yamada, K., Kang, D., *et al.* (2015). Overexpression of TFAM or twinkle increases mtDNA copy number and facilitates cardioprotection associated with limited mitochondrial oxidative stress. *PLoS one* 10, e0119687.
- Ikeda, T., Osaka, H., Shimbo, H., Tajika, M., Yamazaki, M., Ueda, A., Murayama, K., and Yamagata, T. (2018). Mitochondrial DNA 3243A>T mutation in a patient with MELAS syndrome. *Hum Genome Var* 5, 25.
- Isaac, R.S., McShane, E., and Churchman, L.S. (2018). The Multiple Levels of Mitonuclear Coregulation. *Annual review of genetics* 52, 511-533.
- Jankevicius, G., Hassler, M., Golia, B., Rybin, V., Zacharias, M., Timinszky, G., and Ladurner, A.G. (2013). A family of macrodomain proteins reverses cellular mono-ADP-ribosylation. *Nat Struct Mol Biol* 20, 508-514.
- Jazwinski, S.M. (2013). The retrograde response: when mitochondrial quality control is not enough. *Biochimica et biophysica acta* 1833, 400-409.

- Joung, J., Konermann, S., Gootenberg, J.S., Abudayyeh, O.O., Platt, R.J., Brigham, M.D., Sanjana, N.E., and Zhang, F. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nature protocols* 12, 828-863.
- Kanki, T., Ohgaki, K., Gaspari, M., Gustafsson, C.M., Fukuoh, A., Sasaki, N., Hamasaki, N., and Kang, D. (2004). Architectural role of mitochondrial transcription factor A in maintenance of human mitochondrial DNA. *Mol Cell Biol* 24, 9823-9834.
- Karras, G.I., Kustatscher, G., Buhecha, H.R., Allen, M.D., Pugieux, C., Sait, F., Bycroft, M., and Ladurner, A.G. (2005). The macro domain is an ADP-ribose binding module. *EMBO* 24, 1911-1920.
- Kasiviswanathan, R., Collins, T.R., and Copeland, W.C. (2012). The interface of transcription and DNA replication in the mitochondria. *Biochimica et biophysica acta* 1819, 970-978.
- Kauppila, J.H.K., Baines, H.L., Bratic, A., Simard, M.L., Freyer, C., Mourier, A., Stamp, C., Filograna, R., Larsson, N.G., Greaves, L.C., *et al.* (2016). A Phenotype-Driven Approach to Generate Mouse Models with Pathogenic mtDNA Mutations Causing Mitochondrial Disease. *Cell reports* 16, 2980-2990.
- Kauppila, J.H.K., Bonekamp, N.A., Mourier, A., Isokallio, M.A., Just, A., Kauppila, T.E.S., Stewart, J.B., and Larsson, N.G. (2018). Base-excision repair deficiency alone or combined with increased oxidative stress does not increase mtDNA point mutations in mice. *Nucleic Acids Res* 46, 6642-6669.
- Khadka, P., Hsu, J.K., Veith, S., Tadokoro, T., Shamanna, R.A., Mangerich, A., Croteau, D.L., and Bohr, V.A. (2015). Differential and Concordant Roles for Poly(ADP-Ribose) Polymerase 1 and Poly(ADP-Ribose) in Regulating WRN and RECQL5 Activities. *Mol Cell Biol* 35, 3974-3989.
- King, G.A., Hashemi Shabestari, M., Taris, K.H., Pandey, A.K., Venkatesh, S., Thilagavathi, J., Singh, K., Krishna Koppiseti, R., Temiakov, D., Roos, W.H., *et al.* (2018). Acetylation and phosphorylation of human TFAM regulate TFAM-DNA interactions via contrasting mechanisms. *Nucleic Acids Res* 46, 3633-3642.
- Kistemaker, H.A., Nardoza, A.P., Overkleeft, H.S., van der Marel, G.A., Ladurner, A.G., and Filippov, D.V. (2016). Synthesis and macrodomain binding of mono-ADP-ribosylated peptides. *Angew Chem Int Ed Engl* 55, 10634-10638.
- Koc, E.C., Cimen, H., Kumcuoglu, B., Abu, N., Akpınar, G., Haque, M.E., Spremulli, L.L., and Koc, H. (2013). Identification and characterization of CHCHD1, AURKAIP1, and CRIF1 as new members of the mammalian mitochondrial ribosome. *Front Physiol* 4, 183.
- Korhonen, J.A., Gaspari, M., and Falkenberg, M. (2003). TWINKLE Has 5' → 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein. *J Biol Chem* 278, 48627-48632.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.
- Krasich, R., and Copeland, W. C. (2017). DNA polymerases in the mitochondria: A critical review of the evidence. *Front Biosci* 22, 692 - 709.
- Kraus, W.L. (2008). Transcriptional control by PARP-1: chromatin modulation, enhancer-binding, coregulation, and insulation. *Current opinion in cell biology* 20, 294-302.

- Kraus, W.L. (2009). New functions for an ancient domain. *Nature structural & molecular biology* *16*, 904-907.
- Kraus, W.L. (2015). PARPs and ADP-Ribosylation: 50 Years ... and Counting. *Mol Cell* *58*, 902-910.
- Krishnakumar, R., and Kraus, W.L. (2010). The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. *Mol Cell* *39*, 8-24.
- Kuhl, I., Miranda, M., Posse, V., Milenkovic, D., Mourier, A., Siira, S.J., Bonekamp, N.A., Neumann, U., Filipovska, A., Polosa, P.L., *et al.* (2016). POLRMT regulates the switch between replication primer formation and gene expression of mammalian mtDNA. *Sci Adv* *2*, e1600963.
- Kukat, C., Davies, K.M., Wurm, C.A., Spahr, H., Bonekamp, N.A., Kuhl, I., Joos, F., Polosa, P.L., Park, C.B., Posse, V., *et al.* (2015). Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proceedings of the National Academy of Sciences of the United States of America* *112*, 11288-11293.
- Kukat, C., Wurm, C.A., Spahr, H., Falkenberg, M., Larsson, N.G., and Jakobs, S. (2011a). Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proceedings of the National Academy of Sciences of the United States of America* *108*, 13534-13539.
- Kukat, C., Wurm, C.A., Spåhr, H., and Falkenberg, M., Larsson, N.-G., and Jakobs, S. (2011b). Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *PNAS* *108*, 13534–13539.
- Kun, E., Zimber, P.H., Chang, A.C.Y., Puschendorf, B., and Grunicke, H. (1975). Macromolecular Enzymatic Product of NAD<sup>+</sup> in Liver Mitochondria. *Proc Nat Acad Sci U S A* *72*.
- Lai, Y., Chen, Y., Watkins, S.C., Nathaniel, P.D., Guo, F., Kochanek, P.M., Jenkins, L.W., Szabo, C., and Clark, R.S. (2008). Identification of poly-ADP-ribosylated mitochondrial proteins after traumatic brain injury. *J Neurochem* *104*, 1700-1711.
- Laing, S., Koch-Nolte, F., Haag, F., and Buck, F. (2011). Strategies for the identification of arginine ADP-ribosylation sites. *J Proteomics* *75*, 169-176.
- Lakshmipathy, U., and Campbell, C. (1999). The human DNA ligase III gene encodes nuclear and mitochondrial proteins. *Mol Cell Biol* *19*, 3869-3876.
- Larsen, S.C., Hendriks, I.A., Lyon, D., Jensen, L.J., and Nielsen, M.L. (2018). Systems-wide Analysis of Serine ADP-Ribosylation Reveals Widespread Occurrence and Site-Specific Overlap with Phosphorylation. *Cell reports* *24*, 2493-2505 e2494.
- Larsen, S.C., Leutert, M., Bilan, V., Martello, R., Jungmichel, S., Young, C., Hottiger, M.O., and Nielsen, M.L. (2017). Proteome-Wide Identification of In Vivo ADP-Ribose Acceptor Sites by Liquid Chromatography-Tandem Mass Spectrometry. *Methods in molecular biology* *1608*, 149-162.
- Lecrenier, N., Van Der Bruggen, P., and Foury, F. (1997). Mitochondrial DNA polymerases from yeast to man: a new family of polymerases. *Gene* *185*, 147-152.

Leidecker, O., Bonfiglio, J.J., Colby, T., Zhang, Q., Atanassov, I., Zaja, R., Palazzo, L., Stockum, A., Ahel, I., and Matic, I. (2016). Serine is a new target residue for endogenous ADP-ribosylation on histones. *Nature chemical biology* *12*, 998-1000.

Leung, A.K.L. (2017). PARPs. *Curr Biol* *27*, R1256-R1258.

Lewis, M.R., and Lewis, W.H. (1914). Mitochondria in Tissue Culture. *Science* *39*, 330-333.

Lewis, S.C., Uchiyama, L.F., and Nunnari, J. (2016). ER-mitochondria contacts couple mtDNA synthesis with mitochondrial division in human cells. *Science* *353*, aaf5549.

Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* *128*, 707-719.

Lightowers, R.N. (2011). Mitochondrial transformation: time for concerted action. *EMBO reports* *12*, 480-481.

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

Lo Re, O., and Vinciguerra, M. (2017). Histone MacroH2A1: A Chromatin Point of Intersection between Fasting, Senescence and Cellular Regeneration. *Genes (Basel)* *8*.

Lotz, C., Lin, A.J., Black, C.M., Zhang, J., Lau, E., Deng, N., Wang, Y., Zong, N.C., Choi, J.H., Xu, T., *et al.* (2014). Characterization, design, and function of the mitochondrial proteome: from organs to organisms. *J Proteome Res* *13*, 433-446.

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

Luoma, P., Melberg, A., Rinne, J.O., Kaukonen, J.A., Nupponen, N.N., Chalmers, R.M., Oldfors, A., Rautakorpi, I., Peltonen, L., Majamaa, K., *et al.* (2004). Parkinsonism, premature menopause, and mitochondrial DNA polymerase  $\gamma$  mutations: clinical and molecular genetic study. *Lancet* *364*, 875-882.

Luscher, B., Butepage, M., Ecker, L., Krieg, S., Verheugd, P., and Shilton, B.H. (2018). ADP-ribosylation, a multifaceted posttranslational modification involved in the control of cell physiology in health and disease. *Chem Rev* *118*, 1092-1136.

Manning, D.R., Fraser, B.A., Kahn, R.A., and Gilman, A.G. (1984). ADP-ribosylation of transducin by islet-activation protein. Identification of asparagine as the site of ADP-ribosylation. *The Journal of biological chemistry* *259*, 749-756.

Martello, R., Leutert, M., Jungmichel, S., Bilan, V., Larsen, S.C., Young, C., Hottiger, M.O., and Nielsen, M.L. (2016). Proteome-wide identification of the endogenous ADP-ribosylome of mammalian cells and tissue. *Nature communications* *7*, 12917.

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

McBride, H.M., Neuspiel, M., and Wasiak, S. (2006). Mitochondria: more than just a powerhouse. *Curr Biol* *16*, R551-560.

- McCulloch, V., and Shadel, G.S. (2003). Human mitochondrial transcription factor B1 interacts with the C-terminal activation region of h-mtTFA and stimulates transcription independently of its RNA methyltransferase activity. *Mol Cell Biol* 23, 5816-5824.
- McDonald, L.J., and Moss, J. (1994). Enzymatic and nonenzymatic ADP-ribosylation of cysteine. *Mol Cell Biochem* 138, 221-226.
- Mereschkowsky, C. (1910). Theorie der zwei Plasmaarten als Grundlage der Symbiogenese, einer neuen Lehre von der Entstehung der Organismen. *Biol Centralbl* 30, 353–367.
- Meyer, R.G., Meyer-Ficca, M.L., Whatcott, C.J., Jacobson, E.L., and Jacobson, M.K. (2007). Two small enzyme isoforms mediate mammalian mitochondrial poly(ADP-ribose) glycohydrolase (PARG) activity. *Experimental cell research* 313, 2920-2936.
- Miwa, M., and Sugimura, T. (1971). Splitting of the ribose-ribose linkage of poly(adenosine diphosphate-ribose) by a calf thymus extract. *The Journal of biological chemistry* 246, 6362-6364.
- Miwa, M., Tanaka, M., Matsushima, T., and Sugimura, T. (1974). Purification and properties of glycohydrolase from calf thymus splitting ribose-ribose linkages of poly(adenosine diphosphate ribose). *The Journal of biological chemistry* 249, 3475-3482.
- Mo, Z., Hu, M., Yu, F., Shao, L., Fan, K., and Jiao, S. (2018). Leukemia-related protein 16 (LRP16) promotes tumor growth and metastasis in pancreatic cancer. *Onco Targets Ther* 11, 1215-1222.
- Mok, B.Y., de Moraes, M.H., Zeng, J., Bosch, D.E., Kotrys, A.V., Raguram, A., Hsu, F., Radey, M.C., Peterson, S.B., Mootha, V.K., *et al.* (2020). A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. *Nature* 583, 631-637.
- Morgenstern, M., Stiller, S.B., Lübbert, P., Peikert, C.D., Dannenmaier, S., Drepper, F., Weill, U., Höß, P., Feuerstein, R., Gebert, M., *et al.* (2017). Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale. *Cell reports* 19, 2836-2852.
- Nass, S., and Nass, M.M. (1963). Intramitochondrial Fibers with DNA Characteristics. II. Enzymatic and Other Hydrolytic Treatments. *J Cell Biol* 19, 613-629.
- Neuvonen, M., and Ahola, T. (2009). Differential activities of cellular and viral macro domain proteins in binding of ADP-ribose metabolites. *J Mol Biol* 385, 212-225.
- Ngo, H.B., Kaiser, J.T., and Chan, D.C. (2011). The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. *Nat Struct Mol Biol* 18, 1290-1296.
- Niere, M., Kernstock, S., Koch-Nolte, F., and Ziegler, M. (2008). Functional localization of two poly(ADP-ribose)-degrading enzymes to the mitochondrial matrix. *J Mol Cell Biol* 28, 814-824.
- Niere, M., Mashimo, M., Agledal, L., Dolle, C., Kasamatsu, A., Kato, J., Moss, J., and Ziegler, M. (2012). ADP-ribosylhydrolase 3 (ARH3), not poly(ADP-ribose) glycohydrolase (PARG) isoforms, is responsible for degradation of mitochondrial matrix-associated poly(ADP-ribose). *J Biol Chem* 287, 16088-16102.
- Noguchi, S., Arakawa, T., Fukuda, S., Furuno, M., Hasegawa, A., Hori, F., Ishikawa-Kato, S., Kaida, K., Kaiho, A., Kanamori-Katayama, M., *et al.* (2017). FANTOM5 CAGE profiles of human and mouse samples. *Sci Data* 4, 170112.

Nunnari, J., and Suomalainen, A. (2012). Mitochondria: in sickness and in health. *Cell* 148, 1145-1159.

Nurminen, A., Farnum, G.A., and Kaguni, L.S. (2017). Pathogenicity in POLG syndromes: DNA polymerase gamma pathogenicity prediction server and database. *BBA Clin* 7, 147-156.

Olins, D.E., and Olins, A.L. (2003). Chromatin history: our view from the bridge. *Nature reviews Molecular cell biology* 4, 809-814.

Orogo, A.M., Gonzalez, E.R., Kubli, D.A., Baptista, I.L., Ong, S.B., Prolla, T.A., Sussman, M.A., Murphy, A.N., and Gustafsson, A.B. (2015). Accumulation of mitochondrial DNA mutations disrupts cardiac progenitor cell function and reduces survival. *J Biol Chem* 290, 22061-22075.

Palazzo, L., Leidecker, O., Prokhorova, E., Dauben, H., Matic, I., and Ahel, I. (2018). Serine is the major residue for ADP-ribosylation upon DNA damage. *eLife* 2018 7, e34334.

Palazzo, L., Thomas, B., Jemth, A.S., Colby, T., Leidecker, O., Feijs, K.L., Zaja, R., Loseva, O., Puigvert, J.C., Matic, I., *et al.* (2015). Processing of protein ADP-ribosylation by Nudix hydrolases. *Biochem J* 468, 293-301.

Park, C.B., and Larsson, N.G. (2011). Mitochondrial DNA mutations in disease and aging. *The Journal of cell biology* 193, 809-818.

Pascal, J.M., and Ellenberger, T. (2015). The rise and fall of poly(ADP-ribose): An enzymatic perspective. *DNA Repair (Amst)* 32, 10-16.

Pastukh, V., Shokolenko, I., Wang, B., Wilson, G., and Alexeyev, M. (2007). Human mitochondrial transcription factor A possesses multiple subcellular targeting signals. *The FEBS journal* 274, 6488-6499.

Pawar, T., and Eide, L. (2017). Pitfalls in mitochondrial epigenetics. *Mitochondrial DNA A DNA Mapp Seq Anal* 28, 762-768.

Peeva, V., Blei, D., Trombly, G., Corsi, S., Szukszto, M.J., Rebelo-Guiomar, P., Gammage, P.A., Kudin, A.P., Becker, C., Altmuller, J., *et al.* (2018). Linear mitochondrial DNA is rapidly degraded by components of the replication machinery. *Nat Commun* 9, 1727.

Peter, B., and Falkenberg, M. (2020). TWINKLE and other human mitochondrial DNA helicases: structure, function and disease. *Genes* 11.

Peterson, F.C., Chen, D., Lytle, B.L., Rossi, M.N., Ahel, I., Denu, J.M., and Volkman, B.F. (2011). Orphan macrodomain protein (human C6orf130) is an O-acyl-ADP-ribose deacylase: solution structure and catalytic properties. *J Biol Chem* 286, 35955-35965.

Pfeffer, G., and Chinnery, P.F. (2013). Diagnosis and treatment of mitochondrial myopathies. *Ann Med* 45, 4-16.

Poirier, G.G., de Murcia, G., Jongstra-Bilen, J., Niedergang, C., and Mandel, P. (1982). Poly(ADP-ribosyl)ation of polynucleosomes causes relaxation of chromatin structure. *Proceedings of the National Academy of Sciences of the United States of America* 79, 3423-3427.

Ponten, J., and Saksela, E. (1967). Two established in vitro cell lines from human mesenchymal tumours. *Int J Cancer* 2, 434-447.

Prole, D.L., Chinnery, P.F., and Jones, N.S. (2020). Visualizing, quantifying and manipulating mitochondrial DNA in vivo. *The Journal of biological chemistry*.

- Puddu, F., Herzog, M., Selivanova, A., Wang, S., Zhu, J., Klein-Lavi, S., Gordon, M., Meirman, R., Millan-Zambrano, G., Ayestaran, I., *et al.* (2019). Genome architecture and stability in the *Saccharomyces cerevisiae* knockout collection. *Nature* 573, 416-420.
- Rahman, S., and Copeland, W.C. (2019). POLG-related disorders and their neurological manifestations. *Nat Rev Neurol* 15, 40-52.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8, 2281-2308.
- Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* 75, 663-670.
- Ringel, R., Sologub, M., Morozov, Y.I., Litonin, D., Cramer, P., and Temiakov, D. (2011). Structure of human mitochondrial RNA polymerase. *Nature* 478, 269-273.
- Robberson, D.L., Kasamatsu, H., and Vinograd, J. (1972). Replication of mitochondrial DNA. Circular replicative intermediates in mouse L cells. *Proceedings of the National Academy of Sciences of the United States of America* 69, 737-741.
- Roger, A.J., Munoz-Gomez, S.A., and Kamikawa, R. (2017). The Origin and Diversification of Mitochondria. *Curr Biol* 27, R1177-R1192.
- Ropp, P.A., and Copeland, W.C. (1996). Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase gamma. *Genomics* 36, 449-458.
- Rosenthal, F., Feijs, K.L., Frugier, E., Bonalli, M., Forst, A.H., Imhof, R., Winkler, H.C., Fischer, D., Caflisch, A., Hassa, P.O., *et al.* (2013). Macrodomein-containing proteins are new mono-ADP-ribosylhydrolases. *Nat Struct Mol Biol* 20, 502-507.
- Rossi, M.N., Carbone, M., Mostocotto, C., Mancone, C., Tripodi, M., Maione, R., and Amati, P. (2009). Mitochondrial localization of PARP-1 requires interaction with mitofilin and is involved in the maintenance of mitochondrial DNA integrity. *J Biol Chem* 284, 31616-31624.
- Rubio-Cosials, A., and Sola, M. (2013). U-turn DNA bending by human mitochondrial transcription factor A. *Current opinion in structural biology* 23, 116-124.
- Sagan, L. (1967). On the origin of mitosing cells. *J Theor Biol* 14, 255-274.
- Schatz, G., Haslbrunner, E., and Tuppy, H. (1964). Deoxyribonucleic Acid Associated with Yeast Mitochondria. *Biochem Biophys Res Commun* 15, 127-132.
- Sellou, H., Lebeaupin, T., Chapuis, C., Smith, R., Hegele, A., Singh, H.R., Kozlowski, M., Bultmann, S., Ladurner, A.G., Timinszky, G., *et al.* (2016). The poly(ADP-ribose)-dependent chromatin remodeler Alc1 induces local chromatin relaxation upon DNA damage. *Mol Biol Cell* 27, 3791-3799.
- Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., *et al.* (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84-87.
- Sharifi, R., Morra, R., Appel, C.D., Tallis, M., Chioza, B., Jankevicius, G., Simpson, M.A., Matic, I., Ozkan, E., Golia, B., *et al.* (2013). Deficiency of terminal ADP-ribose protein glycohydrolase TARG1/C6orf130 in neurodegenerative disease. *EMBO J* 32, 1225-1237.

Simsek, D., Furda, A., Gao, Y., Artus, J., Brunet, E., Hadjantonakis, A.K., Van Houten, B., Shuman, S., McKinnon, P.J., and Jasin, M. (2011). Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature* 471, 245-248.

Smith, A.C., and Robinson, A.J. (2019). MitoMiner v4.0: an updated database of mitochondrial localization evidence, phenotypes and diseases. *Nucleic acids research* 47, D1225-D1228.

Söllner, F., Kotthoff, C., and Ladurner, A.G. (2020a). A Novel Flow Cytometry-based Assay to Monitor Mitochondrial DNA Replication in Vivo. (manuscript ready for submission).

Söllner, F., Nardoza, A.P., Kotthoff, C., Hosp, F., Smith, R., Moeller, G.K., Tzika, E., Oliver, B.E., Sacco, F., Dreker, T., *et al.* (2020b). Protein ADP-ribosylation Regulates Human Mitochondrial DNA Replication. *Cell Rep* (under revision).

Spadafora, D., Kozhukhar, N., Chouljenko, V.N., Kousoulas, K.G., and Alexeyev, M.F. (2016). Methods for Efficient Elimination of Mitochondrial DNA from Cultured Cells. *PloS one* 11, e0154684.

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

Spelbrink, J.N., Toivonen, J.M., Hakkaart, G.A., Kurkela, J.M., Cooper, H.M., Lehtinen, S.K., Lecrenier, N., Back, J.W., Speijer, D., Foury, F., *et al.* (2000). In vivo functional analysis of the human mitochondrial DNA polymerase POLG expressed in cultured human cells. *J Biol Chem* 275, 24818-24828.

Spinelli, J.B., and Haigis, M.C. (2018). The multifaceted contributions of mitochondria to cellular metabolism. *Nature cell biology* 20, 745-754.

Stewart, J.B., and Chinnery, P.F. (2015). The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nat Rev Genet* 16, 530-542.

Stewart, J.B., and Chinnery, P.F. (2020). Extreme heterogeneity of human mitochondrial DNA from organelles to populations. *Nature Reviews Genetics*.

Stimpfel, M., Jancar, N., and Virant-Klun, I. (2018). New Challenge: Mitochondrial Epigenetics? *Stem Cell Rev Rep* 14, 13-26.

Suarez, J., Hu, Y., Makino, A., Fricovsky, E., Wang, H., and Dillmann, W.H. (2008). Alterations in mitochondrial function and cytosolic calcium induced by hyperglycemia are restored by mitochondrial transcription factor A in cardiomyocytes. *Am J Physiol, Cell Physiol* 295, C1561-1568.

Suomalainen, A., and Battersby, B.J. (2018). Mitochondrial diseases: the contribution of organelle stress responses to pathology. *Nature reviews Molecular cell biology* 19, 77-92.

Suzuki, T., Nagao, A., and Suzuki, T. (2011). Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases. *Annual review of genetics* 45, 299-329.

Szczesny, B., Brunyanszki, A., Olah, G., Mitra, S., and Szabo, C. (2014). Opposing roles of mitochondrial and nuclear PARP1 in the regulation of mitochondrial and nuclear DNA integrity: implications for the regulation of mitochondrial function. *Nucleic Acids Res* 42, 13161-13173.






- Szczesny, B., Tann, A.W., Longley, M.J., Copeland, W.C., and Mitra, S. (2008). Long patch base excision repair in mammalian mitochondrial genomes. *The Journal of biological chemistry* **283**, 26349-26356.
- Taylor, R.W., and Turnbull, D.M. (2005). Mitochondrial DNA mutations in human disease. *Nat Rev Genet* **6**, 389-402.
- Teloni, F., and Altmeyer, M. (2016). Readers of poly(ADP-ribose): designed to be fit for purpose. *Nucleic acids research* **44**, 993-1006.
- Timmis, J.N., Ayliffe, M.A., Huang, C.Y., and Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature reviews Genetics* **5**, 123-135.
- Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740.
- Tynnismaa, H., Sembongi, H., Bokori-Brown, M., Granycome, C., Ashley, N., Poulton, J., Jalanko, A., Spelbrink, J.N., Holt, I.J., and Suomalainen, A. (2004). Twinkle helicase is essential for mtDNA maintenance and regulates mtDNA copy number. *Hum Mol Genet* **13**, 3219-3227.
- Uhler J.P., and M., F. (2021). *In Vitro Analysis of mtDNA Replication.* , Vol 2192. (Springer Protocols: Humana, New York, NY ).
- Venegas, V., and Halberg, M.C. (2012). Measurement of mitochondrial DNA copy number. *Methods in molecular biology* **837**, 327-335.
- Villa, A.M., Fusi, P., Pastori, V., Amicarelli, G., Pozzi, C., Adlerstein, D., and Doglia, S.M. (2012). Ethidium bromide as a marker of mtDNA replication in living cells. *Journal of Biomedical Optics* **17**.
- Vogle, F.N., Burkhart, J.M., Gonczarowska-Jorge, H., Kucukkose, C., Taskin, A.A., Kopczynski, D., Ahrends, R., Mossmann, D., Sickmann, A., Zahedi, R.P., *et al.* (2017). Landscape of submitochondrial protein distribution. *Nature communications* **8**, 290.
- Wallace, D.C. (2010). Colloquium paper: bioenergetics, the origins of complexity, and the ascent of man. *Proceedings of the National Academy of Sciences of the United States of America* **107 Suppl 2**, 8947-8953.
- Wallace, D.C. (2012). Mitochondria and cancer. *Nature reviews Cancer* **12**, 685-698.
- Wang, C., Taki, M., Sato, Y., Tamura, Y., Yaginuma, H., Okada, Y., and Yamaguchi, S. (2019). A photostable fluorescent marker for the superresolution live imaging of the dynamic structure of the mitochondrial cristae. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 15817-15822.
- Wang, J.Q., Chen, Q., Wang, X., Wang, Q.C., Wang, Y., Cheng, H.P., Guo, C., Sun, Q., Chen, Q., and Tang, T.S. (2013a). Dysregulation of mitochondrial calcium signaling and superoxide flashes cause mitochondrial genomic DNA damage in Huntington disease. *The Journal of biological chemistry* **288**, 3070-3084.
- Wang, K.Z., Zhu, J., Dagda, R.K., Uechi, G., Cherra, S.J., 3rd, Gusdon, A.M., Balasubramani, M., and Chu, C.T. (2014). ERK-mediated phosphorylation of TFAM downregulates mitochondrial transcription: implications for Parkinson's disease. *Mitochondrion* **17**, 132-140.

- Wang, Y.E., Marinov, G.K., Wold, B.J., and Chan, D.C. (2013b). Genome-wide analysis reveals coating of the mitochondrial genome by TFAM. *PLoS one* 8, e74513.
- Wang, Z.H., Clark, C., and Geisbrecht, E.R. (2016). Analysis of mitochondrial structure and function in the *Drosophila* larval musculature. *Mitochondrion* 26, 33-42.
- Wu, Y., Li, Q., and Chen, X.Z. (2007). Detecting protein-protein interactions by Far western blotting. *Nature protocols* 2, 3278-3284.
- Yasukawa, T., and Kang, D. (2018). An overview of mammalian mitochondrial DNA replication mechanisms. *J Biochem* 164, 183-193.
- Young, M.J., and Copeland, W.C. (2016). Human mitochondrial DNA replication machinery and disease. *Curr Opin Genet Dev* 38, 52-62.
- Young, M.J., Humble, M.M., DeBalsi, K.L., Sun, K.Y., and Copeland, W.C. (2015). POLG2 disease variants: analyses reveal a dominant negative heterodimer, altered mitochondrial localization and impaired respiratory capacity. *Hum Mol Genet* 24, 5184-5197.
- Žaja, R., Aydin, G., Lippok, B.E., Feederle, R., Lüscher, B., and Feijs, K.L.H. (2020). Comparative analysis of MACROD1, MACROD2 and TARG1 expression, localisation and interactome. *Sci Rep* 10.
- Zang, L., Hong, Q., Yang, G., Gu, W., Wang, A., Dou, J., Mu, Y., Wu, D., and Lyu, Z. (2018). MACROD1/LRP16 Enhances LPS-Stimulated Inflammatory Responses by Up-Regulating a Rac1-Dependent Pathway in Adipocytes. *Cell Physiol Biochem* 51, 2591-2603.
- Zhang, H., Barcelo, J.M., Lee, B., Kohlhagen, G., Zimonjic, D.B., Popescu, N.C., and Pommier, Y. (2001). Human mitochondrial topoisomerase I. *Proc Natl Acad Sci USA* 98, 10608-10613.
- Zhang, H., Zhang, Y.W., Yasukawa, T., Dalla Rosa, I., Khiati, S., and Pommier, Y. (2014). Increased negative supercoiling of mtDNA in TOP1mt knockout mice and presence of topoisomerases I $\alpha$  and I $\beta$  in vertebrate mitochondria. *Nucleic acids research* 42, 7259-7267.
- Zhang, J., Liu, H., Luo, S., Lu, Z., Chavez-Badiola, A., Liu, Z., Yang, M., Merhi, Z., Silber, S.J., Munne, S., *et al.* (2017). Live birth derived from oocyte spindle transfer to prevent mitochondrial disease. *Reprod Biomed Online* 34, 361-368.
- Zheng, W., Khrapko, K., Collier, H.A., Thilly, W.G., and Copeland, W.C. (2006). Origins of human mitochondrial point mutations as DNA polymerase gamma-mediated errors. *Mutat Res* 599, 11-20.
- Zong, W.X., Rabinowitz, J.D., and White, E. (2016). Mitochondria and Cancer. *Mol Cell* 61, 667-676.

## Appendices

## Appendix 1: Affidavit

	<p>LUDWIG- MAXIMILIANS- UNIVERSITÄT MÜNCHEN</p>	<p>Promotionsbüro Medizinische Fakultät</p>		
<b>Affidavit</b>				

Söllner, Flavia

\_\_\_\_\_  
Surname, first name

\_\_\_\_\_  
Street

\_\_\_\_\_  
Zip code, town, country

I hereby declare, that the submitted thesis entitled:

*Regulation of Mitochondrial DNA Replication by MacroD1 and ADP-ribosylation*

.....

is my own work. I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the submitted thesis or parts thereof have not been presented as part of an examination degree to any other university.

Oslo, 28.09.2021

place, date

Flavia Söllner

Signature doctoral candidate

## Appendix 2: Abbreviations

ARH3 ADP-ribosylhydrolase 3

ARTs ADP-Ribosyltransferases

ATP adenosine triphosphate

ChIP Chromatin immunoprecipitation

ECAR extracellular acidification rate

ETC electron transport chain

GLUD1 glutamate dehydrogenase

HPF1 Histone PARylation Factor 1

HSP heavy-strand promoter

H-strand heavy-strand

IMM inner mitochondrial membrane

IMS intermembrane space

LIG3 mitochondrial Ligase 3

LSP light-strand promoter

L strand light-strand

MAR mono-ADP-ribose

MIP Mitochondrial DNA Immunoprecipitation

MRT mitochondrial replacement therapy

mtDNA mitochondrial DNA

MTS mitochondrial targeting sequence

mtSSBs mitochondrial single strand binding proteins

NCR noncoding region

nDNA nuclear DNA

nts nucleotides

OCR oxygen consumption rate

OH Origin of replication heavy-strand

OL Origin of replication light-strand

OMM outer mitochondrial membrane

OXPHOS oxidative phosphorylation

PAR poly-ADP-ribose

PARG Poly-ADP-ribose Glycohydrolase

PARPs Poly-ADP-ribose polymerases

PBMs PAR-binding linear motifs

PBZs PAR-binding zinc-fingers

PEO progressive external ophthalmoplegies

POLG DNA polymerase [?](#)

POLRMT mitochondrial DNA-directed RNA polymerase

PTM post-translational modification

ROS reactive oxygen species

TARG1 ADP-ribose protein glycohydrolase TARG1

TFAM mitochondrial transcription factor A

TFB2M mitochondrial transcription factor B2

TOP1 topoisomerase I

TOP2A topoisomerase IIA

TOPOs topoisomerases

TWINKLE mitochondrial DNA helicase

**Table 7.1: Nucleotide Abbreviations**

Nucleotide Abbreviations	
One letter code	Name
A	Adenine
G	Guanine
C	Cytosine
T	Thymine
U	Uracil
R	Purine (A or G)
Y	Pyrimidine (C or T)
N	Any nucleotide
W	Weak (A or T)
S	Strong (G or C)
M	Amino (A or C)
K	Keto (G or T)
B	Not A (G, C or T)
H	Not G (A, C or T)
D	Not C (A, G or T)
V	Not T (A, G or C)

**Table 7.2: Amino Acid Abbreviations**

Amino Acid Abbreviations		
One letter code	Three letter code	Name
G	Gly	Glycine
A	Ala	Alanine
V	Val	Valine
I	Ile	Isoleucine
L	Leu	Leucine
S	Ser	Serine
T	Thr	Threonine
C	Cys	Cysteine
M	Met	methionine
D	Asp	Aspartate
E	Glu	Glutamate
N	Asn	Asparagine
Q	Gln	Glutamine
K	Lys	Lysine
R	Arg	Arginine
H	His	Histidine
F	Phe	phenylalanine
Y	Tyr	tyrosine
W	Trp	tryptophan
P	Pro	proline
Ψ	large hydrophobic amino acid residue	
x	any amino acid residue	



## Appendix 3: Recipes

**100x Complete protease inhibitor (500  $\mu$ L)**, 1 tablet complete protease inhibitor is dissolved in 500  $\mu$ L sdH<sub>2</sub>O

**1 M DTT (20 mL)**, 3.09 g DTT, 20 mL sdH<sub>2</sub>O, Sterile filtrate, Store at -20°C.

**0.5 M EDTA (250 ml)**, 46.53g NA<sub>2</sub>EDTA\*2H<sub>2</sub>O, 200 mL sdH<sub>2</sub>O, 5g NaOH pellets, (the pH must be just above 7 for EDTA to be dissolved), Adjust pH to 8.0 with 1 M NaOH, Adjust volume to 250 mL with sdH<sub>2</sub>O, Autoclave

**50% Glycerol (100 mL)**, 50 mL 99.5% Glycerol, 50 mL sdH<sub>2</sub>O, Autoclave

**5 M KAc (250 mL)**, 122.6 g KAc, 250 mL sdH<sub>2</sub>O, Autoclave

**LB medium (1000 mL)**, 10 g Trypton, 5.0 g Yeast extract, 10 g NaCl, sdH<sub>2</sub>O to 1000 mL, Adjust pH to 7.2 with 1 M NaOH, Autoclave

**LB plates (20 plates)**, 400 mL LB medium, 6.0 g agar, Autoclave, *Supplement with antibiotic (if needed) when temperature is down to approximately 50°C.*

**2 M MgSO<sub>4</sub> (250 mL)**, 123.24 g MgSO<sub>4</sub>, 250 mL sdH<sub>2</sub>O

**5M NaCl (500 mL)**, 146.1 g NaCl, 500 mL sdH<sub>2</sub>O, Autoclave

**10 M NaOH (50 mL)**, 20 g NaOH pellets, 50 mL sdH<sub>2</sub>O

**50 mM PMSF (10 mL)**, 87 mg PMSF, 10 mL isopropanol, freeze at -20°C

**RIPA**, 50 mM Tris/Cl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.25% Na-Deoxycholate + Complete PI

**50x TAE (500mL)**, 121 g Tris-base, 50 mL 0.5 M EDTA, 28.5 mL acetic acid, dH<sub>2</sub>O to 500 mL

**1x TE buffer (500 mL)**, 5 mL 1 M Tris-HCl pH 8.0, 1 mL 0.5 M EDTA, dH<sub>2</sub>O to 500 mL, Autoclave

**1 M Tris HCl pH 6.8, pH7, pH7.6, pH 8 and pH8.8 (1000 mL)**, 121.15 g Tris base, 800 ml dH<sub>2</sub>O, Adjust pH with HCl, sdH<sub>2</sub>O to 1000 mL, Autoclave

**10% Triton X-100 (50 mL)**, 5 mL 100% Triton X-100, dH<sub>2</sub>O to 50 mL

### Solutions for SDS-PAGE

**10% APS (10 mL)**, 1 g APS, 10 mL H<sub>2</sub>O, store at -20°C

**10% SDS (500 mL)**, 50 g SDS, H<sub>2</sub>O to 500 mL

**1x SDS electrophoresis buffer (1000 mL)**, 15 g glycine, 3 g Tris base, 10 mL 10% SDS, dH<sub>2</sub>O to 1000 mL

**3x SDS loading buffer (10 mL)**, 3.75 mL 0.5 M Tris HCl pH 6.8, 0.69 g SDS, 3 mL 99.5% glycerol, dH<sub>2</sub>O to 10 mL, Add a few grains bromphenol blue. Before use add 10% 1 M DTT.

**10% Separating Gel Buffer for Immunoblotting (4x)**, 8 mL 40% acrylamide, 8 mL 1.5 M Tris HCl pH 8.8, 16 mL sdH<sub>2</sub>O, 320 μL 20% SDS, 100 μL 10 % APS, 20 μL TEMED

**Stacking gel**, 2.5 mL 40% acrylamide, 5 ml 0.5 M Tris HCl pH 8.8. 12.5 mL sdH<sub>2</sub>O, 200 μL, 20% SDS, 60 μL 10 % APS, 20 μL TEMED

### **Solutions for WB**

**10x TBS-T (1000 mL)**, 50 mL 1 M Tris HCl pH 8.0, 150 mL 5 M NaCl, 2.5 mL Tween 20, sdH<sub>2</sub>O to 500 mL

**10x PBS (1000 mL)**, 80 g NaCl, 2 g KCl, 17.4 g Na<sub>2</sub>HPO<sub>4</sub> • 2H<sub>2</sub>O, 2.4 g KH<sub>2</sub>PO<sub>4</sub> in 800 mL of sdH<sub>2</sub>O, top up to 1 L. When diluted to 1x PBS the buffer should be at pH 7.4. Addition of Tween 20 as required.

**1x TBS-T with 5% milk (100 mL)**, 25.0 g skim milk powder, 50 mL 10x TBS-T, dH<sub>2</sub>O to 500 mL

**Mild stripping buffer (1000 mL)**, 15 g glycine, 1 g SDS, 10 ml Tween20, Adjust pH to 2.2, dH<sub>2</sub>O to 1000 mL

**Harsh stripping buffer (100 mL)**, *To be prepared and used under fumehood*, 20 ml SDS 10%, 12.5 ml Tris HCl pH 6.8 0.5M, 67.5 ml dH<sub>2</sub>O water, 0.8 ml β-mercaptoethanol

**Protein binding buffer**, 100mM NaCl, 20mM Tris pH 7.6, 0.5mM EDTA, 10% glycerol, 0.1% Tween, 2% Milk, *supplement with fresh 1mM DTT*

**MIP - lysis buffer**, 50 mM Tris-HCl pH 8, 10 mM EDTA, 1% SDS, 1 x Complete Protease Inhibitor

**MIP - RIPA Buffer**, 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate, 140 mM NaCl, 1 x Complete Protease Inhibitor

**FACS sorting buffer**, 1xPBS, 2% FCS, EDTA 2 mM, HEPES 25 mM

**FACS and Imaging - cell fixation**, methanol 70%, acetone 30%, store and use at -20°C

**Mass spectrometry - buffer 1**, 50 mM Tris/Cl pH 7.5, 150 mM NaCl, 1 mM EDTA

**Mass spectrometry - buffer 2**, 50 mM Tris/Cl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.25% Na-Deoxycholate

**Mass spectrometry - elution buffer**, 2 M urea, 20 mM Tris-HCl pH 8.0, 1 mM DTT, 100 ng sequence-grade modified trypsin

**Mass spectrometry - alkylation buffer**, 2 M urea, 20 mM Tris HCl pH 8.0, 5 mM iodoacetamide

**Mass spectrometry - urea buffer**, 2 M urea, 20 mM Tris HCl pH 8.0

**Mass spectrometry - buffer A**, 0.1% [v/v] formic acid

**Mass spectrometry - buffer B**, 0.1% [v/v] formic acid, 80% [v/v] acetonitrile

## Appendix 4: Antibodies

In support of this thesis, we used both mono- and polyclonal antibodies in immunoprecipitation, western blot and immunofluorescent staining experiments. The antibodies used are listed with source, clone or lot number and company (table 7.3).

**Table 7.3: Antibodies used in this thesis.**

Antibody	Source	Clone/lot	Company
$\alpha$ -Flag	rabbit	F7425	Sigma-Aldrich
$\alpha$ -GAPDH	rabbit	ab9485	abcam
$\alpha$ -GFP	goat		in-house polyclonal
$\alpha$ -Hsp60	mouse	Clone 24	BD Transduction Laboratories
$\alpha$ -LaminB1	rabbit	ab16048	abcam
$\alpha$ -LIG3	mouse	Clone 7	BD Transduction Laboratories
$\alpha$ -MacroD1	mouse		In-house monoclonal
$\alpha$ -PARP1	rabbit	ab6079-1	abcam
$\alpha$ -POLG1	rabbit	D1Y6R	Cell Signalling
$\alpha$ -TFAM	rabbit	HPA063684	Atlas Antibodies
$\alpha$ -Tim17b			Pineda
$\alpha$ -Twinkle	mouse		gift from Anu Suomalainen
$\alpha$ -V5	mouse		gift from Elizabeth Kremmer
$\alpha$ -goat-HRP			
$\alpha$ -mouse-HRP			
$\alpha$ -rabbit-HRP			
$\alpha$ -mouse-488			
$\alpha$ -mouse-647			
$\alpha$ -rabbit-488			
$\alpha$ -rabbit-647			

We would like to especially acknowledge A. Suomalainen and E. Kremmer for gifting us antibodies used in this study.

## Appendix 5: Primers

The primers used for the various assays and techniques described in this thesis can be found in table 7.4. In the last column the experiments primers are used for are specified.

**Table 7.4: Primers used in this thesis.**

Name	Sequence	Assay
Target1 f	CACCGTCCGGCACTCGTCGGTAAGCAGG	CRISPR/Cas9
Target1 r	AAACCCTGCTTACCGACGAGTGCCGGAC	CRISPR/Cas9
Target2 f	CACCGCCAAGATCACCGGCGGCTATCGG	CRISPR/Cas9
Target2 r	AAACCCGATAGCCGCCGGTGATCTTGGC	CRISPR/Cas9
nDNA β2-m f	TGCTGTCTCCATGTTTGATGTATCT	qPCR/ChIP
nDNA β2-m r	TCTCTGCTCCCCACCTCTAAGT	qPCR/ChIP
mtDNA 16SrRNA f	GCCTTCCCCGTAAATGATA	qPCR/ChIP
mtDNA 16SrRNA r	TTATGCGATTACCGGGCTCT	qPCR/ChIP
mtDNA tRNA L f	CACCCAAGAACAGGGTTTGT	qPCR
mtDNA tRNA L r	TGGCCATGGGTATGTTGTTA	qPCR
mtDNA 685- fwdEcoRI	TATAGAATTCGGTGCAGCCGCTATTAAGGTCG	mutation load
mtDNA 685- revHindIII	TATAAAGCTTCCGATCAGGGCGTAGTTTG	mutation load
8.9kb f	TCTAAGCCTCCTTATTCGAGCCGA	mtDNA integrity
221bp f	CCCCACAAACCCCACTACTAAACCCA	mtDNA integrity
8.9kb r 221bp r	TTTCATCATGCGGAGATGTTGGATGG	mtDNA integrity
pmCherry4008 NdeI fwd	TATCATATGCCAAGTACGCCCCCTATTGACG	subcloning
pmCherry4008 XhoI rev	TATACTCGAGGCGAGACGACCTCCCCGTCGTCG	subcloning

## Appendix 6: Plasmids

The following plasmids were used for the work presented in this thesis (table 7.5).

All plasmids were verified by restriction restriction digest and sequencing analysis.

**Table 7.5: Plasmids used in this thesis.**

Plasmid	Note	Database
pSpCas9(BB)-2A-GFP	(Ran et al., 2013)	3216
pmEGFP-C1 empty		
pmCherry-C1 empty		4276
pcDNA3.1 empty C-HA		3416
Frt/TO NEO empty		3461
pcDNA3.1 His V5		
pETMCN HisV5		
MacroD1-GFP (N3)	codon optimised	3459
MacroD1-GFP (Frt/to)	codon optimised	3455
MacroD1-GFP (Frt/to-NEO)	codon optimised	3434
MacroD1-mCherry	codon optimised	4264
MacroD1TM (Frt/to)	codon optimised	3429
MacroD1TM-mCherry	codon optimised	4265
MacroD1G270E-GFP (Frt/to)	codon optimised	3425
MacroD1G270E-mCherry	codon optimised	4266
MTS-GFP (N3)	MTS from MacroD1	3440
MTS-GFP (Frt/To)	MTS from MacroD1	3430
MTS-GFP (Frt/To- NEO)	MTS from MacroD1	3438
ME2 in pCMV-SPORT6	from MPI	3414
ME2-HA in pcDNA3.1		
ME2-Flag		3475
ME2-GFP		
ME2-mCherry		
POLG1 in pCMV-SPORT6	from MPI	3413
POLG1-HA in pcDNA3.1		
POLG1-Flag		3478
POLG1-GFP		
POLG1-mCherry		4271
hPOLG1 MycHis	(Spelbrink et al., 2000)	4028

hPOLG1 D198A MycHis	(Spelbrink et al., 2000)	4029
NGFP-POLG2	(Young et al., 2015)	4022
LIG3 in pCMV-SPORT6	from MPI	3411
LIG3-HA in pcDNA3.1		
LIG3-Flag		3476
LIG3-GFP		
LIG3-mCherry		4272
TWINKLE-Flag		
TWINKLE-GFP		
TWINKLE-mCherry		4270
hTFAM-GFP	(Pastukh et al., 2007)	4021
hPARP2_isf1-GFP	from Blessing, C.	
pMA3790	(Spadafora et al., 2016)	4262
pMA4008	(Spadafora et al., 2016)	4263
pMA3790-mCherry		4268
pMA4008-mCherry		4269
Triple-TARG D125A		4274
Triple- MD1-TM		4275
NUDT16		
pGEX4-PARP10 cat domain (819-1025)		3974
$\alpha$ -helical TIM -Cherry C1		

We would like to especially acknowledge J. Spelbrink, W. Copeland and M. Alexeyev for gifting us constructs used in this study.



## Appendix 7: Cell lines

The cell lines generated and / or used to for this body of work are listed in table....

The last column of the table is reserved for notes, be it the specific names of clones or how many different monoclonal cell lines were used.

**Table 7.6: Cell lines used and/or generated in this thesis.**

Cell line	Notes
HEK293-T-REx	
HEK293-T-Rex MD1KO	cl29
HEK293-T-Rex mtGFP	
HEK293-T-Rex MacroD1-GFP	
HEK293-T-Rex MacroD1-G270E-GFP	
HEK293-T-Rex MacroD1-TM-GFP	
HEK293-T-Rex mtTARG1-GFP	
HEK293-T-Rex mtTARG1-K84A-G123E -GFP	
U2-OS	
U2-OS MD1KO	1cl11
U2-OS + mtGFP	
U2-OS MD1KO + mtGFP	
U2-OS + MD1-GFP	
U2-OS MD1KO + MD1-GFP	
U2-OS + MD1G270E-GFP	
U2-OS MD1KO + MD1G270E-GFP	
U2-OS + POLG1-GFP	
U2-OS MD1KO + POLG1-GFP	
U2-OS + LIG3-GFP	
U2-OS MD1KO + LIG3-GFP	
U2-OS + TWINKLE-GFP	
<i>U2-OS MD1KO + TWINKLE-GFP</i>	
U2-OS + NGFP-POLG2	
U2-OS MD1KO + NGFP-POLG2	

## Appendix 8: Fly lines

**Table 7.7: Fly lines used in this thesis.**

Fly line	Specification
2u	2202U (wild-type) (Boynton and Tully, 1992)
Mef2-G4	27390 from BDSC
Mef2-G4 UAS-Mito-HA-GFP	Mef2-G4 recombined with 8443 from BDSC
CG33054RNAi	KK108302 from VDR

## Appendix 9: Data Links & MTS Predictions

### Original Mass Spectrometry Analyses Sheets

#### MacroD1

<https://github.com/Flavinoid/mts/tree/master/actions/analysis/data/in/macrod1>

#### TARG1

<https://github.com/Flavinoid/mts/tree/master/actions/analysis/data/in/targ1>

### Data Links

#### Uniprot

<https://www.uniprot.org/>

#### Human Protein Atlas

<http://www.proteinatlas.org>

#### iPSORT

<http://ipsort.hgc.jp/>

#### MitoCarta2.0

[www.broadinstitute.org/files/shared/metabolism/mitocarta/human.mitocarta2.0.html](http://www.broadinstitute.org/files/shared/metabolism/mitocarta/human.mitocarta2.0.html)

#### MitoMiner v4.0

<http://mitominer.mrc-mbu.cam.ac.uk/release-4.0/begin.do>

#### MitoProtII

<https://ihg.gsf.de/ihg/mitoprot.html>

#### PhosphoSitePlus

<https://www.phosphosite.org/homeAction>

ShinyGO

<http://bioinformatics.sdstate.edu/go/>

TargetP

<http://www.cbs.dtu.dk/services/TargetP-1.1/index.php>

MTS predictions: MS Data

<https://github.com/Flavinoid/mts>

Use:

The process of generating the graphs has been broken down into a sequence of actions.

The initial action `fetch_fasta_data_from_uniprot` assumes that you have some initial input data. Currently this data is hardcoded in the file `protein_ids.csv`.

An action directory normally contains a `run.sh` script that guides you through that step. There will also be a `README.md` file that will provide you with more information.

The three main steps required to generate the data required for creating the visualizations are the following.

1. `fetch_fasta_data_from_uniprot`
2. `upload_fasta_files`
3. `build_gene_predictions`

every new step builds upon the data returned by the previous ones.

Once the prediction files have been created it is possible to start with the visualizations and analysis.

Fetch FASTA data from Uniprot

[https://github.com/Flavinoid/mts/tree/master/actions/fetch\\_fasta\\_data\\_from\\_uniprot](https://github.com/Flavinoid/mts/tree/master/actions/fetch_fasta_data_from_uniprot)

To use this action, you need to make sure that the `run.sh` script is executable

```
$ chmod 755 run.sh
```

to check that it works it is advisable to then run the following command

```
$ ./run.sh --dev
```

The --dev flag will use a file called protein\_ids\_sample.csv. This file is a small selection of the full set of data protein\_ids.csv.

Using the main file takes a while and generates a roughly 5mb output file.

After the data has been fetched from Uniprot and the resulting csv written to either data/out/uniprot\_results\_sample.csv or data/out/uniprot\_results.csv a directory /out/fasta will be created that contains FASTA files suitable for uploading to ?.

#### *working with existing data*

If you already have a data/out/uniprot\_results.csv file and just want to generate the FASTA files for uploading - you can use the command.

```
$ ./generate_fasta_files.sh
```

Again you will need to make sure that it is executable. By default this will use the file data/out/uniprot\_results.csv as input. If you call the command with the --dev flag (./generate\_fasta\_files.sh --dev) then it will read the file from data/out/uniprot\_results\_sample.csv.

Upload FASTA files to TargetP

[https://github.com/Flavinoid/mts/tree/master/actions/upload\\_fasta\\_files](https://github.com/Flavinoid/mts/tree/master/actions/upload_fasta_files)

Unfortunately, the first half of this action is currently a manual step.

Using <http://www.cbs.dtu.dk/services/TargetP-1.1/index.php>

You can upload the files generated by fetch\_fasta\_data\_from\_uniprot one by one. The server should take about 30 seconds or so before returning a result.

## Step 1

Once the computation is complete you should be shown an output as follows:

### SUBMISSION

Paste a single sequence or several sequences in **FASTA** format into the field below:

Submit a file in **FASTA** format directly from your local disk:

Choose file targetp\_input1.fasta

#### Organism group

- Non-plant  
 Plant

#### Prediction scope

- Perform cleavage site predictions

#### Cutoffs

- no cutoffs; winner-takes-all (default)  
 specificity >0.95 (predefined set of cutoffs that yielded this specificity on the TargetP test sets)  
 specificity >0.90 (predefined set of cutoffs that yielded this specificity on the TargetP test sets)  
 define your own cutoffs (0.00 - 1.00): **cTP:**  **mTP:**  **SP:**  **other:**

Submit

Clear fields

#### Restrictions:

At most 2,000 sequences and 200,000 amino acids per submission; each sequence not more than 4,000 amino acids.

#### Confidentiality:

The sequences are kept confidential and will be deleted after processing.

## Step 2



## TargetP 1.1 Server - prediction results

Technical University of Denmark

```
### targetp v1.1 prediction results #####
Number of query sequences: 480
Cleavage site predictions not included.
Using NON-PLANT networks.
```

Name	Len	mTP	SP	other	Loc	RC
sp_AOFGR8-5_ESYT2_HU	328	0.077	0.148	0.869	-	2
sp_AOFGR8-4_ESYT2_HU	527	0.101	0.133	0.824	-	2
sp_AOFGR8-6_ESYT2_HU	942	0.904	0.013	0.230	M	2
sp_AOFGR8-2_ESYT2_HU	893	0.753	0.021	0.336	M	3
tr_H7BXI1_H7BXI1_HUM	866	0.101	0.481	0.666	-	5
sp_AOFGR8_ESYT2_HUMA	921	0.904	0.013	0.230	M	2
sp_ALL0T0_HACL2_HUMA	632	0.015	0.943	0.180	S	2
tr_E9PL44_E9PL44_HUM	166	0.015	0.943	0.180	S	2
tr_MOQZX5_MOQZX5_HUM	229	0.057	0.316	0.753	-	3
tr_MOR1B5_MOR1B5_HUM	165	0.461	0.345	0.174	M	5
tr_E9PJS0_E9PJS0_HUM	262	0.015	0.943	0.180	S	2
tr_MOR026_MOR026_HUM	525	0.461	0.345	0.174	M	5
tr_A2A274_A2A274_HUM	805	0.831	0.042	0.147	M	2
sp_Q99798_ACON_HUMAN	780	0.831	0.042	0.147	M	2
tr_D6RF25_D6RF25_HUM	142	0.249	0.062	0.731	-	3
tr_E7EV99_E7EV99_HUM	632	0.313	0.051	0.663	-	4
sp_P35611_ADDA_HUMAN	737	0.313	0.051	0.663	-	4
sp_P35611-2_ADDA_HUM	631	0.313	0.051	0.663	-	4
tr_HOYFD8_HOYFD8_HUM	105	0.093	0.067	0.917	-	1
tr_D6RAH3_D6RAH3_HUM	137	0.313	0.051	0.663	-	4
tr_E7ENY0_E7ENY0_HUM	663	0.313	0.051	0.663	-	4
tr_D6RJE2_D6RJE2_HUM	69	0.251	0.062	0.726	-	3
tr_HOYG19_HOYG19_HUM	114	0.174	0.039	0.865	-	2
sp_P35611-3_ADDA_HUM	768	0.313	0.051	0.663	-	4
tr_HOY9H2_HOY9H2_HUM	444	0.070	0.259	0.824	-	3
sp_P35611_ADDA_HUMAN	737	0.313	0.051	0.663	-	4

These results should be added to the file /data/out/predictions.txt.

Mapping Protein-ID predictions onto Genes

[https://github.com/Flavinoid/mts/tree/master/actions/build\\_gene\\_predictions](https://github.com/Flavinoid/mts/tree/master/actions/build_gene_predictions)

*cleaning up the Data*

By now you should have the file actions/upload\_fasta\_files/data/out/predictions.txt

```
$/run.sh
```

This script performs two steps. Firstly, it will extract the protein-ids that are predicted as being a "Mitochondrial targeting". The output is then written to /data/out/mito\_targeting\_proteinIDs.txt. After this it will use the data in

./data/in/gene\_and\_associated\_proteins.csv in combination with the newly generated file to create the file /data/out/predicated\_mito\_targeting\_genes.txt.

The files are just normal text files with a name per row.

*note*

the returned values from targetP look like: sp\_Q96RR1\_PEO1\_HUMAN. We use the UniProtKB section (Q96RR1) to link the protein-ids back to the file ./data/in/gene\_and\_associated\_proteins.csv

In general, it seems like the long versions have the form

<sp|tr>\_<UniProtKB>\_<GENE-NAME>\_<HUMAN|Hm>(\_<number>)

Volcano Plot

<https://github.com/Flavinoid/mts/tree/master/actions/analysis>



MTS predictions: PARPs

PARP2/ARTD2 (Q9UGN5-1 and Q9UGN5-2)

3/5/2019

MitoProt II - v1.101

## MitoProt II - v1.101

---

Sequence name: PARP2  
Input sequence length : 583 aa

---

### VALUES OF COMPUTED PARAMETERS

Net charge of query sequence : +9  
Analysed region : 18  
Number of basic residues in targeting sequence : 6  
Number of acidic residues in targeting sequence : 0  
**Cleavage site** : **25**  
Cleaved sequence : MAARRRRSTGGGRARALNESKRVN

---

### HYDROPHOBIC SCALE USED

	GES	KD	GVH1	ECS
H17	0.900	1.165	-0.060	0.432
MesoH	-0.231	0.229	-0.324	0.221
MuHd_075	15.534	10.168	3.794	4.300
MuHd_095	9.967	7.654	2.121	2.206
MuHd_100	10.582	9.121	3.157	2.719
MuHd_105	6.176	7.657	2.819	1.957
Hmax_075	-2.700	3.900	-1.005	1.420
Hmax_095	-11.375	-0.700	-3.612	-1.391
Hmax_100	-2.900	-0.200	-2.121	0.020
Hmax_105	-16.400	-2.333	-4.021	-1.700

---

### PROBABILITY

of export to mitochondria: **0.2492**

---

[\[HELP\]](#)

---

Please cite as reference:

M.G. Claros, P. Vincens. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur. J. Biochem. 241, 770-786.

<https://ihg.gsf.de/cgi-bin/paolo/mitofilter?seq=MAARRRRSTGGGRARALNESKRNVNNGNTAPEDSSPAKTRRCQRQESKKMPVAGGKANKD%0D%0ARTEDK...> 1/1

# MitoProt II - v1.101

Input sequence length : 573 aa

## VALUES OF COMPUTED PARAMETERS

Net charge of query sequence : +11  
 Analysed region : 18  
 Number of basic residues in targeting sequence : 6  
 Number of acidic residues in targeting sequence : 0  
**Cleavage site** : **25**  
 Cleaved sequence : MAARRRRSTGGGRARALNESKRVN

## HYDROPHOBIC SCALE USED

	GES	KD	GVH1	ECS
H17	1.247	1.471	0.087	0.567
MesoH	0.137	0.322	-0.220	0.286
MuHd_075	15.534	10.168	3.794	4.300
MuHd_095	9.967	7.654	2.121	2.206
MuHd_100	10.582	9.121	3.157	2.719
MuHd_105	6.176	7.657	2.819	1.957
Hmax_075	-2.700	3.900	-1.005	1.420
Hmax_095	-11.375	-0.700	-3.612	-1.391
Hmax_100	-2.900	-0.200	-2.121	0.020
Hmax_105	-16.400	-2.333	-4.021	-1.700

## PROBABILITY

of export to mitochondria: **0.2502**

[\[HELP\]](#)

Please cite as reference:

M.G. Claros, P. Vincens. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur. J. Biochem. 241, 770-786.



## TargetP 1.1 Server - prediction results

Technical University of Denmark

---

```
### targetp v1.1 prediction results #####  
Number of query sequences: 1  
Cleavage site predictions included.  
Using NON-PLANT networks.
```

Name	Len	mTP	SP	other	Loc	RC	TPlen
Sequence	583	0.596	0.042	0.479	M	5	7
cutoff		0.000	0.000	0.000			

**Explain** the output. Go [back](#).

---



## TargetP 1.1 Server - prediction results

Technical University of Denmark

---

```
### targetp v1.1 prediction results #####  
Number of query sequences: 1  
Cleavage site predictions included.  
Using NON-PLANT networks.
```

Name	Len	mTP	SP	other	Loc	RC	TPlen
Sequence	570	0.552	0.035	0.538	M	5	7
cutoff		0.000	0.000	0.000			

**Explain** the output. Go [back](#).

---

## iPSORT Prediction

**Predicted as:** having a mitochondrial targeting peptide

Sequence (Type: nonplant)

```

1 MAARR RRSTG GGRAR ALNES KRVNN GNTAP EDSSP AKKTR RCQRQ ESKKM
51 PVAGG KANKD RTEDK QDGMP GRSWA SKRVS ESVKA LLLKG KAPVD PECTA
101 KVGKA HVYCE GNDVY DVMLN QTNLQ FNNNK YYLIQ LLEDD AQRNF SVWMR
151 WGRVG KMGQH SLVAC SGNLN KAKEI FQKKF LDKTK NNWED REKFE KVP GK
201 YDMLQ MDYAT NTQDE EETKK EESLK SPLKP ESQLD LRVQE LIKLI CNVQA
251 MEEMM MEMKY NTKKA PLGKL TVAQI KAGYQ SLKKI EDCIR AGQHG RALME
301 ACNEF YTRIP HDFGL RTPPL IRTQK ELSEK IQLLE ALGDI EIAIK LVKTE
351 LQSPE HPLDQ HYRNL HCALR PLDHE SYEFK VISQY LQSTH APTHS DYTMT
401 LLDLF EVEKD GEKEA FREDL HNRML LWHGS RMSNW VGILS HGLRI APPEA
451 PITGY MFGKG IYFAD MSSKS ANYCF ASRLK NTGLL LLSEV ALGQC NELLE
501 ANPKA EGLLQ GKHST KGLGK MAPSS AHFVT LNGST VPLGP ASDTG ILNPD
551 GYTLN YNEYI VYNPN QVRMR YLLKV QFNFL QLW

```

Values used for reasoning

Node	Answer	View	Substring	Value(s)	Plot
1. Signal peptide?	No	Average Hydropathy (KYTJ820101)	[6,20]	-1.40667 ( >= 0.953? No)	<a href="#">show</a>
2. Mitochondrial ?	Yes	Average Net Charge (KLEP840101)	[1,30]	0.233333 ( >= 0.083? Yes)	<a href="#">show</a>
		Indexing: AI1 Pattern: 221121122 (ins/del <= 3)	[1,30]	MAARR-RRSTGGGRARALNESKRVNNGNTAP 22211-1122000121220020120000222 221121122	--

\* This color means "not used".

Name	Alphabet Indexing		
	0	1	2
AI1	DEGHKN	IR	ACFLMPQSTVWY
AI2	ACDEFGHLMNQSTVWY	KR	IP

[Return to iPSORT Home](#)

## iPSORT Prediction

**Predicted as:** [having a mitochondrial targeting peptide](#)

Sequence (Type: nonplant)

```

1 MAARR RRTSG GGRAR ALNES KRVNN GNTAP EDSSP AKKTR RCQRQ ESKKM
51 PVAGG KANKD RTEDK QDESV KALLL KGKAP VDPEC TAKVG KAHVY CEGND
101 VYDVM LNQTN LQFNN NKYYL IQLLE DDAQR NFSVW MRWGR VGKMG QHSLV
151 ACSGN LNKAK EIFQK KFLDK TKNNW EDREK FEKVP GKYDM LQMDY ATNTQ
201 DEEET KKEES LKSPL KPESQ LDLRV QELIK LICNV QAMEE MMEM KYNTK
251 KAPLG KLTVA QIKAG YQSLK KIEDC IRAGQ HGRAL MEACN EFYTR IPHDF
301 GLRTP PLIRT QKELS EKIQL LEALG DIEIA IKLVK TELQS PEHPL DQHYR
351 NLHCA LRPLD HESYE FKVIS QYLQS THAPT HSDYT MTLTD LFEVE KDGEK
401 EAFRE DLHNR MLLWH GSRMS NWWGI LSHGL RIAPP EAPIT GYMFG KGIYF
451 ADMSS KSANY CFASR LKNTG LLLLS EVALG QCNEL LEANP KAEBL LQKGH
501 STKGL GKMAP SSAHF VTLNG STVPL GPASD TGILN PDGYT LNYNE YIVYN
551 PNQVR MRYLL KVQFN FLQLW

```

Values used for reasoning

Node	Answer	View	Substring	Value(s)	Plot
1. Signal peptide?	No	Average Hydropathy (KYTJ820101)	[6,20]	-1.40667 ( >= 0.953? No)	<a href="#">show</a>
2. Mitochondrial ?	Yes	Average Net Charge (KLEP840101)	[1,30]	0.233333 ( >= 0.083? Yes)	<a href="#">show</a>
		Indexing: AII Pattern: 221121122 (ins/del <= 3)	[1,30]	MAARR-RRSTGGGRARALNESKRVNNGNTAP 22211-1122000121220020120000222 221121122	--

\* This color means "not used".

Name	Alphabet Indexing		
	0	1	2
AII	DEGHKN	IR	ACFLMPQSTVWY
AI2	ACDEFGHLMNQSTVWY	KR	IP

[Return to iPSORT Home](#)

## MitoProt II - v1.101

---

Sequence name: PARP5A  
Input sequence length : 1330 aa

---

### VALUES OF COMPUTED PARAMETERS

Net charge of query sequence : -5  
Analysed region : 66  
Number of basic residues in targeting sequence : 3  
Number of acidic residues in targeting sequence : 0  
**Cleavage site : not predictable**  
Cleaved sequence :

---

### HYDROPHOBIC SCALE USED

	GES	KD	GVH1	ECS
H17	1.618	1.882	0.490	0.615
MesoH	0.855	0.557	0.039	0.337
MuHd_075	23.974	13.369	5.632	5.377
MuHd_095	20.229	16.842	5.792	4.205
MuHd_100	24.421	16.480	6.519	5.230
MuHd_105	24.485	17.799	6.372	5.166
Hmax_075	-6.650	9.333	-8.246	-0.297
Hmax_095	-7.525	9.300	-5.165	-0.997
Hmax_100	-3.400	11.900	-5.540	0.460
Hmax_105	-6.563	11.500	-6.325	-0.128

---

### PROBABILITY

of export to mitochondria: **0.0077**

---

[\[HELP\]](#)

---

Please cite as reference:

M.G. Claros, P. Vincens. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur. J. Biochem. 241, 770-786.



# TargetP 1.1 Server - prediction results

Technical University of Denmark

---

```
### targetp v1.1 prediction results #####  
Number of query sequences: 1  
Cleavage site predictions included.  
Using NON-PLANT networks.
```

Name	Len	mTP	SP	other	Loc	RC	TPlen
Sequence	1327	0.823	0.033	0.223	M	2	7
cutoff		0.000	0.000	0.000			

[Explain the output.](#) [Go back.](#)

---



## iPSORT Prediction

**Predicted as: not having signal or mitochondrial targeting peptide**

Sequence (Type: nonplant)

```

1 MAASR RSQHH HHHHQ QQLQP APGAS APPPP PPPPL SPGLA PGTTP ASPTA
51 SGLAP FASPR HGLAL PEGDG SRDPP DRPRS PDPVD GTSCC STTST ICTVA
101 AAPVV PAVST SSAAG VAPNP AGSGS NNSPS SSSSP TSSSS SSPSS PGSSL
151 AESPE AAGVS STAPL GPGAA GPGTG VPAVS GALRE LLEAC RNGDV SRVKR
201 LVDAA NVNAK DMAGR KSSPL HFAAG FGRKD VVEHL LQMGa NVHAR DDGGL
251 IPLHN ACSFG HAEVV SLLLC QGADP NARDN WNYTP LHEAA IKGKI DVCIV
301 LLQHG ADPNI RNTDG KSALD LADPS AKAVL TGEYK KDELL EAARS GNEEK
351 LMALL TPLNV NCHAS DGRKS TPLHL AAGYN RVRIV QLLLQ HGADV HAKDK
401 GGLVP LHNAC SYGHY EVTEL LLKHG ACVNA MDLWQ FTPLH EAASK NRVEV
451 CSLLL SHGAD PTLVN CHGKS AVDMA PTPEL RERLT YEFKG HSLLQ AAREA
501 DLAKV KKTLA LEIIN FKQPQ SHETA LHCAV ASLHP KRKQV TELL LKGAN
551 VNEKN KDFMT PLHVA AERAH NDVME VLHKH GAKMN ALDTL QOTAL HRAAL
601 AGHLQ TCRL LSYGS DPSII SLQGF TAAQM GNEAV QQILS ESTPI RTSDV
651 DYRL L EASKA GDLET VKQLC SSQNV NCRDL EGRHS TPLHF AAGYN RVSVV
701 EYLLH HGADV HAKDK GGLVP LHNAC SYGHY EVAEL LVRHG ASVNV ADLWK
751 FTPLH EAAAK GKYEI CKLLL KHGAD PTKKN RDGNT PLDLV KEGDT DIQDL
801 LRGDA ALLDA AKKGC LARVQ KLCTP ENINC RDTQG RNSTP LHLAA GYNL
851 EVAEY LLEHG ADVNA QDKGG LIPLH NAASY GHVDI AALLI KYNTC VNATD
901 KWAF T PLHEA AQKGR TQLCA LLLAH GADPT MKNQE GQTPL DLATA DDIRA
951 LLIDA MPPEA LPTCF KPQAT VVSAS LISPA STPSC LSAAS SIDNL TGPLA
1001 ELAVG GASNA GDGAA GTERK EGEVA GLDMN ISQFL KSLGL EHLRD IFETE
1051 QITLD VLADM GHEEL KEIGI NAYGH RHKLI KGVER LLGGQ QGTNP YLTFH
1101 CVNQG TILLD LAPED KEYQS VEEEM QSTIR EHRDG GNAGG IFNRY NVIRI
1151 QKVVN KKLRE RFCHR QKEVS EENHN HHNER MLFHG SPFIN AIIHK GFDER
1201 HAYIG GMFGA GIYFA ENSSK SNQYV YGIGG GTGCP THKDR SCYIC HRQML
1251 FCRVT LGKSF LQFST MKMAH APPGH HSVIG RPSVN GLAYA EYVIY RGEQA
1301 YPEYL ITYQI MKPEA PSQTA TAAEQ KT

```

Values used for reasoning

Node	Answer	View	Substring	Value(s)	Plot
1. Signal peptide?	No	Average Hydropathy (KYTJ820101)	[6,20]	-2.65333 ( >= 0.953? No)	<a href="#">show</a>
2. Mitochondrial	No	Average Net Charge	[1,30]	0.0666667 ( >= 0.083? No)	<a href="#">show</a>

## MitoProt II - v1.101

Sequence name: PARP16  
 Input sequence length : 322 aa

### VALUES OF COMPUTED PARAMETERS

Net charge of query sequence : +10  
 Analysed region : 10  
 Number of basic residues in targeting sequence : 1  
 Number of acidic residues in targeting sequence : 0  
**Cleavage site** : **12**  
 Cleaved sequence : MQPSGWAAARE

### HYDROPHOBIC SCALE USED

	GES	KD	GVH1	ECS
H17	2.329	2.776	0.485	0.858
MesoH	0.231	0.445	-0.327	0.306
MuHd_075	37.661	20.166	10.260	7.175
MuHd_095	18.154	7.747	3.320	3.440
MuHd_100	5.133	2.572	0.637	1.114
MuHd_105	6.015	5.795	2.070	0.631
Hmax_075	12.400	11.300	1.709	4.340
Hmax_095	4.375	6.213	-0.703	2.922
Hmax_100	-1.800	0.500	-2.889	1.420
Hmax_105	4.600	4.700	-1.108	2.560

### PROBABILITY

of export to mitochondria: **0.1372**

[\[HELP\]](#)

Please cite as reference:

M.G. Claros, P. Vincens. Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur. J. Biochem. 241, 770-786.



# TargetP 1.1 Server - prediction results

Technical University of Denmark

```

### targetp v1.1 prediction results #####
Number of query sequences: 1
Cleavage site predictions included.
Using NON-PLANT networks.

```

Name	Len	mTP	SP	other	Loc	RC	TPlen
Sequence	322	0.459	0.033	0.553	_	5	-
cutoff		0.000	0.000	0.000			

[Explain](#) the output. Go [back](#).

## iPSORT Prediction

**Predicted as:** not having signal or mitochondrial targeting peptide

Sequence (Type: nonplant)

---

```

1 MQPSG WAAAR EAAGR DMLAA DLRCs LFASA LQSYK RDSVL RPFPA SYARG
51 DCKDF EALLA DASKL PNLKE LLQSS GDNHK RAWDL VSWIL SSKVL TIHSA
101 GKAEF EKIQK LTGAP HTPVP APDFL FEIEY FDPAN AKFYE TKGER DLIYA
151 FHGSR LENFH SIIHN GLHCH LNKTS LFGEG TYLTS DLSLA LIYSP HGHGW
201 QHSLL GPILS CVAVC EVIDH PDVKC QTKKK DSKEI DRRRA RIKHS EGGDI
251 PKKYF VVTNN QLLRV KYLLV YSQKP PKRAS SQLSW FSSHW FTVMI SLYLL
301 LLLIV SVINS SAFQH FWNRA KR
  
```

---

Values used for reasoning

Node	Answer	View	Substring	Value(s)	Plot
<b>1. Signal peptide?</b>	No	Average Hydropathy (KYTJ820101)	[6,20]	0.0666667 ( >= 0.953? No)	<a href="#">show</a>
<b>2. Mitochondrial ?</b>	No	Average Net Charge (KLEP840101)	[1,30]	0 ( >= 0.083? No)	<a href="#">show</a>
		Indexing: AI1 Pattern: 221121122 (ins/del <= 3)	[1,30]	MQPSGWAAAREEAAGRDMLAADLRCsLFASA 22220222210220102222021222222 NOMATCH	--

\* This color means "not used".

Name	Alphabet Indexing		
	0	1	2
<b>AI1</b>	DEGHKN	IR	ACFLMPQSTVWY
<b>AI2</b>	ACDEFGHLMNQSTVWY	KR	IP

[Return to iPSORT Home](#)

---