New properties of mitochondrial aminoacyl-tRNA synthetases and their connection to mitochondrial diseases

Hagen Schwenzer

▶ **To cite this version:**

Hagen Schwenzer. New properties of mitochondrial aminoacyl-tRNA synthetases and their connection to mitochondrial diseases. Subcellular Processes [q-bio.SC]. Université de Strasbourg, 2013. English. NNT : 2013STRAJ045 . tel-01124189

HAL Id: tel-01124189
https://theses.hal.science/tel-01124189

Submitted on 6 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE des SCIENCES DE LA VIE ET DE LA SANTE

Architecture et Réactivité de l'ARN – UPR 9002 du CNRS

THÈSE

présentée par :

Hagen SCHWENZER

pour obtenir le grade de :

Docteur de l'université de Strasbourg

Discipline / Spécialité : Aspects moléculaires de la biologie

"Aminoacyl-ARNt synthétases mitochondriales humaines: aspects fondamentaux et contribution à la compréhension de pathologies reliées"

soutenue le : 21 octobre 2013

THÈSE dirigée par :

[Dr. Sissler Marie]	Directrice de Recherche au CNRS, Université de Strasbourg
[Pr. Florentz Catherine]	Professeur des Universités, Université de Strasbourg

RAPPORTEURS :

[Pr. Meisinger Chris]	Professeur des Université, Université de Freiburg, Allemagne
[Dr. Mirande Marc]	Directeur des Recherche au CNRS, Laboratoire d'Enzymologie et Biochimie Structurales, Gif-sur-Yvette

AUTRES MEMBRES DU JURY :

[Dr. Lombès Anne]	Directrice de Recherche, INSERM, Institue Cochin Paris
[Dr. Tarassov Ivan]	Directeur de Recherche, CNRS, Université de Strasbourg

English version of the title:

“New properties of mitochondrial aminoacyl-tRNA synthetases and their connection to mitochondrial diseases”

Für meine Familie

Acknowledgment

First of all, I would like to thank Dr. Marc Mirande, Laboratoire d'Enzymologie et Biochemie Structurales, Gif-sur-Yvette, Dr. Anne Lombes, Institue Cochin-Université Paris Descarts; Prof. Dr. Chris Meisinger, Institute for Biochemistry and Molecular Biology, Freiburg, and Dr. Ivan Tarassov, GMGM, Université de Strasbourg for their interest and evaluation of my work.

I would like to thank Prof. Dr Eric Westhof, Director of the UPR9002, to give me the opportunity to work in the Unit.

I would like to express my gratitude to Prof. Catherine Florentz and Dr. Marie Sissler for having given me the opportunity to perform my PhD as part of their research group. In addition I would like to thank Marie for supervising me over the period of my PhD, for all the discussions and help.

I would like to thank the Team Florentz (Joern, Franzi, Agnès, Loukmane, Bernard and Claude) for all their help during my thesis. In addition, I would like to thank Frank for all the constructive discussions and Pizza evenings we had during his stay in Strasbourg.

I also want to thanks Redmond Smyth for all the stylistic improvements in English (or Irish) in this manuscript and all other texts I had to write.

Last but not least, I like to thanks all my colleagues and friends at the UPR9002, who really enriched my life in Strasbourg. THANKS.

Abbreviations

A	ampere
aa	amino acide
aaRS	aminoacyl-tRNA synthetase
Asp	aspartic acid
ATP	adenosine triphosphat
BN-PAGE	blue native PAGE
cDNA	complementary DNA
Ci	Curie
cpm	counts per minute
d	day
Da	dalton
DMP	dimethyl pimelimidate
DMSO	dimethylsulfoxide
DNA	desoxyribonucleic acid
dNTPs	desoxynucleoside triphosphate
E.coli	Escherichia coli
EDTA	Ethlenediaminetetraacetic acid
EF	elongation factor
FCS	foetal calf serum
g	*gravitational force*
GFP	green fluoresent protein
h	hours
ID	identification
IF	initiation factor
Ig	immune globulin
k	kilo
Kac	potassium acetate
kb	kilobase
kV	kilo V
LB	luria broth
M	molarity
MARS	multi aminoacyl-tRNA synthetase complex
min	minute
MIP	mitochondrial intermediate peptidase
MPP	mitochondrial processing peptidase
mRNA	messenger RNA
mRNA	messenger RNA
MRPL39	mitochondrial ribosomal protein L39
mt	mitochondrial
MTS	mitochondrial targeting sequence

MVA	Modified Vaccinia Ankara
MW	molecular weight
n	nano
NaOAc	sodium acetate
nc	noncoding
OD	optical density
OXOPHOS	Oxidative phosphorylation
PAGE	Polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PVDF	polyvinylidene difluoride
RC	respiratory chain
RNA	ribonucleic acid
rRNA	ribosomal RNA
RT-PCR	reverse-transcribed PCR
s	second
SDS	sodium dodecylsulfate
shRNA	short hairpin RNA
siRNA	small interfering RNA
SOD2	Superoxide dismutase 2
TIM	translocase of the inner membrane
TOM	translocase of the outer mitochondrial membrane
tRNA	transfere RNA
U	unit
UV	ultra-violet
V	volt
VDAC	Voltage-dependent anion channel
W	Watt
wt	wild type
µ	micro
µg	microgramm

Table of Contents

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

1 MITOCHONDRIA AT A GLANCE 1

2 THE MAMMALIAN MITOCHONDRIAL TRANSLATION MACHINERY 5

2.1 HUMAN MITOCHONDRIAL GENOME 5

2.2 KEY ACTORS OF MITOCHONDRIAL TRANSLATION MACHINERY 7

2.3 PECULIAR ASPECTS OF MITOCHONDRIAL TRANSLATION 9

2.4 AMINOACYL-TRNA SYNTHETASES ARE KEY MOLECULES OF TRANSLATION 10

3 IMPORT OF NUCLEUS-ENCODED MITOCHONDRIAL PROTEINS 13

3.1 THE MITOCHONDRIAL PROTEOME 13

3.2 MITOCHONDRIAL IMPORT MACHINERY 14

4 HUMAN TRANSLATION SYSTEM AND MITOCHONDRIAL DISEASES 20

4.1 INTEGRATION OF MT-AARSS INTO THE MITOCHONDRIAL GENOME EXPRESSION 20

BOOK CHAPTER #1 22

BOOK CHAPTER #2 23

5 NON-TRANSLATIONAL FUNCTIONS OF AMINOACYL-TRNA SYNTHETASES 24

5.1 ORGANIZATION OF CYTOSOLIC AARSS 24

5.2 AARSS BEYOND TRANSLATION 25

5.3 GENERAL CONCEPTS OF AARSS TO GAIN NEW FUNCTIONS 26

5.4 NON-TRANSLATIONAL FUNCTIONS AND THEIR LINK TO DISEASES 30

OBJECTIVES OF THE THESIS 31

RESULTS

CHAPTER I: BACTERIAL INSERTION DOMAIN OF MITOCHONDRIAL ASPRS 34

1 INTRODUCTION 34

2 ARTICLE #1 36

3 DETECTION OF THE PUTATIVE PROTEIN CORRESPONDING TO MT-ASPRSΔEXON13 37

4 RNAI KNOCK-DOWN OF MT-ASPRS AND MT-ASPRSΔEXON13 40

4.1 PROLIFERATION OF TRANSIENT TRANSFECTED CELLS WITH SHRNA TARGETING MT-ASPRS OR MT-ASPRSΔEXONS13 40

4.2 PROLIFERATION OF STABLY TRANSFECTED CELLS WITH SIRNA TARGETING MT-ASPRS OR MT-ASPRSΔEXONS13 43

5 ACHIEVEMENTS AT A GLANCE: 48

6 GRAPHICAL SUMMARY 48

CHAPTER II: SUB MITOCHONDRIAL LOCALIZATION OF AARSS 49

1 INTRODUCTION 49

2 SUB-MITOCHONDRIAL FRACTIONATION AND DETECTION OF AARSS 52

2.1 VALIDATION OF THE MITOCHONDRIAL FRACTIONATION PROTOCOL 52

2.2 ESTABLISHMENT OF THE SUB-MITOCHONDRIAL LOCALIZATION OF MT-ASPRS 54

2.3 DECIPHERING THE NATURE OF THE INTERACTION OF MT-ASPRS TO OR WITHIN THE MEMBRANE 55

3 DECIPHERING THE DISTRIBUTION OF THE FULL SET OF MT-AARSS INSIDE MITOCHONDRIA 62

4 SUB-MITOCHONDRIAL LOCALIZATION OF TRNAS 65

5 PRELIMINARY ANALYSIS OF THE SUB-MITOCHONDRIAL PROTEOME 69

5.1 PROTEOMIC ANALYSIS OF THE SUB-MITOCHONDRIAL LOCALIZATION OF AARSS 69

5.2 ADVANCED PROTEOMIC ANALYSIS OF FRACTIONATED MITOCHONDRIA 71

5.3 PROTEOMIC ANALYSIS OF MITOPLAST FRACTIONATION 73

5.4 FUTURE POWER OF THE PROTEOMIC ANALYSIS OF THE SUB-MITOCHONDRIAL FRACTIONS 75

6 ACHIEVEMENTS AT A GLANCE: 77

7 GRAPHICAL SUMMARY 77

CHAPTER III: TOWARDS THE DEFINITION OF N-TERMINAL MATURATION SITE OF MT-AARSS 78

1 INTRODUCTION 78

2 ENGINEERED EXPRESSION SYSTEM FOR NATURALLY MATURATED MAMMALIAN PROTEINS UPON MITOCHONDRIAL IMPORT 79

3 ARTICLE #2 81

4 VARIATION OF THE THEORETICAL PI WITH THE LENGTH OF N-TERMINAL SEQUENCE OF MT-AARSS 82

5 EXPERIMENTAL DETERMINATION OF THE N-TERMINAL PROCESSING SITE OF MT-ASPRS 87

6 MUTATIONAL ANALYSIS OF MTS OF MT-ASPRS AND SOD2 88

7 ACHIEVEMENTS AT A GLANCE 92

8 GRAPHICAL SUMMARY 92

CHAPTER IV: BIOLOGICAL PARTNERS OF THE HUMAN MT-ASPRS 93

1 INTRODUCTION 93

2 DATA MINING 95

2.1 STRING: FUNCTIONAL PROTEIN ASSOCIATION NETWORK 95

2.2 DATA MINING OF EXISTING HIGH-THROUGHPUT DATA 100

3 IDENTIFICATION OF PUTATIVE PARTNER PROTEINS OF MT-ASPRS 102

3.1 INVOLVEMENT OF MT-ASPRS IN MACROMOLECULAR COMPLEXES 102

3.2 IDENTIFICATION OF PUTATIVE PROTEIN PARTNERS 106

4 ACHIVEMENTS AT A GLANCE 114

5 GRAPHICAL SUMMARY 114

CHAPTER V DISORDERS OF THE HUMAN MITOCHONDRIAL AMINOACYLATION SYSTEM 115

1 *IN VITRO* ANALYSES OF THE AMINOACYLATION PROPERTIES OF A MUTANT TRNA^{ASP} FOUND IN PATIENTS WITH SIDEROBLASTIC ANEMIA 117

2 ANALYSIS OF MUTATIONS FOUND IN PATIENTS WITH A LEIGH-LIKE SYNDROME 120

3 PATHOLOGY-RELATED MUTATIONS IN MT-ASPRS: BIOPHYSICAL PROPERTIES AND IN VIVO CHARACTERIZATION 126

3.1 BIOPHYSICAL CHARACTERIZATION OF MUTATIONS IN MT-ASPRS 126

3.2 ARTICLE # 3 132

3.3 IN VIVO CHARACTERIZATION OF MUTANTS OF MT-ASPRS 133

4 ACHIEVEMENTS AT A GLANCE 140

5 GRAPHICAL SUMMARY 140

CONCLUSION

1 SUMMARY AT A GLANCE 141

2 EVALUATION OF THE OBTAINED RESULTS 143

2.1 RELEASED SELECTIVE PRESSURE ON A STRUCTURAL DOMAIN GIVES NEW INSIGHTS ON THE FUNCTIONAL RELAXATION OF MITOCHONDRIAL ASPARTYL-TRNA SYNTHETASE 143

2.2 SUB-MITOCHONDRIAL DISTRIBUTION OF MT-AARSS SUPPORT THE IDEA OF A MEMBRANE LOCALIZED TRANSLATION MACHINERY 145

2.3 INVESTIGATION OF NATURALLY MATURED MITOCHONDRIAL PROTEINS 147

2.4 THE INTEGRATION OF MT-AARSS INTO FUNCTIONAL NETWORKS REMAINS OPEN 150

2.5 CHARACTERIZATION OF PATHOLOGY-RELATED MUTATION 151

3 LIMITATION AND PERSPECTIVES 154

3.1 TECHNICAL AND SCIENTIFIC LIMITATIONS 154

4 RECOMMENDATION FOR FUTURE SHORT- AND LONG-TERM RESEARCH DIRECTIONS 157

4.1 DECIPHERING THE SUB-MITOCHONDRIAL ORGANIZATION 157

4.2 DECIPHERING NEW FUNCTIONAL NETWORKS 158

4.3 GENERAL SUGGESTIONS FOR FURTHER RESEARCH DIRECTION: 158

MATERIALS AND METHODS

MATERIALS 160

1 CHEMICAL PRODUCTS AND MATERIALS 160

2 ENZYMES 161

3 BIOLOGICAL MATERIAL 161

4 PRIMERS 162

5 LIST OF USED ANTIBODIES 163

METHODS 164

1 MAMMALIAN CELL CULTURE 164

1.1 HUMAN CELL LINES 164

1.2 OTHER CELL LINES 165

2 BHK21/VACCINIA VIRUS EXPRESSION SYSTEM 166

2.1 VECTOR 166

2.2 VIRUS PRODUCTION 167
2.3 PROTEIN PRODUCTION 168
3 IN VIVO KNOCK DOWN OF MRNA 168
3.1 STABLE CELL LINES EXPRESSING SHRNA AGAINST SPECIFIC MRNAS 168
3.2 CELL PROLIFERATION ASSAY 169
4 SUB-CELLULAR FRACTIONATION 169
4.1 PREPARATION OF MITOCHONDRIA 169
4.1 PREPARATION OF MITOPLAST 170
4.2 SUB-MITOCHONDRIAL FRACTIONATION 171
5 ENRICHMENT OF PROTEINS BY IMMUNO-PRECIPITATION 172
5.1 PRODUCTION OF ANTIBODIES IN RABBITS 172
5.2 PREPARATION OF COVALENTLY-LINKED ANTIBODIES: 172
5.3 PREPARATION OF PROTEIN EXTRACTS 173
5.4 IMMUNO-PRECIPITATION 174
6 DNA MANIPULATION 175
6.1 PREPARATION OF CDNA FROM TOTAL RNA 175
6.2 AMPLIFICATION OF MT-AARS SEQUENCES AND CLONING 175
6.3 SITE DIRECTED MUTAGENESIS 176
7 RNA MANIPULATION 176
7.1 RNA ISOLATION FROM HUMAN CELL LINES 176
7.2 EXTRACTION OF MRNA FROM PURIFIED POLYSOMES 177
7.3 NORTHERN BLOT ANALYSIS 177
8 AMINOACYLATION OF *IN VITRO* TRANSCRIBED TRNAS 179
8.1 *IN VITRO* TRANSCRIPTION OF TRNAS 179
8.2 AMINOACYLATION ASSAY USING IN VITRO TRANSCRIBED TRNAS 180
9 ANALYSIS OF PROTEIN SAMPLES 181
9.1 WESTERN BLOT 181
9.2 NATIVE GEL ELECTROPHORESIS 182
9.3 2D-GELELECTROPHORESIS 183
9.4 MASS SPECTROMETRY ANALYSIS: 183

BIBLIOGRAPHY 186

List of tables

TABLE 1: COMPARISON OF THE CONSTITUTION FROM BACTERIAL RIBOSOME AND MAMMALIAN MT-RIBOSOME. 8

TABLE 2: OVERVIEW ABOUT HUMAN MT-AARS. SYNTHETASES ARE SORTED ACCORDING TO THE TWO MAIN CLASSES. 12

TABLE 3: SUMMARY OF EXPERIMENTALLY DETERMINED AND *IN SILICO* PREDICTED SUB-MITOCHONDRIAL DISTRIBUTION OF AARSS IN MEMBRANE AND MATRIX FRACTIONS. 63

TABLE 4: CLUSTERING OF ALL KNOWN DATA IN REGARD TO THE CLASSIFICATION AFTER THE PI OF THE MT-AARS. 91

TABLE 5: INTERACTOME OF MT-ASPRS. 101

TABLE 7: SUMMARY OF IDENTIFIED PROTEINS BY CO-IP EXPERIMENTS.. 107

TABLE 8: SUMMARY OF IDENTIFIED PROTEINS FROM FOUR DIFFERENT CO-IPS. 109

TABLE 8: KINETIC PARAMETERS OF WILD TYPE AND MUTANT TRNA^{ASP} 119

TABLE 9: SUMMARY OF PREVIOUSLY REPORTED EFFECTS OF MISSENSE MUTATIONS IN MT-ASPRS. 127

TABLE 10: LIST OF ALL USED COMMERCIAL ANTIBODIES 163

List of figures

FIGURE 1: SCHEMATIC ILLUSTRATION OF THE HUMAN CIRCULATED MITOCHONDRIAL GENOME. 6

FIGURE 2: THE MITOCHONDRIAL PROTEIN IMPORT MACHINERY. 15

FIGURE 3: PROTEOLYTIC CLEAVAGES OF MITOCHONDRIAL PRECURSOR PROTEINS AFTER MITOCHONDRIAL IMPORT 17

FIGURE 4: CONSENSUS SEQUENCES FOR THE CLEAVAGE OF MITOCHONDRIAL PROTEINS BY MPP AND MIP. 19

FIGURE 5: FROM THE ENCODING OF MITOCHONDRIAL AMINOACYL-TRNA SYNTHETASES TO THE MITOCHONDRIAL ATP SYNTHESIS. 21

FIGURE 6: ORGANIZATION OF THE HUMAN CYTOSOLIC MULTI-AARS COMPLEX (MARS) 25

FIGURE 7: GENERAL CONCEPTS TO OF CYTOSOLIC AARSS TO GAIN NEW FUNCTIONS. 26

FIGURE 8: HUMAN MITOCHONDRIAL PROTEINS SEPARATED BY 2D-GEL ELECTROPHORESIS. 38

FIGURE 9: SCHEMATIC ILLUSTRATION OF THE HYBRIDIZATION SITES OF SIRNA. 41

FIGURE 10: VERIFICATION OF SIRNA-MEDIATED KNOCKDOWN OF MRNA ENCODING MT-ASPRS AND MT-ASPRSΔEXON13 IN HEK293T CELL BY RT-PCR. 41

FIGURE 11: GROWTH CURVES OF HEK293T CELLS TRANSIENTLY TRANSFECTED WITH SIRNAS 43

FIGURE 13: SHRNA-MEDIATED KNOCK DOWN OF MT-ASPRS AND MT-ASPRSΔEXON13.. 45

FIGURE 14: PROLIFERATION RATE OF CELLS EXPRESSING SHRNA-FL AND SHRNA-Δ13 OVER 6 DAYS. 46

FIGURE 15: SCHEME OF THE POSSIBLE DISTRIBUTION OF AARS PROTEINS WITHIN THE MITOCHONDRIAL MATRIX OR ALONG THE INNER MEMBRANE. 50

FIGURE 16: FLOWCHART OF MITOCHONDRIA PURIFICATION AND FRACTIONATION. 52

FIGURE 17: QUALITY CONTROL OF MITOCHONDRIAL FRACTIONS 53

FIGURE 18: TRANSMEMBRANE HELICES PREDICTION USING THE TMHMM SERVER V 2.0 55

FIGURE 19: SENSITIVITY OF THE MEMBRANE BOUND MT-ASPRS TO DIFFERENT SALT CONCENTRATION AND CHAOTROPIC AGENT.. 57

FIGURE 20: DIGITONIN TREATMENT OF MITOCHONDRIAL MEMBRANE. 58

FIGURE 21: MITOPLAST PREPARATION USING INCREASING AMOUNT OF DIGITONIN. 60

FIGURE 22: MITOPLAST PREPARATION USING OSMOTIC SHOCK. 61

FIGURE 23: SCHEMATIZED DISTRIBUTION OF MT-AARSS IN MATRIX AND/OR ALONG THE MEMBRANE OF MITOCHONDRIA.. 65

FIGURE 24: DISTRIBUTION OF TRNAS INSIDE MITOCHONDRIA. 67

FIGURE 25: QUANTIFICATION OF TRNA AMOUNTS IN MEMBRANE AND MATRIX FRACTIONS. 68

FIGURE 26: PROTEOMIC ANALYSIS OF THE SUB-MITOCHONDRIAL DISTRIBUTION OF AARSS. 70

FIGURE 27: VENN DIAGRAM OF IDENTIFIED PROTEINS BY LC-MS/MS IN DIFFERENT SUB MITOCHONDRIAL FRACTIONS.. 71

FIGURE 28: MITOCHONDRIAL ANNOTATION OF IDENTIFIED PROTEINS 72

FIGURE 29: CLUSTERING OF PROTEINS IDENTIFIED IN MATRIX, PERIPHERAL AND INNER MEMBRANE FRACTION IN REGARD TO THEIR SUB-MITOCHONDRIAL LOCALIZATION. 74

FIGURE 30 FUNCTIONAL CLUSTERING OF PROTEINS FOUND IN DIFFERENT SUB-MITOCHONDRIAL FRACTIONS. 75

FIGURE 31: PREDICTED PIS AS A FUNCTION OF THE SEQUENCE SIZE OF MT-AARSS. 82

FIGURE 32: RELATIONSHIP OF PI AND THE MOLECULAR WEIGHT (MW) OF THE MT-AARSS. 84

FIGURE 33: SCHEMATIZED CLASSIFICATION OF AARSS ACCORDING TO THEIR PREDICTED NET CHARGE. 85

FIGURE 34: SCHEMATIZED N-TERMINAL CLEAVAGE SITES OF THE MT-ASPRS UPON IMPORT INTO MITOCHONDRIA. 87

FIGURE 35: EXPRESSION OF PROTEINS WITH MUTATED MTS AND DETECTION IN SUB-MITOCHONDRIAL FRACTION. 89

FIGURE 36A: STRING-PREDICTED FUNCTIONAL INTERACTION 96

FIGURE 37: IDENTIFICATION OF THE OLIGOMERIC STATUS OF THE NATIVE AND RECOMBINANT MT-ASPRS BY BN-PAGE. 103

FIGURE 38: CROSS-LINKING OF MITOCHONDRIAL PROTEINS AND SUBSEQUENT SEPARATION ON BN-PAGE. 104

FIGURE 39: CROSS-LINKING OF MITOCHONDRIAL PROTEINS AND SUBSEQUENT FRACTIONATION OF MATRIX AND MEMBRANE FRACTIONS. 105

FIGURE 40: SCHEMATIZED FLOWCHART OF CO-IMMUNO PRECIPITATION (CO-IP). 106

FIGURE 41: MASCOT SCORES OF THE IDENTIFIED PROTEINS. 112

FIGURE 42: CONSTRUCTION OF THE $TRNA^{ASP}$ TRANSCRIPT. 118

FIGURE 43: FAMILY PEDIGREE OF THE AFFECTED FAMILY. 120

FIGURE 44: ANALYSIS OF ASPARAGINYL-$TRNA^{ASN}$ BY NORTHERN BLOT HYBRIDIZATION. 121

FIGURE 45: ANALYSIS OF ASPARAGINYL-TRNA^{ASN} BY NORTHERN BLOT HYBRIDIZATION..... 122

FIGURE 46: QUANTIFICATION OF ASPARAGINYL-TRNA^{ASN} NORTHERN BLOT HYBRIDIZATION..... 123

FIGURE 47: ANALYSIS OF MODIFIED TRNAS BY NORTHERN BLOT HYBRIDIZATION..... 125

FIGURE 48: LOCALIZATION OF MUTATIONS WITHIN THE CRYSTALLOGRAPHIC STRUCTURE OF HUMAN MT-ASPRS..... 129

FIGURE 49: SOLUBILITY TEST OF RECOMBINANT EXPRESSED MUTANTS..... 131

FIGURE 50: SUB-MITOCHONDRIAL LOCALIZATION OF MUTANT ASPRSS..... 137

FIGURE 51: REPRESENTATION OF THE EXPRESSION VECTOR CONSTRUCT AND CORRESPONDING RESTRICTION SITES NDEI, NSII AND NCOI FLANKING THE ORF FOR MT-ASPRS..... 167

FIGURE 52: PERCOL-SUCROSE GRADIENT. MITOCHONDRIAL FRACTION APPEARS AS A YELLOW RING IN TOP AREA OF THE GRADIENT..... 170

List of Publications

Book Chapter #1 *22*

Translation in mammalian mitochondria: Order and disorder linked to tRNAs and aminoacyl-tRNA synthetases *(in press)*

Book Chapter #2 *23*

Pathogenic Implications of Human Mitochondrial Aminoacyl-tRNA Synthetases *(published 2013)*

Article #1 *36*

Released selective pressure on a structural domain gives new insights on the functional relaxation of mitochondrial aspartyl-tRNA synthetase *(submitted)*

Article #2 *81*

Expression system for naturally maturated mammalian proteins imported into mitochondria *(submitted)*

Article #3 *132*

Comparative biophysical investigations of pathogenic mutants of human mitochondrial aspartyl-tRNA synthetase *(in preperation)*

Article #4 136

A human pathology-related mutation prevents import of an aminoacyl-tRNA synthetase into mitochondria *(published 2011)*

INTRODUCTION

Introduction

1 Mitochondria at a glance

Compartmentalization is one major property of a eukaryotic cell. Compartments (also called organelles) form a membrane-enclosed lumen, which ensures the many different and highly specialized processes that can occur in parallel inside a cell. Major intracellular compartments in eukaryotic cells include e.g. the nucleus, mitochondria, chloroplasts (solely in plants) and lysosomes, and occupy about half of the intracellular space. Each organelle has a separate specialized set of proteins responsible for their unique functions. Mitochondria are one of these highly specialized organelles. They are enclosed by double membrane and serve as the major source of ATP for the cell (Alberts et al., 2002).

The mitochondrial ***origin*** is nowadays well explained by the endosymbiont hypothesis. The mitochondrion is hypothesized to originate from a single α-Proteobacterial lineage and to have undergone large genomic reduction(s) compared to the progenitor genome. A massive transfer of genes from the endosymbiont to the nuclear host genome has also been postulated. Surprisingly, proteomic analyses revealed that less than 20% of mitochondrial proteins originate clearly from α-Proteobacterial proteome, more than 50% have neither α-Proteobacterial homolog nor necessarily a bacterial one, and approximately 30% have even neither bacterial nor archaeal homolog (reviewed in e.g. Gray, 2012).

The **s*tructure and morphology*** of mitochondria can be observed under an electron microscope and are characterized by outer and inner membranes and cristea. The latter one is a characteristic inversion of the mitochondrial inner membrane, which increases significantly the membrane surface. The mitochondrial lumen is called matrix. It is assumed that the morphology of mitochondria is highly variable and changes with the

functional requirements. For example, the size of the inner mitochondrial surface area has been correlated with the energy demand of the cells (Scheffler, 2008).

Fusion and fission of mitochondria are processes that occur under normal cellular conditions. Fusion is when the outer membrane and subsequently the inner membrane of two mitochondria fuse in a coordinated manner, maintaining the integrity of the two membranes, the inter membrane space and the matrix. Fission is the coordinated division of one mitochondrion into two mitochondria (review in e.g. Perfettini et al., 2005). Mitochondrial fusion and fission typically counterbalance each other. The distribution of the mitochondria is highly variable between different organisms and even cell types. Unicellular organism have typically one mitochondria while human liver cells can have up to 2 000 mitochondria per cell (Alberts et al., 2002).

The ***biogenesis*** of mitochondria requires well-coordinated processes of transcription and translation from two separate genomes. Mitochondria and chloroplast (in plants) have, beside the nucleus, their own genomes and corresponding transcription/translation machineries. The mitochondrial genome differs in size from approximately 6kb (*Plasmodium falciparum*) to up to more than 69kb (*Reclinomonas Americana*) (Gray, 2012). The human mitochondrial genome is composed of 16 568bp. It codes for 37 genes comprising, 13 messenger RNAs (mRNAs; corresponding proteins being all subunits of the respiratory chain complexes), 2 ribosomal RNAs (rRNAs) and 22 transfer RNAs (tRNAs) (for complete sequence, see www.mitomap.com). Since the mitochondrial proteome has been estimated to contain ~1500 proteins (Meisinger et al., 2008), the vast majority of mitochondrial proteins are encoded in the nucleus, produced in the cytosol and imported into the mitochondria.

Mitochondrial ***metabolism and function*** is mainly known for the production of the cellular free energy in the form of ATP, by a process named oxidative phosphorylation (OXPHOS). A human adult has an average ATP turnover of approximately 200kg per day; despite this, only a few grams are available at a given time (Brown et al., 2006). OXPHOS takes place in five large multi-subunits complexes named respiratory chain complexes and is located in the mitochondrial inner membrane. Beside this prominent process, mitochondria take part in other fundamental metabolic functions such as the Krebs cycle,

fatty acid metabolism, urea cycle, and iron or calcium metabolism. Finally, mitochondria also play a role in other cellular functions in e.g. cell signaling, apoptosis or heat production (reviewed in e.g.Scheffler, 2008).

Mitochondrial ***dysfunctions*** are involved in a wide and diverse spectrum of human disorders such as cancer, Alzheimer, Parkinson, muscular dystrophy, neurodegeneration and mitochondrial myopathies. Classical 'mitochondrial diseases' are defined as a clinically heterogeneous group of disorders that are characterized by a dysfunction of the OXPHOS pathway (Chinnery, 1993). Primary mitochondrial defects are those affecting directly the respiratory chain and due to mutations in proteins of the OXPHOS pathway (either nucleus- or mitochondrial- encoded). Secondary mitochondrial defects result from impaired mitochondrial metabolism affecting indirectly OXPHOS (reviewed in e.g. Chinnery, 1993 or Schapira, 2002). In both primary and secondary mitochondrial defects, a deficiency in ATP production is thought to be the dominant cause of disease.

Thus, Mitochondria are unique organelles in a eukaryotic cells forming an own ***compartment*** of α-Proteobacterial ***origin*** with a characteristic ***structure and morphology***, regulated ***fusion and fission*** processes, own ***biogenesis*** machinery, essential ***functions*** and pathology-related ***dysfunctions***.

The present thesis will focus on a family of proteins, the aminoacyl-tRNA synthetases (aaRSs), implicated in the **biogenesis** (translation) of mitochondrial-encoded proteins. All Members of this protein family are nucleus-encoded and transported from the cytosolic to the mitochondrial **compartment**, where they are integrated into the mitochondrial translation machinery. The ***origin*** of this protein family is discussed in regard to its evolutionary history. Furthermore, we integrate the aaRSs in the ***structure and morphology*** of mitochondria (sub-mitochondrial localization) and new auxiliary ***functions*** will be discussed. All this knowledge shall help to integrate this protein family in the emerging field of mitochondrial **dysfunction**.

The following introduction focuses on the involvement of the aaRSs in the mitochondrial translation of proteins and the consequences of reported pathology-related

mutations. The introduction is divided into four main parts. In the first part, the **mitochondrial translation machinery** will be presented with an emphasis on the aaRSs. Due to the dual origin of mitochondrial proteins, we will center our attention in the second part on the nucleus-encoded proteins and their **import** into the mitochondria. In the third part, the impact of **pathology-related mutations** in mt-aaRSs will be analyzed. Finally, **non-translational functions** of cytosolic aaRSs will be recalled in the fourth part. The main objectives of this thesis will then be defined in line with the emerging needs for further investigations.

Of note, I contributed during my PhD to two book chapters. As such, these works form a center part of this introduction. For sake of better understanding, some aspects discussed in the book chapters will be considered as well in other parts of the introduction.

2 The mammalian mitochondrial translation machinery

Mitochondria have their own translation machinery to decode their mitochondrial genome. This machinery is distinct from the cytosolic one and possesses its own set of key actors needed for correct translation. Despite a bacterial origin (for most of them), molecules from the human/mammalian mitochondrial translation machinery have some peculiarities when compared to the bacterial ancestors or to cytosolic homologues. One key family of molecules of translation, the aminoacyl-tRNA synthetases, exemplifies this peculiarity and will be the focus of the following paragraphs.

2.1 Human mitochondrial genome

Human mitochondria contain their own small circular DNA genome (mt-DNA) with a size of 16 568bp (**Figure 1**) (Anderson et al., 1981). This genome codes for 22 tRNAs, 2 rRNAs and 13 proteins, all subunits of the respiratory chain complexes. Most of the protein coding genes are separated by one or several tRNA genes, illustrating not only that tRNA are highly dispersed over the mitochondrial genome, but also that they serve as punctuation between the genes.

It has been shown that several copies of mammalian mt-DNA are organized into a dynamic mitochondrial network with discrete protein–DNA complexes, named nucleoids. Those are shown to be membrane associated. Also, the amount of nucleoids and the number of mt-DNA copy vary and depend on the cell type, the developmental stage, or on the species (reviewed in Spelbrink, 2010).

Mitochondrial DNA replication is not fully understood but in one longstanding hypothesis it is bidirectional and follows a “strand-displacement” model involving two origins of replication (O_H and O_L) and a handful of proteins. In contrast, more recent work supports a coupled model of mtDNA replication involving RNA incorporation (reviewed in Holt and Reyes, 2012). Of note, the H and L characters make references to the “heavy” and

"light" DNA strands obtained after the separation of mt-DNA on a cesium chloride gradient reflecting the biased distribution of GC within the two strands.

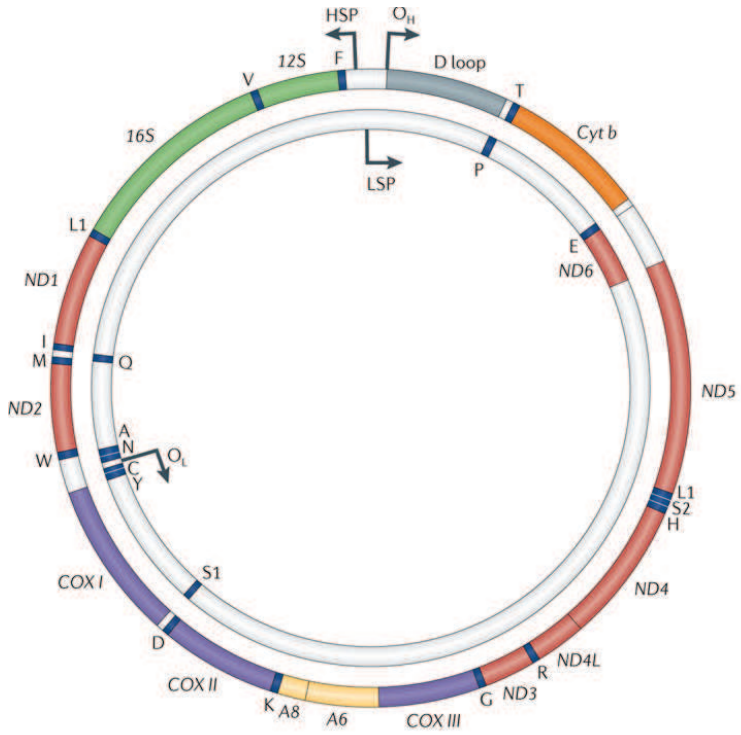


Figure 1: Schematic illustration of the human circulated mitochondrial genome. Encoded genes are represented, tRNA are represented by single letter code. Start of replication initiation is indicated by O_H and O_L.Transcirption promotor sites are indicated by HSP (heavy strand promotor) and LSP (light strand promotor) (Schon et al., 2012).

The transcription of the mt-genome occurs in both directions resulting in three (two for the heavy strand and one from the light strand) polycistronic transcripts. Within those primary transcripts, tRNAs flank the rRNAs and the coding mRNA (reviewed in Asin-Cayuela and Gustafsson, 2007). Thus, following the tRNA punctuation model, the endonucleolytic excision of tRNAs guide the correct processing release of mRNA and rRNA with in most cases proper ends (Ojala et al., 1981). 5'- and 3'- endonucleolytic excisions of tRNAs are performed by RNase P and RNase Z, respectively. However, functional tRNAs are obtained in mammalian mitochondria only after the CCA-addition at the 3'-ends (reviewed e.g. Levinger et al., 2004). In addition, several posttranscriptional

modifications are added to increase the tRNA stability or guarantee the proper function (Suzuki et al., 2011). Since mt-mRNA does not contain introns, the further processing is restricted to constitutive polyadenylation, which create UAA stop at least in some animal mt-mRNAs lacking a stop codon and protecting mt-mRNA from exonuclease activity (Nagaike et al., 2008 and Temperley et al., 2010).

After transcription, the pathways followed by the coding and the non-coding RNAs are different. The rRNAs contribute to the small and large subunits of the mitochondrial ribosome. The mt-tRNAs are enzymatically loaded with the corresponding amino acid by the cognate aaRS. Furthermore, the aminoacyl-tRNAs serve as adaptors for mRNA decoding and protein synthesis at the level of the mitochondrial ribosome. This process is called mitochondrial translation and will be discussed in more detail.

2.2 Key actors of mitochondrial translation machinery

The mitochondrial ribosome is the central player in the mitochondrial translation machinery. Despite its endosymbiotic origin, the size and the composition of mitochondrial ribosome differ from those of the bacterial ancestor (**Table 1**). The mitochondrial ribosome has a smaller fraction of RNAs and a higher amount of proteins. The RNA:protein ratio evolved from 2:1 in bacteria to 1:2 in the mammalian mitochondria (Agrawal and Sharma, 2012). Interestingly, structural data indicates that new proteins do not replace the missing RNA content, but instead that additional proteins occupy new positions in both subunits of the mt-ribosome (Sharma et al., 2003).

Ribosome source	**Bacteria (*E.coli*)**	**Mammalian mitochondria (*B.taurus*)**
Molecular mass	2.3 MDa	2.7 MDa
Sedimentation coefficient	70S	55S
RNA:protein ratio	2:1	1:2
Subunits	30S + 50S	28S + 39S

Small subunit composition	16S rRNA (1542 nt) + 21 proteins	12S rRNA (950 nt) + 29 proteins
Large subunit composition	23S rRNA and 5S rRNA (total 3024 nt) + 34 proteins	16S rRNA (1560 nt) + 50 proteins
Diameter	~260 Å	~320 Å

Table 1: Comparison of the constitution from bacterial ribosome and mammalian mt-ribosome. Table adapted from (Agrawal and Sharma, 2012).

In addition, data obtained from mitochondrial fractionation suggest that the mt-ribosome is localized close to the mitochondrial inner membrane (Liu and Spremulli, 2000) and the nascent polypeptide chain is co-translationally released into the mitochondrial inner membrane(Gruschke and Ott, 2010).

So far, two mitochondrial initiation factors (IF-2_{mt} and IF-3_{mt}) and three elongation factors (EF-Tu_{mt}, EF-ts_{mt} and EF-G_{mt}) have been characterized (reviewed in Scheffler, 2008). Briefly, IF-2mt interacts with the initiator Met-$tRNA^{Met}$ and the ribosome. IF-3mt has been shown to support the dissociation of the mt-ribosome subunits (Koc and Spremulli, 2002). EF-Tu_{mt}, EF-Ts_{mt} and EF-G_{mt} have been shown to be able to replace their bacterial homologues, suggesting similar activities. EF-Tu_{mt} binds specifically the aminoacyl-tRNAs and guides them from the cognate aaRSs to the ribosome. EF-Ts_{mt} supports the release of the aminoacyl-tRNAs by promoting the GDP exchange with GTP on EF-Tu_{mt}. EF-G_{mt} catalyzes the translocation process from the P-site to the E-site on the active core of the ribosome (Cai et al., 2000).

The aminoacyl-tRNA synthetases catalyze the important step of tRNA aminoacylation and provide aminoacyl-tRNAs as substrates for the protein biosynthesis. A further detailed description of those key molecules of translation will be given below and within the two book chapters prepared during my PhD and included in this introduction.

2.3 Peculiar aspects of mitochondrial translation

As already mentioned, translation is a process, in which the ribosome and several cofactors (e.g. initiation or elongation factors) facilitate the correct binding of the anticodon of the aminoacyl-tRNA to the complimentary codon region of the mRNA. Consequently, the ribosome helps to form the peptide bond between the amino acid and the nascent peptide. Thus, the loading of each tRNA by the correct amino acid is a crucial step for a correct decoding of the mRNA. This reaction is catalyzed by the aaRSs and governed by what has been defined as the 'second genetic code' (De Duve, 1988). Several features distinguish the mitochondrial translation system of metazoans (animals) from those of other translation systems but also of mitochondria of other species.

The most obvious peculiarity of the mitochondrial translation machinery is the dual origin of its constitutive molecules. Altogether more than 80 different factors are involved in the mitochondrial translation but only the set of tRNAs and two of the rRNAs are encoded by the mitochondrial genome All other factors (e.g. translation factors, proteins from the ribosome, aaRSs) are encoded in the nucleus, translated in the cytosol and imported inside mitochondria. Proteins of the mitochondrial translation machinery are encoded by a distinct set of genes (with some exceptions) and thus differ from those of cytosolic localization.

Another outstanding peculiarity is that the genetic code is not conserved in mitochondria. As examples, the "universal" termination codon UGA codes in mammalian mitochondria a Trp, and the "universal" Arg codons AGA and AGG are used as termination codons in mitochondria (Watanabe, 2010). Beside these codon variations, significant reduction of the amount of ncRNAs has been observed (Anderson et al., 1981; Matthews et al., 1982). The most striking example concerns the set of tRNAs, reduced to a minimal number of 22 in mammalian mitochondria. Most of them have sequence and structural peculiarities and a tendency to a shortening that led scientists to designate them as 'bizarre' (Florentz et al., 2003). However, specific post-transcriptional modifications have arisen in animals that strengthen the tRNA structure (e.g. methylation at the A9 position of $tRNA^{Lys}$; (Helm et al., 1999) or the codon-anticodon reading (e.g. 7-

methylguanosine at wobble position of $tRNA^{SER}_{GCU}$; reviewed in e.g. Yokobori et al., 2001; Watanabe, 2010) Other discrepancies are demonstrated concerning the functional properties of mitochondrial aaRSs. Those include reduced catalytic efficiencies (Bonnefond et al., 2005) or increased intrinsic structural plasticities (Klipcan et al., 2012; Neuenfeldt et al., 2013) for mt-aaRSs when compared to bacterial homologues. Unilateral recognition relationships have been described at least for some of the key factors of translation. For example, the ancestral bacterial aaRSs are not able to cross-aminoacylate the corresponding mitochondrial tRNAs (Fender et al., 2012). Exhaustive lists of discrepancies concerning the functional properties of mitochondrial aaRSs are described in both integrated book chapters.

In general, it has been postulated that the conservation of a functional translation in mitochondria requires a continuous adaptation of proteins in response to reduced or degenerated RNAs. This is illustrated by a gain of plasticity (Klipcan et al., 2012; Neuenfeldt et al., 2013) or an enlarged protein composition (Matthews et al., 1982). It has been hypothesized that those unique features of the mt-translation system that have arisen from the endosymbiont (reviewed in Watanabe 2010) reflect a unique mitochondria/nucleus co-evolution and may result from a faster evolutionary rate of mitochondrial genome compared to the nuclear one (Castellana et al., 2011).

2.4 Aminoacyl-tRNA synthetases are key molecules of translation

Aminoacyl-tRNA Synthetases (aaRSs), among the above described other actors, are some of the key actors of the translation machinery. Those housekeeping enzymes interact specifically with cognate tRNAs and esterify them with corresponding amino acids through a reaction named aminoacylation. Aminoacylated-tRNAs are then carried by elongation factors (e.g. Ef-Tu) to the ribosome where the amino acid is incorporated into the nascent peptide chain. The aminoacylation reaction involves a two-step process including first the activation of the specific amino acid into an adenylate in the presence of ATP; and second, the specific recognition of the cognate tRNA followed by transfer of the activated amino

acid (Ibba et al., 2005). In addition to their classical role in translation, unexpected alternative functions have been reported for many of the cytosolic aaRSs over the last decade (will be discussed in a separated part below).

The human nuclear genome encodes two distinct sets of aaRSs. One set of aaRS is specific for cytosolic translation and one set for mitochondrial translation. The two differ in their sequences, and mt-aaRS harbor in addition a mitochondrial targeting sequence (MTS)**.** The nuclear gene annotation of the full set of human mt-aaRSs showed that 17 of the 19 mt-aaRSs are encoded by a distinct set of genes compared to the cytosolic counterparts (Bonnefond et al., 2005). The two exceptions concern the GlyRS and LysRS. GlyRSs are generated from two translation initiation sites on the same gene (Mudge et al., 1998; Shiba et al., 1994) and LysRSs follow an alternative mRNA splicing pathway (Tolkunova et al., 2000). Both possibilities generate proteins with and without MTS, and no gene for mt-GlnRS has been found so far.

In line with the dual origin of the mitochondrial translation machinery, the evolutionary history of the mt-aaRSs has been investigated. Despite the conventional view of the endosymbiotic origin of mitochondria (Gray, 2012) the source of nuclear gene for mt-addressed aaRSs is diverse, reflecting numerous post-endosymbiotic and/or lateral gene transfer events (Brindefalk et al., 2007). Nevertheless, many of the mt-aaRSs originate from the bacterial domain. This is the case for the human mt-aspartyl-tRNA synthetase (mt-AspRS) (Bonnefond et al., 2005).

A first comprehensive analysis of mt-aaRS was performed in 2005 with the annotation of the full set of mt-aaRSs and functional characterization of the mt-AspRS and mt-TyrRS (Bonnefond et al., 2005). So far, experimental characterization have been performed only for 9 additional enzymes, i.e. mt-PheRS, mt-SerRS, mt-IleRS, mt-LeuRS, mt-LysRS, mt-TrpRS, mt-MetRS, mt-GluRS and mt-HisRS (**Table 2**). As for cytosolic aaRSs, mt-aaRSs are grouped into the two classes of aaRSs, as originally defined (Cusack et al., 1990; Eriani et al., 1990). This classification is based on signature motifs being respectively HIGH and KMSK for class I enzymes and motifs 1, 2 and 3 for class II enzymes.

	Gene	Protein	Protein length in aa	Predicted cleavage site	Kinetic parameters	Related publications*
Class I	RARS2	mt-ArgRS	578	after 16	no	
	CARS2	mt-CysRS	564	after 62	no	
	EARS2	mt-GluRS	523	after 25	yes	(Nagao et al., 2009)
	IARS2	mt-IleRS	993	after 61	yes	(Kelley et al., 2000)
	LARS2	mt-LeuRS	903	39 aa	yes	(Bullard et al., 2000) (Sohm et al., 2003) (Wittenhagen et al., 2003)
	MARS2	mt-MetRS	593	18 aa	no	
	WARS2	mt-TrpRS	360	18 aa	yes	(Jorgensen et al., 2000)
	YARS2	mt-TyrRS	477	after 16	yes	(Bonnefond et al., 2005) (Bonnefond et al., 2007)
	VARS2	mt-ValRS	993	after 29	no	
Class II	AARS2	mt-AlaRS	985	after 68	no	
	NARS2	mt-AsnRS	477	after 14	no	
	DARS2	mt-AspRS	645	after 47	yes	(Bonnefond et al., 2005) (Fender et al., 2012) (Messmer et al., 2009) (Neuenfeldt et al., 2013) (Gaudry et al., 2012)
	GARS	mt-GlyRS	739	54 aa	no	
	HARS2	mt-HisRS	506	after 34	yes	(Pierce et al., 2011)
	KARS	mt-LysRS	625	16 aa	yes	(Dias et al., 2012)
	FARS2	mt-PheRS	451	37 aa	yes	(Bullard et al., 1999) (Levin et al., 2007) (Klipcan et al., 2008) (Klipcan et al., 2012)
	PARS2	mt-ProRS	475	after 47	no	
	SARS2	mt-SerRS	518	after 20	yes	(Yokogawa et al., 2000)
	TARS2	mt-ThrRS	718	after 39	no	

Table 2: Overview about human mt-aaRS. Synthetases are sorted according to the two main classes. Cleavage site indicated as « after xx » was predicted. Cleavage site indicated by « Xaa » is experimental obtained. References in right column refer to studies in which kinetic parameters are determined. *Given literature refer to kinetic analysis. This Table is adapted from Bonnefond et al., 2005.

Nowadays, the crystal structures of the human mt-TyrRS, mt-PheRS and mt-AspRS (Bonnefond et al., 2007, Klipcan et al., 2012, Neuenfeldt et al., 2013) and of the bovine mitochondrial mt-SerRS (Chimnaronk et al., 2005) have been solved. It has been observed

distinguish themselves by more electropositive surface potentials (especially along the tRNA binding surface) or by enlarged grooves for tRNA accommodation, displaying an adaption to degenerated mt-tRNA (Neuenfeldt et al., 2013). Today, restrictions in solving the crystal structure of mt-aaRSs are mainly due to the difficulty in producing large amounts of stable mitochondrial proteins with a well-defined N-terminus leading to soluble proteins (further discussed in a following paragraph).

More information in regard to structural, biophysical and functional peculiarities of the mammalian mitochondrial tRNAs and aaRSs, and of their partnership in their wild-type state will be given in the book chapter #1. Structural, biophysical and functional peculiarities about the human mitochondrial translation system, with a special intention on the human aaRSs, are discussed in the book chapter #2.

3 Import of nucleus-encoded mitochondrial proteins

3.1 The mitochondrial proteome

Mt-aaRSs belong to the numerous nucleus-encoded proteins, which are translated in the cytosol and imported into the mitochondria, thanks to the presence of MTS. Several research groups have analyzed the total mammalian mitochondrial proteome. It has been estimated to comprise more than 1500 proteins (Meisinger et al., 2008). With approximately 1100 mitochondrial proteins, the MitoCarta database is currently the largest collection of mammalian mitochondrial genes and corresponding expressed proteins of assumed mitochondrial localization (Pagliarini et al., 2008). The largest compendium of distinct human mitochondrial proteins, identified by mass spectrometric analysis, was obtained from highly purified human heart mitochondria (Taylor et al., 2003). This initial set of protein sequences together with sequences derived from public sequence databases is integrated within the MitoProteome database (Cotter et al., 2004). However, approximately 400 proteins of the estimated mitochondrial proteome have not yet been identified. Indeed,

it seems difficult to define the precise number of mitochondrial proteins, since mitochondria are dynamic organelles, with proteins possessing dual or multiple localizations, and showing different expression patterns depending on cell types or on physiological conditions. It has been postulated that a precise knowledge of mitochondrial proteome will be of help to understand mitochondrial complexity, tissue heterogeneity, mitochondrial evolution, but also mitochondrial diseases (e.g. Calvo and Mootha, 2010; Gray, 2012; Gregersen et al., 2012).

3.2 Mitochondrial Import machinery

3.2.1 Cytosolic expression of mitochondrial precursor protein

Nucleus-encoded mitochondrial proteins are translated as mitochondrial precursor proteins in the cytosol. Co-translational or post-translational mechanisms for the translocation of the nascent precursor protein have been proposed (Fujiki and Verner, 1993; Verner, 1993). Nowadays, a post-translational mechanism for translocation of mitochondrial precursor proteins is widely accepted (Hoogenraad et al., 2002; Schleiff and Becker, 2011; Wiedemann et al., 2004). However, several publications indicate that some mRNA-bound ribosomes are associated with the mitochondria (Kellems et al., 1975; Marc et al., 2002). After translation, cytosolic chaperones guide the precursor protein to Tom20, Tom22 and Tom70 receptors of the translocase of the outer mitochondrial membrane (TOM) complex localized in the outer membrane (Chacinska et al., 2009).

3.2.2 Translocation of the precursor protein through the mitochondrial outer and inner membranes

Precursor proteins are imported through the translocase of the outer mitochondrial membrane (TOM) and, depending on sorting signals, either integrated in the outer membrane, directed to the intermembrane space, or further internalized through the

translocase of the inner membrane (TIM) **Figure 2**. Two general classes of sorting signals are known: (i) the classical cleavable N-terminal localized MTSs (that lead to the translocation of proteins into the matrix, but also for some proteins to the inter-membrane space or to the inner membrane). N-terminal MTSs are characterized by amphipathic α-helix, and are typically cleaved after the import; (ii) Non-cleavable internal sorting signals are found in proteins of the mitochondrial outer membrane, in the majority proteins of inter-membrane space, and in a few proteins of mitochondrial matrix (reviewed in e.g. Schmidt et al., 2010) .

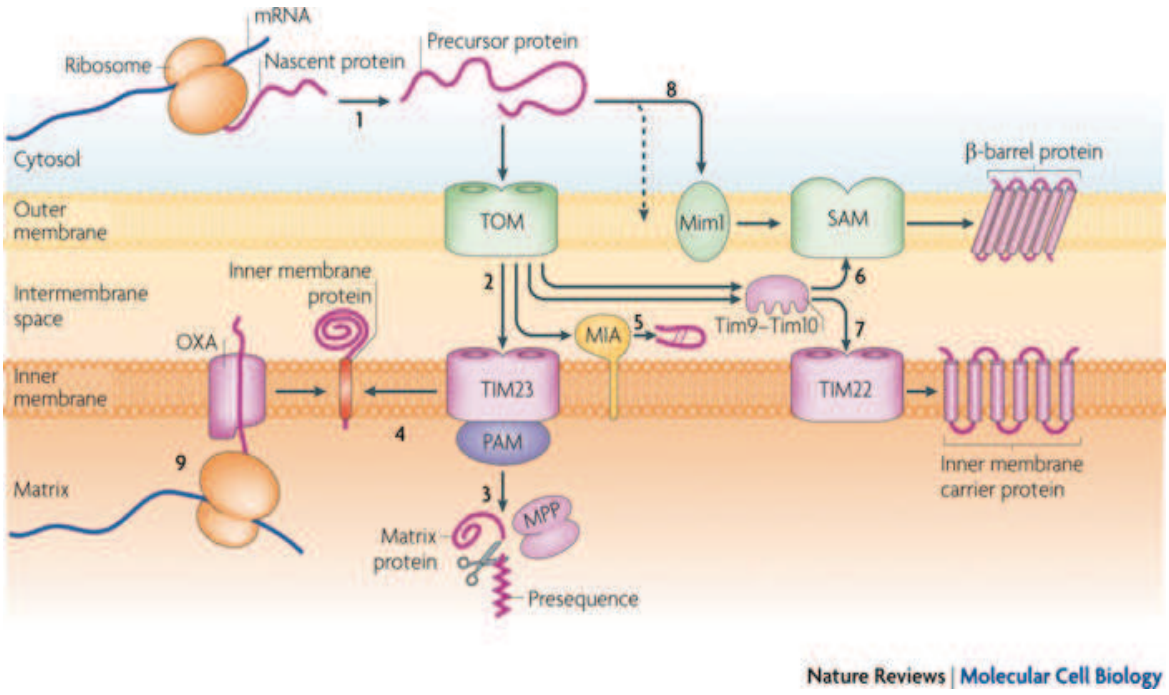


Figure 2: The mitochondrial protein import machinery (Figure taken from Schmidt et al., 2010). Mitochondrial proteins are translated in the cytosole as precursor protein and imported by translocase of the outer mitochondrial membrane (TOM) or integrated directly in the outer membrane by a protein called mitochondrial import 1 (Mim1). The general translocase TOM is the entry gate for nearly all types of mitochondrial proteins in the intermembrane space. Depending on the MTS, proteins are translocated to mitochondrial matrix by the translocase of the inner membrane and presequence translocase-associated motor (TIM23/PAM) complex. Proteins of the outer membrane are integrated in the membrane by the sorting and assembly machinery (SAM). Proteins of the inner membrane are integrated in the membrane by the carrier translocase of the inner membrane (TIM22). Proteins of the intermembrane space are released and folded by the mitochondrial intermembrane spaces assembly (MIA) machinery.

3.2.3 Sorting of mitochondrial proteins

More than five different import pathways to translocate mitochondrial precursor proteins into their predetermined compartment of mitochondria (outer membrane, inter-membrane space, inner membrane, matrix) are reported (reviewed in (Chacinska et al., 2009, Schmidt et al., 2010, Becker et al., 2012). These pathways were studied using mainly fungal model organisms. Mechanisms of mitochondrial import in higher eukaryotic organisms are not fully deciphered, but expected to be of enormous importance for the field of mitochondrial biogenesis (Chacinska et al., 2009).

The outer membrane has three import complexes (TOM, MIM and SAM). The general translocase TOM, which is the entry gate for nearly all types of mitochondrial proteins, translocate the proteins in the inter-membrane space. Depending on the MTS, proteins can be further translocated to mitochondrial matrix through the TIM23/PAM complex, integrated in the outer membrane by the sorting and assembly machinery (SAM) complex, integrated in the inner membrane by the carrier translocase of the inner membrane (TIM22) or released and folded by the mitochondrial inter-membrane spaces assembly (MIA) machinery (Chacinska et al., 2009).

3.2.4 Proteolytic cleavages of the precursor protein

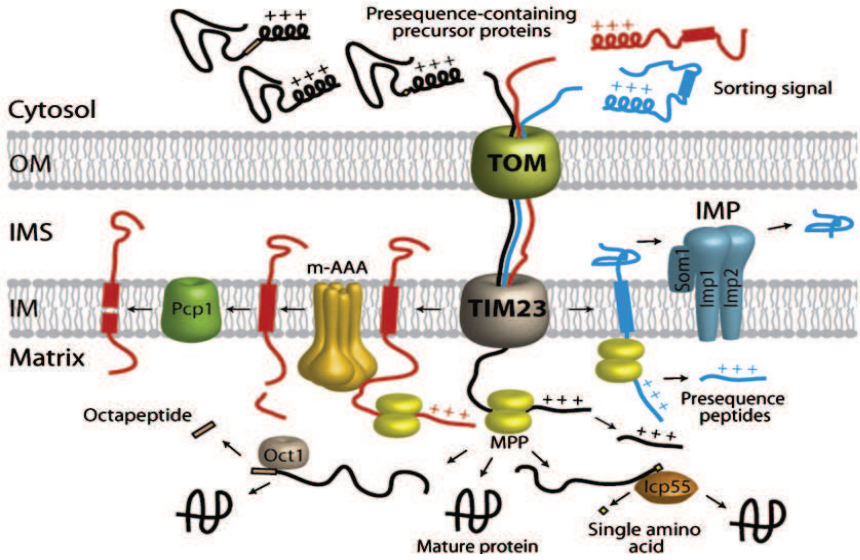


Figure 3: Proteolytic cleavages of mitochondrial precursor proteins after mitochondrial import (Adapted from Mossmann et al., 2012)**.** After internalization of mitochondrial precursor proteins, mitochondrial proteins with a N-terminal MTS are constitutively cleaved by mitochondrial processing peptidase (MPP). Subsequently, proteins can be cleaved by the matrix localized oxtapeptidyl aminopeptidase 1 (Oct1) or the intermediated cleaving peptidase (Icp55). Proteins of mitochondrial inner membrane can be cleaved Pcp1 (processing of cytochrome c peroxidase). The inner membrane peptidase (IMP) cleaves hydrophobic sorting signals of proteins translocated to intermembrane space.

3.2.5 Mitochondrial Processing enzymes

The constitutive proteolytic cleavage of N-terminal MTS occurs after the internalization of the precursor protein into mitochondria and is not yet fully understood. Several peptidases have been identified that cleave with different specificity the N-terminal part. The mitochondrial processing peptidase (MPP) removes the MTS. Different additional cleavages can subsequently occur to remove newly exposed residues such as that performed by the N-terminal distinct octapeptide by Oct1 (octapeptidyl aminopeptidase 1; Gakh et al., 2002) or destabilizing residues by Icp55 (intermediated cleaving peptidases; Vögtle et al., 2009). In addition, the membrane bound inner membrane peptidase (IMP) removes, after MPP cleavage, the newly exposed hydrophobic sorting signal of proteins translocated to

inter-membrane space. The second known membrane bound peptidase Pcp1 (processing of cytochrome c peroxidase) cleaves internal hydrophobic sorting signals. Cleavage motifs in mitochondrial precursor proteins were studied using mainly fungal model organisms.

3.2.6 The N-terminal definition

Several studies investigated in more detail the maturation process of mitochondrial proteins. For example, some studies analyzed important cleavage motifs of the MPP (e.g. reviewed Gakh et al., 2002, Mossmann et al., 2012) whereas others profiled the proteolytic events *in vivo* (e.g. Timmer et al., 2007). In addition, the first comprehensive analysis of the N-terminal proteome of yeast revealed the N-terminal processing site of 615 proteins (Vögtle et al., 2009). It turns out that cleavage site selection follows very loose consensus sequences making predictions difficult (Gakh et al., 2002). However, several *in silico* tools exist to predict the mitochondrial localization and the mitochondrial targeting sequence e.g. TargetP (Emanuelsson et al., 2000), Predotar (Small et al., 2004), Mitoprot (Claros and Vincens, 1996).

In general, MTSs cleaved by MPP have a size of ~20-60 aa, are characterized by overall positive charge, a predicted ability to form amphiphilic α-helices and the presence of argine at position -2 of the cleavage site. Four different cleavage sites motifs have been identified, which contain cleavage signals for MPP and respectively one for Oct1 and Icp55 (reviewed in e.g. Gakh et al., 2002,Mossmann et al., 2012).

Of note, it is assumed that the precursor proteins are translocated through the mitochondrial membrane as unfolded polypeptide chains. This facilitates the correct recognition of the MTS and the physical translocation through the narrow import machinery. Thus, after processing and sub-mitochondrial sorting, a chaperone complex will refold the protein in its active form (Voos, 2013).

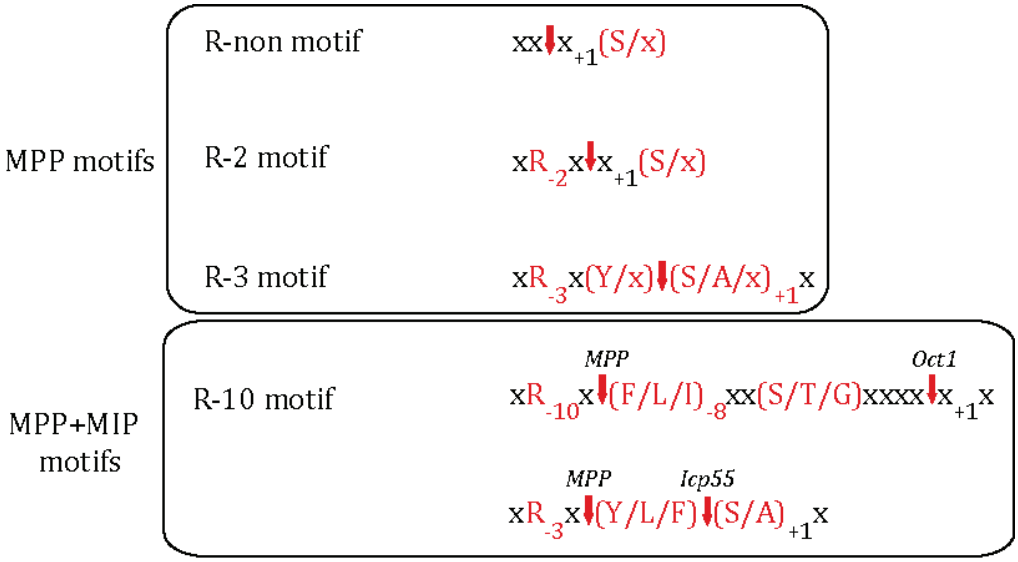


Figure 4: Consensus sequences for the cleavage of mitochondrial proteins by MPP and MIP. Amino acid sequences are displayed. Red residues indicate highly conserved amino acids responsible for the recognition. Red arrows indicate the cleavage position.

4 Human translation system and mitochondrial diseases

4.1 Integration of mt-aaRSs into the mitochondrial genome expression

Mt-aaRSs are encoded in the nucleus, translated in cytosol, targeted to the mitochondria thanks to N-terminal MTSs, as previously predicted (Bonnefond et al., 2005). Thus, it is assumed that mt-aaRSs follow classical import pathway of precursor proteins with a MTS. Upon the import into mitochondria, mt-aaRSs are integrated into the mt-translation machinery and contribute to the translation of the 13 protein sub-units of the mitochondrial respiratory chain complexes. Beside these mitochondrial-encoded sub-units, all other proteins of the respiratory chain are nucleus encoded and imported into mitochondria. Only a correct constitution of the mitochondrial respiratory complex leads to the proper production of ATP. This requires a coordinated expression of nuclear and mitochondrial encoded proteins, but also a coordinate association and functioning of the mitochondrial translation machinery. Above all, this involves efficient partnerships between mt-DNA-encoded RNAs and nucleus-encoded proteins, especially between rRNAs and ribosomal proteins to form active ribosomes and between tRNAs and aaRSs to allow for accurate synthesis of aa-tRNAs. Accordingly, there are key links between the aminoacylation activity of aaRSs in charge of the synthesis of the 13 mt-DNA-encoded proteins, and the activity of respiratory chain complexes. It can be anticipated that any dysfunction of a single macromolecule of the translation machinery may have severe impacts on the activity of the respiratory chain complexes and subsequently lead to mitochondrial defects. The two following book chapters shed light on mitochondrial translation system, highlight order and disorder in the partnership of tRNAs and aaRSs, and summarize the pathogenic implication of the latter one.

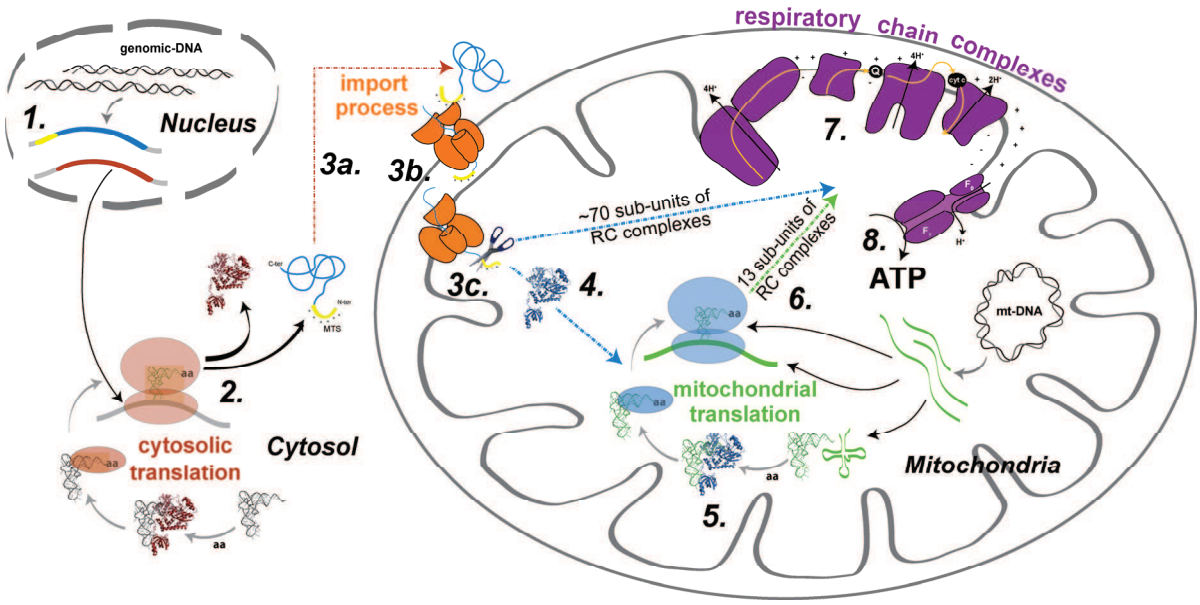


Figure 5: From the encoding of mitochondrial aminoacyl-tRNA synthetases to the mitochondrial ATP synthesis. The route from the place of encoding of mt-aaRSs (the nucleus) to their place of biosynthesis (the cytosol) and their place of use (the mitochondria) is schematized.Mt-aaRSs biogenesis comprises mRNAs expression and processing (1), mt-aaRSs synthesis 2), import process into mitochondria (3a addressing; 3b translocation; 3c processing), and proper folding, oligomerization, and stability upon entry to mitochondria (4). Mt-aaRSs functioning includes amino acid activation, tRNA recognition, tRNA charging (5). Mt-aaRSs are devoted to the mitochondrial translation, and thus, to the synthesis of the 13 mt-DNA-encoded respiratory chain (RC) complexes (6), for which the activity (7) ultimately lead to ATP production (8). Of note, all other sub-units of the RC complexes (~70) are also imported from the cytosol. cIllustration was taken from Schwenzer et al., 2013.

Book chapter #1

Translation in mammalian mitochondria: Order and disorder linked to tRNAs and aminoacyl-tRNA synthetases

Catherine Florentz, Joern Pütz, Frank Jühling, H. Schwenzer, Peter Stadler, Bernard Lorber, Claude Sauter, Marie Sissler

2013

Published in "Translation in mitochondria and other organelles"
Springer Berlin Heidelberg (Publisher), Anne-Marie Duchêne (Editor)
(in press)

Layout: **T1 Standard SC** | Book ID: 307498_1_En | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **55/83**

Editor Proof

Chapter 3
Translation in Mammalian Mitochondria: Order and Disorder Linked to tRNAs and Aminoacyl-tRNA Synthetases

Catherine Florentz, Joern Pütz, Frank Jühling, Hagen Schwenzer, Peter F. Stadler, Bernard Lorber, Claude Sauter and Marie Sissler

Abstract Transfer RNAs (tRNAs) and aminoacyl-tRNA synthetases (aaRSs) are key actors in all translation machineries. AaRSs aminoacylate cognate tRNAs with a specific amino acid that is transferred to the growing protein chain on the ribosome. Mammalian mitochondria possess their own translation machinery for the synthesis of 13 proteins only, all subunits of the respiratory chain complexes involved in the synthesis of ATP. While 22 tRNAs and two ribosomal RNAs are also coded by the mitochondrial genome, aaRSs are nuclear encoded and become imported. The fact that the two cellular genomes, nuclear and mitochondrial,

C. Florentz (✉) · J. Pütz · F. Jühling · H. Schwenzer · B. Lorber · C. Sauter · M. Sissler
Institut de Biologie Moléculaire et Cellulaire du CNRS, Architecture et Réactivité de l'ARN, Université de Strasbourg, 15 rue René Descartes, 67084 Strasbourg, Cedex, France
e-mail: C.Florentz@ibmc-cnrs.unistra.fr

J. Pütz
e-mail: J.Puetz@ibmc-cnrs.unistra.fr

F. Jühling
e-mail: frank@bioinf.uni-leipzig.de

H. Schwenzer
e-mail: H.Schwenzer@ibmc-cnrs.unistra.fr

B. Lorber
e-mail: B.Lorber@ibmc-cnrs.unistra.fr

C. Sauter
e-mail: C.Sauter@ibmc-cnrs.unistra.fr

M. Sissler
e-mail: M.Sissler@ibmc-cnrs.unistra.fr

F. Jühling · P. F. Stadler
Bioinformatics Group, Department of computer Science and Interdisciplinary Center for Bioinformatics, University of Leipzig, Härtelstrasse 16-18, 04107 Leipzig, Germany
e-mail: studla@bioinf.uni-leipzig.de

UNCORRECTED PROOF

DOI: 10.1007/978-3-642-39426-3_3, © Springer-Verlag Berlin Heidelberg 2013

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **56/83**

Editor Proof

evolve at different rates raises numerous questions as to the co-evolution of partner macromolecules. Herein we review the present state-of-the-art on structural, biophysical, and functional peculiarities of mammalian mitochondrial tRNAs and aaRSs, and of their partnership in their wild-type state. Then, we oppose this mitochondrial "order" to the "disorder" generated by the presence of a variety of mutations occurring in the corresponding human genes that have been correlated to an increasing number of diseases. So far, more than 230 mutations in mitochondrial tRNA genes and a rapidly growing number of mutations in mitochondrial aaRS genes have been reported as causative of a large variety of pathologies. The molecular incidence of mutations on structural, biophysical and functional properties of the related macromolecules will be summarized. Mutations in mitochondrial tRNA genes lead to complex mosaic effects with a major impact on tRNA structure. Some mutations affecting mitochondrial aaRS genes do not interfere with the housekeeping aminoacylation activity, suggesting that mitochondrial aaRSs, alike cytosolic aaRSs are involved in other processes than translation. This opens new research lines.

3.1 Introduction

Mammalian mitochondria possess a small circular genome, coding for 13 polypeptide chains only. These are subunits of the respiratory chain complexes involved in the oxidative phosphorylation of ADP into ATP. The synthesis of these 13 subunits requires a complete mitochondrial translation machinery. Interestingly, the RNA components of this machinery are encoded by the mitochondrial genome. This is the case of 11 mRNAs (leading to the 13 proteins), 2 rRNA, and 22 tRNAs representing the minimal and sufficient set to read all codons (Anderson et al. 1981). All protein components are coded by the nuclear genome, synthesized in the cytosol, and imported into the organelle. They include the full sets of ribosomal proteins, as well as tRNA maturation and modification enzymes, aminoacyl-tRNA synthetases (aaRSs), and also translation initiation, elongation, and termination factors. The dual genetic origin of the macromolecules constituting the translation machinery has raised numerous questions as to their properties, mechanism, and specificities of their partnerships, regulation of expression and mechanisms of co-evolution. Indeed, the mitochondrial genome evolves 15–20 times more rapidly than the nuclear genome (Brown et al. 1979; Castellana et al. 2011), generating highly variable sequences so that the RNAs coded by the mammalian mitochondrial genomes are peculiar. All have lost some information as compared to their bacterial homologs. For instance, mRNAs miss 3′ and 5′ untranslated regions, ribosomal RNAs are significantly shorter than bacterial counterparts and tRNAs present a range of peculiarities, from the absence of a few nucleotide signature motifs to the absence of full structural domains (Willkomm and Hartmann 2006).

Herein we first focus on human mitochondrial tRNAs (mt-tRNAs) and aaRSs (mt-aaRSs) and on their partnerships, to illustrate the present knowledge on

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **57/83**

Editor Proof

peculiar structural, biophysical, and functional properties of two partner macromolecules encoded by two different genomes. Peculiarities of human mt-tRNA genes as highlighted by recent powerful bioinformatics approaches will be summarized and compared to mitochondrial tRNA genes in metazoan. Analysis of tRNAs will illustrate large sequence and structural variability, low thermodynamic stability, and high structural flexibility of this family of RNAs. Current knowledge on human mt-aaRSs will be summarized from genes to structural and functional properties of the corresponding proteins. The partnerships of aaRSs with their substrates will be described not only in terms of aminoacylation properties but also in terms of thermodynamics of substrate binding, a parameter that allowed for clear distinction from bacterial aminoacylation systems. As will be highlighted, the enlarged plasticity of the mitochondrial enzymes as compared to that of aaRS from other origins might be the result of an evolutionary adaptation of the nuclear-encoded protein to the rapidly evolving mitochondria-encoded RNAs. The human mitochondrial aspartylation system, studied in our laboratory, will serve as case study all along the review. After this first part, illustrating the status of wild-type macromolecules and thus referring to "order" in mitochondrial translation, "disorder" will be considered as well as a molecular perturbation as well as a mitochondrial pathology.

In the last two decades, mt-tRNA genes were recurrently reported as hosting point mutations linked to a variety of neuromuscular and neurodegenerative disorders. More recently, nuclear genes coding for mt-aaRSs also became the center of attention, with mutations causative of further disorders. The present understanding of their impact on the molecular level of "disorder" induced within tRNAs and aaRSs molecules will be summarized. While many mutations do interfere with the housekeeping aminoacylation reaction, several others have no detectable effect on it. This is in favor of the existence of additional mt-tRNA and mt-aaRS functions or properties that are altered by these mutations. These supplementary functions and properties need still to be determined. New research lines in this direction will be suggested.

3.2 Mammalian Mitochondrial tRNAs

3.2.1 Structural Properties of Mammalian Mitochondrial tRNAs

The gene content of the human mitochondrial genome is similar to what found in other metazoan mitochondria. It encodes for 22 mt-tRNAs (Anderson et al. 1981), one for each of the 20 amino acid specificities, and two additional ones for leucine and serine, respectively. This minimal set of tRNAs is sufficient for reading of all codons despite the genetic code is different in mitochondria from nuclear genomes. Already at the stage of bovine and human mitochondrial DNA sequencing (Anderson et al. 1981) peculiar structural properties of tRNAs were noticed,

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **58/83**

Editor Proof

marking significant differences with so-called canonical tRNAs present in bacteria or in the cytosol of eukaryotic cells. Most striking was the absence of a possible full cloverleaf structure for some of the expressed gene products, noticeably the absence of a complete structural domain for tRNASer1 (the D-arm of the cloverleaf), and, despite the presence of a potential cloverleaf for the other tRNAs, the absence of typical signature motifs in many sequences, motifs so far known as being conserved in all tRNAs. Mammalian mt-tRNAs fall into two groups according to the location of the corresponding gene on either of the mt-DNA coding strand. "Heavy" tRNAs (eight cases) are G-rich, while "Light" tRNAs (14 cases) are A, U, and C rich, leading to a series of typical structural characteristics such as biased base-pair content (Helm et al. 2000). Further, the full set of mammalian (and metazoan) mt-tRNAs do have a short variable region (Fig. 3.1) while canonical tRNAs also include structures with large variable regions (Giegé et al. 2012). Bioinformatic alignments to additional mammalian mt-tRNA genes confirmed and extended these properties and allowed for a fine-tuned description of detailed structural characteristics of each of the 22 tRNA families (Helm et al. 2000). As an outcome, mammalian mt-tRNAs structural properties range from canonical to highly degenerated types. Figure 3.1a indicates the most common degenerations of the classical cloverleaf structure in human mitochondrial sequences, particularly the loss of D/T loop interactions. Mt-tRNAs are expressed in cells to very low levels as compared to their cytosolic counterparts (estimated as 1 mt-tRNA per 160 cytosolic tRNA), rendering access for biochemical investigation of these RNAs very limited (Enriquez and Attardi 1996). Detailed experimental secondary structures were however established on in vitro transcripts and revealed large structural flexibility, with alternative folds to the cloverleaf co-existing in equilibrium (Bonnefond et al. 2005b; Helm et al. 1998; Sohm et al. 2003). This is due to a severe bias in nucleotide content (A, U, and C-rich, and G-poor) especially for the 14 tRNAs coded by the light DNA strand, leading to very low numbers of stable G-C base-pairs, and at opposite, to high levels of weaker A-U and G-U pairs. To be noticed also that the thermodynamic stabilities calculated for the cloverleaf folds are twice as weak as compared to those of canonical tRNAs (Fender et al. 2012). Post-transcriptional modifications (only characterized in a limited number of sequences; Suzuki et al. 2011) were found to be the triggers for stabilization of the cloverleaf fold as demonstrated with the case study highlighting the role of m1A9 modification in human mt-tRNA^{Lys} (Helm and Attardi 2004; Helm et al. 1998; Motorin and Helm 2010). This modification hinders base pairing of residue 9 (and of neighboring nucleotides 8 and 10) in the 3′-end domain of the tRNA with residues 64 (and neighboring residues 65 and 63, respectively) in the T-stem. Post-transcriptional modifications in mammalian mt-tRNAs (only characterized in a limited number of sequences; Suzuki et al. 2011) remain however quantitatively far more limited as compared to other tRNAs (7 % of modified nucleotides as compared to 13 %) suggesting that they play crucial roles (Helm et al. 1999; Suzuki et al. 2011). Interestingly, a number of specific modifications, such as taurine-dependent modifications of anticodon nucleotides, were reported to be key actors in mitochondrial codon reading (Kirino et al. 2005).

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **59/83**

Editor Proof

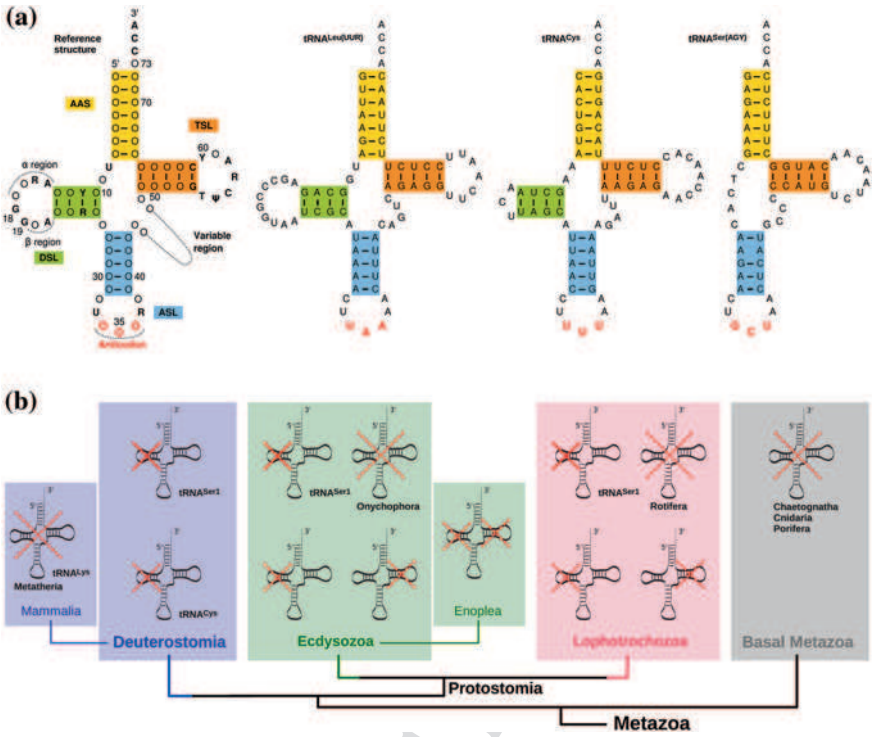


Fig. 3.1 Structural diversity and evolution of human mitochondrial tRNAs. **a**. Cloverleaf structures of three typical human mt-tRNA secondary structures as compared to the canonical cloverleaf reference structure (*left*). *AAS* amino acid acceptor stem, *DSL* D-stem and loop, *ASL* anticodon stem and loop, *TSL* T-stem and loop. The three examples presented, support the large range of structural profiles covered. $tRNA^{Leu(UUR)}$ is of classical type, with the full cloverleaf domains as well as the full set of conserved nucleotides involved in tertiary folding (see, in particular, the presence of residues G18, G19 in the D-loop, and of residues T54T55C56 in the T-loop, allowing for interaction between the corresponding domains). $tRNA^{Lys}$ illustrates the family of mt-tRNAs still presenting the four domains of the cloverleaf, but with serious variations in the size of the D- and T-loops and their nucleotide content, missing the conserved elements, as well as the unusual nucleotides in the short connector between the acceptor and the D-stems. Finally, $tRNA^{Ser(AGY)}$ is the typical tRNA missing a full structural domain, namely the D-arm. **b**. Simplified view on the evolution of tRNA sets and tRNA secondary structures in Metazoa. Each panel highlights the most striking and representative structural deviation present within the considered evolutionay group. Accordingly, basal metazoan lost the full set of mt-tRNA genes and mammals, lost the gene for $tRNA^{Lys}$ in methateria

The question as to the 3D fold of mammalian mt-tRNAs has retained the attention of many investigators. Indeed, whatever the secondary structural properties of these RNAs, it is expected that their 3D structures allow for the recognition by partner macromolecules, in particular the cognate aaRS for aminoacylation and the ribosome for codon reading and protein synthesis. An L-shaped structure based on an acceptor branch (T-arm and acceptor arm) and an anticodon branch (D-arm and anticodon

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **60/83**

Editor Proof

arm) is thus expected as for tRNAs of all kingdoms of life (Giegé et al. 2012). This is considered despite the general absence of signature nucleotides (the so-called conserved and semi-conserved nucleotides) known to support such a fold in canonical tRNAs. Birefringence measurements, NMR, and chemical structure probing in solution were the leading approaches confirming the existence of L-shape-like folds for mt-tRNAs, bringing the two functional extremities, namely the acceptor 3′-end and the anticodon, at the expected distance from each other (reviewed in Giegé et al. 2012). The 3D structure of canonical tRNAs is based on a set of tertiary interactions between nucleotides at distance in the secondary structure, in particular on a network of interactions defining and stabilizing the "central core" of the L-shaped structure. Fine-tuned chemical structural probing on two case studies, human mt-tRNA^{Asp} and bovine mt-tRNA^{Phe} confirmed the existence of these networks (Messmer et al. 2009; Wakita et al. 1994). It remains to be verified if such networks can take place in all mt-tRNAs, an hypothesis that could not be supported by nucleotide conservation, but that will need a more detailed consideration of nucleotide partnership rules as tackled by Leontis and Westhof (Leontis et al. 2002). Initial data are available suggesting that at least all mammalian mt-tRNA^{Asp} would benefit from the set of tertiary network interactions despite nonconservation of involved nucleotides (Messmer et al. 2009). Extension of the analysis to the full set of mammalian mt-tRNAs is in progress. In regard of the elbow of the L, no hydrogen bonds could be detected between the D and T-loops, so that there may be no stabilization of the global L at this level. Accordingly, the angle formed by the two branches of the L may vary according to the tRNA, as also demonstrated by birefringence methods (Friederich and Hagerman 1997), and likely may vary according to its functional state, allowing for various structural adaptations to partner macromolecules (reviewed in Giegé et al. 2012). The role of post-transcriptional modifications in the global flexibility of the tRNA remains to be determined.

In summary, mammalian mt-tRNAs present a large diversity of structural types, ranging from canonical stable type for a few tRNA families only, to highly "bizarre" versions, characterized by an unusually high flexibility and plasticity of the full macromolecule, along an L-shaped-like global fold.

3.2.2 Comparison with Mitochondrial tRNA Genes from Other Metazoan

Due to their peculiarities, metazoan mt-tRNA genes are difficult to detect with the available search programs such as tRNAscan-SE (Lowe and Eddy 1997) and ARWEN (Laslett and Canbäck 2008), so that an approach using INFERNAL and covariance models was developed and implemented (Nawroki et al. 2009). A first systematic overview on nearly 2,000 metazoan RefSeq genomes (Pruitt et al. 2007) and their tRNA gene content allowed for interesting insights on the evolution of mt-tRNA genes (Jühling et al. 2012a). An overview on tRNA genes deprived of information coding for a full structural domain of the expressed RNA (D- or T-arm), and

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **61/83**

Editor Proof

on loss of full mt-tRNA genes is illustrated in Fig. 3.1b. While mt-tRNA genes are encoded by mt-genomes over all metazoan clades, some organisms miss mt-tRNA genes and form exceptions. The loss of mt-tRNA genes was shown to be related with the loss of the corresponding aaRS genes in Cnidarians (e.g., jellyfish) (Haen et al. 2010). Members of Ceractinomorpha (sponges) (Wang and Lavrov 2008), Chaetognatha (arrow worms) (Miyamoto et al. 2010), and Rotifera (freshwater zooplankton) (Suga et al. 2008) also lost mt-tRNA genes. To be mentioned is the specific loss of $tRNA^{Lys}$ genes in Metatheria (e.g., marsupials) where only a pseudo-gene remains (Dörner et al. 2001). It is assumed that in these organisms, by nuclear tRNAs (Alfonzo and Söll 2009; Duchêne et al. 2009).

Highly truncated tRNA gene sequences are known for Onychophora (velvet worms) (Braband et al. 2010), where unusual editing mechanisms are assumed to repair the tRNAs (Segovia et al. 2011) while other genes are completely missing. Other mt-tRNA genes lost only small sequence stretches so that the canonical cloverleaf of the tRNA can form without the need of repairing mechanisms. However, more than 90 % of mt-tRNAs share a four-arm cloverleaf structure, and the 10 % remaining mostly lost either the T-or the D-stem and developed replacement loops. The corresponding genes are found throughout all metazoan clades but are locally conserved. Hotspots of organisms where the full set of mt-tRNA genes have lost a complete domain are Ecdysozoa (arthropoda and nematoda) and Lophotrochozoa (molluscs and worms). In contrast, in Lepidosauria (turtles, snakes, and lizards) only the genes for $mt\text{-}tRNA^{Cys}$ seem to spontaneously have lost the D-stem coding region along several parallel events (Macey et al. 1997; Seutin et al. 1994). Another well-known example of truncated mt-tRNA genes concerns Nematoda, where all mt-tRNAs lost their T-stem, except those of both $tRNA^{Ser}$, which instead lost their D-stem (Wolstenholme et al. 1994). While these truncated secondary structures in *C. elegans* are conserved within all other members of its family (round worms), a recent study on their sister group Enoplea detected even genes for "armless" tRNAs, namely sequences without both D- and T-stems (Jühling et al. 2012b). These sequences are conserved and are thereby the shortest functional tRNA sequences known so far (Jühling et al. 2012b). Compared to other metazoan mt-tRNA genes, mammalian genes lead to nearly canonical cloverleaf structures, and encode the full gene content with the exception of $tRNA^{Lys}$, which is missing in marsupials. However, hotspots of gene losses and of genes of truncated structures are distributed in other parts of the metazoan tree. Only the $tRNA^{Ser1}$ gene, coding for a tRNA deprived of a D-domain, shows equal distribution all along the metazoans (Arcari and Brownlee 1980; de Bruijn et al. 1980).

3.3 Mitochondrial aaRSs

3.3.1 Genes and Proteins

Present-day mitochondrial genomes do not code for aaRSs, so that the full set of enzymes is necessarily nuclear encoded and that the expressed proteins become

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **62/83**

Editor Proof

imported into mitochondria. A set of specific genes for mammalian mt-aaRSs has been recorded (Bonnefond et al. 2005a). This set distinguishes from the set of cytosolic synthetases for most enzymes, with the sole exceptions of LysRS and GlyRS, coded by a same gene but with distinguishing features allowing for dual location of the protein. The gene for LysRS undergoes alternative splicing so that the two final mature proteins distinguish by a few N-terminal amino acids only (Tolkunova et al. 2000). GlyRSs are generated from two translation initiation sites on the same gene, so that the two mature proteins are the same, but differ by the presence or the absence of a mitochondrial targeting sequence (MTS) (Mudge et al. 1998; Shiba et al. 1994). A further exception concerns mt-GlnRS for which no specific gene was found so far. As an alternative to the existence of a specific mt-GlnRS, the possible import of the cytosolic enzyme was proposed (Rinehart et al. 2005) as well as the existence of an indirect pathway based on misaminoacylation of mt-tRNA^{Gln} with mt-GluRS, followed by transamidation of the charged glutamic acid into glutamine. Such a pathway exists both in yeast (Frechin et al. 2009a) and human mitochondria (Nagao et al. 2009). In yeast, the dual localization of GluRS is controlled by binding to Arc1p, a tRNA nuclear export cofactor that behaves as a cytosolic anchoring platform. When the metabolism of the yeast cell switches from fermentation to respiration, the expression of Arc1p is down-regulated and this increases the import of GluRS to satisfy a higher demand of mt glutaminyl-tRNA^{Gln} for mitochondrial protein synthesis (Frechin et al. 2009b).

Almost all mammalian mt-aaRSs have about the same size (once the MTS is removed upon import into mitochondria) and the same structural 2D organization as their homologs from the three kingdoms of life (Bonnefond et al. 2005a). The only exception is PheRS which is classically a tetramer ($\alpha 2\beta 2$) but only a monomer in mitochondria (Klipcan et al. 2008). Also, all mt-aaRSs contain the expected signature motifs for amino acid specificity and signature motifs of either class I or class II of aaRSs (Bonnefond et al. 2005a; Brindefalk et al. 2007; Sissler et al. 2005; Woese et al. 2000). The evolutionary origin of the different genes however remains intriguing. Sequence alignments have indeed not identified sequence stretches or signature motifs that could originate from alpha-proteobacterial ancestors, favoring various horizontal gene transfers events along evolution (Brindefalk et al. 2007).

During import into the mitochondria, the MTS is cleaved enzymatically. The detailed process is still under exploration for aaRSs but the trend for other proteins follows a two-step process. Once the polypeptide chain has crossed the mitochondrial membranes and entered the mitochondrion, a first cleavage occurs. This cleavage might be definite. Alternatively, the partially matured protein then localizes either in the matrix or anchors into (or at the surface of) the inner mitochondrial membrane before another part of the MTS is cut off. Interestingly, ribosomes and elongation factors have been found close to the inner membrane suggesting that probably the entire machinery required for protein biosynthesis is located there, and accordingly aaRSs too. A particularity of MTSs is their nonconserved length, sequence and amino acid compositions (e.g., Chacinska et al. 2009). So far, no reliable consensus sequence was found that could be used to accurately

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **63/83**

Editor Proof

predict the positions of maturation. This was highlighted by a wrongly predicted cleavage site of human mt-LeuRS, which led to a poor expression of a tentative mature protein in *E. coli* while a variant, shortened by further 39 N-terminal amino acids overproduced well in *E. coli* and was purified as an active enzyme while the one deprived of only 21 amino acids was insoluble (Bullard et al. 2000; Yao et al. 2003). Another well-documented example is human mt-AspRS. The predicted mature protein had a very low solubility when overexpressed in *E. coli* cells. Dynamic light scattering analyses revealed that aggregation proceeded during purification. A comparative analysis of a set of variants differing by their N-terminal sequences revealed that expression of the protein was actually enhanced when the N-terminus was extended by seven natural amino acids of the predicted mature N-terminus (Gaudry et al. 2012). The redesigned protein was highly soluble, monodisperse and functionally active in tRNA aminoacylation. It yielded crystals that were suitable for structure determination (Gaudry et al. 2012; Neuenfeldt et al. 2013). These results suggest that additional criteria should be taken into account for the prediction of the correct MTS cleavage sites and that the definition of the precise N-terminus of mature mt-aaRS should be determined experimentally.

3.3.2 Crystallographic Structures

As already highlighted, despite various evolutionary origins of the genes coding for mammalian mt-aaRSs, several enzymes have a same modular organization than their bacterial homologs. This is illustrated in Fig. 3.2 with three of the four crystal structures that have been determined so far for mt-aaRSs for exclusive mitochondrial location (additional crystallographic structures are available for human GlyRS (Cader et al. 2007; Xie et al. 2006) and for LysRS (Guo et al. 2008), aaRSs of dual cytosolic and mitochondrial location). Bovine mt-SerRS, human mt-TyrRS, and human mt-AspRS show an overall architecture close to that of their prokaryotic homologs (Bonnefond et al. 2007; Chimnaronk et al. 2005; Neuenfeldt et al. 2013). Mt-PheRS is again the exception. Instead of forming complex heterodimeric assemblies as bacterial, archaeal, and cytosolic enzymes, it forms a two-domain monomer, which only maintains the catalytic domain characteristic of class II. This human mitochondrial version is the smallest known aaRS (Yadavalli et al. 2009).

Along with the reduction of the tRNA pool (22 in human mitochondria) and the simplification of identity rules, several of these enzymes have adapted the way they recognize their tRNA substrate, especially when the latter display a non canonical 2D fold leading to higher flexibility. Positively charged patches at the surface of mt-SerRS were redistributed to bind tRNAs lacking an extended variable region, the hallmark and major identity element of prokaryotic, eukaryotic, and archaeal $tRNAs^{Ser}$ (Chimnaronk et al. 2005). The three other mt-aaRSs display a more electropositive tRNA-binding interface, which may favor interactions

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **64/83**

Editor Proof

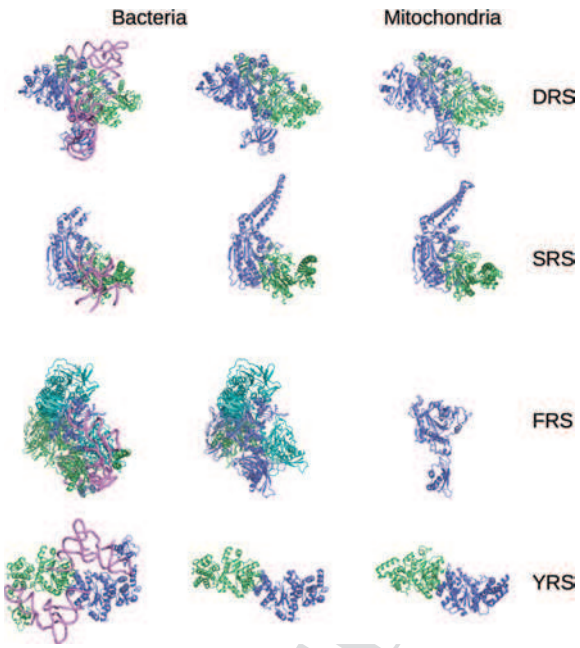


Fig. 3.2 Cyrstal structures of mammalian mitochondrial aaRSs (bovine or human) and of their bacterial homologs (*E. coli -Eco-* or *Thermus thermophiles -Tth-*). On the left, bacterial complexes (PDBids: 1C0A for *Eco*DRS/tRNA, 1SRS for *Tth*SRS/tRNA, 2IY5 for *Tth*FRS/tRNA and 1H3E for *Tth*YRS/tRNA) are shown with monomers A in blue, monomers B in green and tRNAs in pink, indicating the binding site of the cognate substrate. In the case of tetrameric bacterial FRS, monomers C and D are depicted in cyan. On the right, free forms of bacterial aaRSs (1EQR for *Eco*DRS, 1SRY for *Tth*SRS, 1B7Y for *Tth*FRS, 1H3F for *Tth*YRS without its C-terminal domain) and mitochondrial enzymes (4AH6 for *Hsa*DRS2, 1WLE for *Bta*SRS2, 3TUP for *Hsa*FRS2, and 2PID for *Hsa*YRS2) are represented in the same orientation and same color code. Except *Hsa*FRS which exhibits a totally different structural organization (monomer instead of heterotetramer), mitochondrial aaRSs have retained the overall architecture of their bacterial relatives. Names of organisms are abbreviated in a three-letter code (e.g., *Homo sapiens*: *Hsa*). Mitochondrial enzymes are referred to by the addition of the number 2

313 with the sugar-phosphate backbone of the substrate to compensate for a reduction
314 of specific contacts with identity elements (Neuenfeldt et al. 2013). The ability
315 of mt-TyrRS, mt-PheRS, and mt-AspRS to aminoacylate heterologuous tRNAs
316 (Bonnefond et al. 2005a; Klipcan et al. 2012; Neuenfeldt et al. 2013) indicates a
317 much higher substrate tolerance, which may be linked to an increased structural
318 plasticity. For instance, mt-PheRS undergoes a large movement of its anticodon-
319 binding domain upon tRNA binding, switching from a closed to an open confor-
320 mation (Klipcan et al. 2012). The higher thermal sensitivity of human mt-AspRS
321 as compared *E. coli* AspRS, its more open catalytic groove in the absence of
322 tRNA and the amplitude of thermodynamic terms associated with tRNA binding
323 are also in favor of a more dynamic structure (Neuenfeldt et al. 2013). Altogether,

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **65/83**

Editor Proof

it appears that mt-aaRS properties have evolved to accompany the sequence and structure drift of mt-tRNAs. Enlarged intrinsic plasticity within a conserved architectural framework is one striking feature along this line. The underlying mechanisms enabling the crosstalk between nuclear and mitochondrial genomes remain to be explored.

3.4 tRNA/aminoacyl-tRNA Synthetase Partnerships in Mammalian Mitochondria

3.4.1 Aminoacylation of tRNAs, the Housekeeping Function of aaRSs

The partnership of tRNAs and aaRSs is a key event in translation. Each synthetase recognizes specifically its tRNA or family of isoaccepting tRNAs, and esterifies its 3′-CCA end with the specific amino acid. The charged tRNA enables delivery of the amino acid to the ribosome where translation takes place. The aminoacylation reaction involves a two-step process including first the activation of the specific amino acid into an adenylate in the presence of ATP, and second, the specific recognition of the cognate tRNA followed by transfer of the activated amino acid (Ibba et al. 2005). Deciphering the detailed mechanisms of these steps for bacterial, archeal, and eukaryotic cytosolic aminoacylation has retained the attention of a large number of research groups over several decades (Ibba et al. 2005). Analysis of mammalian mitochondrial aminoacylation systems is only at initial stages. Due to the dual origin of the two partner macromolecules and to the diverging structural properties of mt-tRNAs, the mechanisms of reciprocal recognition and of co-evolution of these macromolecules deserve much attention.

3.4.2 Mammalian Mitochondrial Synthetases have Low Catalytic Activities

Only a limited number of recombinant mammalian aaRSs have been obtained so far, allowing for biochemical and enzymatic characterization in vitro. As already reviewed elsewhere (Florentz et al. 2003; Suzuki et al. 2011) these enzymes present a 20- to 400- fold lower catalytic activity than their cytoplasmic and bacterial homologs. In the specific case of human mt-AspRS as compared to *E. coli* AspRS, the affinity for the substrate tRNA is higher by an order of magnitude as measured by isothermal titration calorimetry (ITC) while the affinity for an analog of the activated amino acid is of same level. However, the catalytic rate kcat for aminoacylation is 40-fold lower for the mt-aaRS (Neuenfeldt et al. 2013). The molecular reasons explaining the lower rate remain however elusive. Indeed,

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **66/83**

Editor Proof

overimposition of the catalytic sites in the crystallographic structures of both enzymes, lead to an important overlap (less than 2 Å rmsd) not allowing to pinpoint intrinsic differences inline with a different catalytic activity (Fender et al. 2006; Neuenfeldt et al. 2013).

3.4.3 Identity Elements in Mitochondrial tRNAs are Limited

Specific recognition of tRNAs by aaRSs is driven by identity elements present in the tRNA (Giegé 2008; Giegé et al. 1998). These elements have been searched by mutagenic approaches on in vitro transcripts for a few mammalian mt-tRNAs (Florentz et al. 2003; Suzuki et al. 2011). Interestingly, while these sets are generally conserved along different organisms and even along kingdoms for a given amino acid specificity, they were found distinct in mt-tRNAs. A striking example concerns identity elements for aspartylation, one or the rare systems so far investigated. Major identity elements (elements for which strongest effects are observed upon mutation) are conserved all along evolution, as residues G73 (the so-called discriminator residue near the 3′- acceptor end), residue G10 in the D-stem, and residues G34, U35, and C36 forming the anticodon triplet. Transfer of this set of residues into host tRNAs of different specificities, converts these tRNAs into aspartic acid accepting species (Giegé et al. 1996). A mutagenic analysis performed on human mt-tRNA^{Asp} revealed that only residues U35 and G36 are important elements for specific recognition and aspartylation by mt-AspRS while the other elements can be replaced by any other nucleotide without influencing the efficiency of tRNA recognition and aspartylation (Fender et al. 2006). Figure 3.3 illustrates this point. The striking non-importance of residue 73, otherwise highly conserved as G in tRNA^{Asp} over all kingdoms of life, is a signature of mt-tRNA degeneration, and at the same time, of evolutionary adaptation of the synthetase. Deep insight into the structural environment of residue 73 in the catalytic site of human mt-AspRS reveals an enlarged space as compared to *E. coli* AspRS and other AspRS, allowing the fit of any of the four nucleotides, rather than the exclusive fit of a G residue at this position. A mutagenic analysis of the enzyme has confirmed this view (Fender et al. 2006). Another example of the peculiarity of identity elements in a mitochondrial system concerns human mt-tRNA^{Tyr} (Bonnefond et al. 2005b). Base-pair G1-C72, forms an important identity element in archaeal and eukaryal tRNA^{Tyr} (Bonnefond et al. 2005b). The mitochondrial tyrosine identity disobeys this rule, since mt-TyrRS is able to aminoacylate as well a tRNA with the G1-C72 pair as the opposite pair C1-G72. Other examples have been reviewed previously (Florentz et al. 2003) and will not be further discussed herein. They indicate that sequence analysis of mammalian mt-tRNAs lead to the conclusion that only a limited number of identity elements known for non mitochondrial aminoacylation systems are present (Florentz et al. 2003).

Despite the limited number of aminoacylation identity elements in mt-tRNAs, translation in mitochondria needs to be accurate so that the 13 synthesized proteins

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **67/83**

Editor Proof

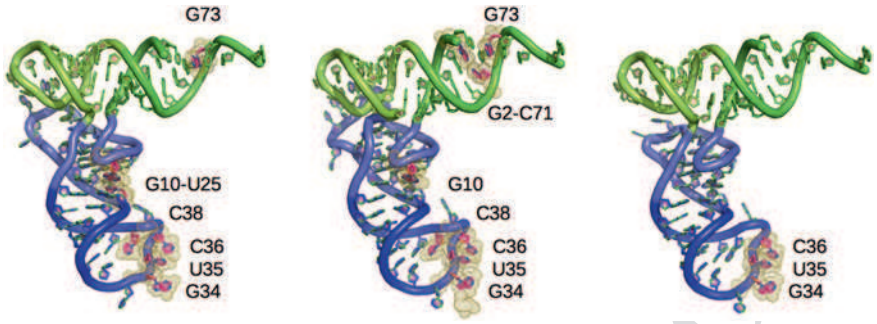


Fig. 3.3 Evolution of $tRNA^{Asp}$ structure and identity elements. From the left to the right: $tRNA^{Asp}$ from the yeast *Saccharomyces cerevisiae* (conformation in the complex with AspRS, PDBid: 1ASY), from *E. coli* (conformation in the complex with AspRS, PDBid: 1C0A), from human mitochondria (homology model—(Messmer et al. 2009)). The three molecules are shown in the same orientation with the acceptor stem in green, the D stem loop in light green, the anticodon stem loop in dark blue, the variable loop and the T stem loop in light blue, respectively. Nucleotides defining the aspartate identity (Fender et al. 2012; Giegé et al. 1996) in each system are depicted in pink with a dotted surface. The human mitochondrial $tRNA^{Asp}$ is characterized by a shortening of D and T loops, leading to an absence of bases interactions at the corner of the L scaffold, and by a reduced set of identity elements

(all subunits of respiratory chain complexes that are partners of more than 80 nuclear encoded subunits) are prepared without mistakes. It is hypothesized that the small competition created by 22 tRNAs only toward about the same number of aaRSs in the mitochondrial environment, as compared to more complex translation machineries in bacteria or eukaryotic cytosol with several hundreds of tRNAs for 20 aaRSs, can deal with a restricted number of identity elements. Further, selection by the elongation factor EF-Tu of properly charged tRNAs only may represent an additional process toward accurate protein synthesis (Nagao et al. 2007), as is the case in bacteria (LaRiviere et al. 2001).

3.4.4 Unprecedented Plasticity of Mitochondrial aaRSs and tRNAs

The discovery of mt-tRNAs that have lost critical structural information raised the question about how the nuclear-encoded mt-aaRSs have adapted to be able to deal with their partners? As already discussed, first insights were provided by resolution of crystal structures of mt-aaRSs (Bonnefond et al. 2007; Chimnaronk et al. 2005; Klipcan et al. 2008; Klipcan et al. 2012; Neuenfeldt et al. 2013). In the specific case of human mt-AspRS, a typical homodimeric bacterial-type AspRS, the 3D architecture is very close to that of the *E. coli* enzyme of same specificity, with the exception that it has a wider catalytic groove, a more electropositive surface potential, and an alternate interaction network at the subunits interface, a set of properties in line with

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **68/83**

Editor Proof

support to facilitated tRNA partnership. An additional biophysical property, namely thermostability, illustrates further the originality of the protein. Comparative differential scanning fluorimetry analyses indicated that the mitochondrial protein is far less stable with regard to temperature than its bacterial homolog. It has a 12 °C lower melting point (Neuenfeldt et al. 2013). These properties are summarized in Table 3.1.

The partnership of a mt-aaRS (human mt-AspRS) with its substrates was investigated by ITC, an approach allowing for direct measurement of affinity (Kd) and of the thermodynamic parameters ΔH (variation in enthalpy), ΔS (variation in entropy), and ΔG (variation in free energy) (Neuenfeldt et al. 2013). Comparative analyses between human mt-AspRS and *E. coli* AspRS revealed a one order of magnitude higher affinity of the mitochondrial enzyme for cognate and noncognate tRNAs (cross binding studies of mt-AspRS with *E. coli* tRNA, and of *E. coli* AspRS with mt-$tRNA^{Asp}$), but with highly different entropy and enthalpy contributions (Table 3.1). Binding parameters of the cognate mitochondrial partners requires far larger enthalpic and entropic contributions than binding of the cognate bacterial partners, underlining reciprocal reorganization along complex formation. Such an adaptation is still possible when the mt aaRS meets the bacterial tRNA but is not possible in the opposite situation, namely when the bacterial enzyme and the mt-tRNA face each other. Thermodynamics thus contribute to explain the well-known unilateral aminoacylation of bacterial synthetases for bacterial tRNAs (Kumazawa et al. 1991). Interestingly, ITC measurements of small substrate binding, revealed that both enzymes bind a synthetic analog of the aspartyl-adenylate by a cooperative allosteric mechanism between the two subunits of the dimeric enzymes, but again with different thermodynamic contributions (Neuenfeldt et al. 2013) (Table 3.1).

Altogether, presently available structural, biophysical, and thermodynamic data support the view of so far unsuspected greater flexibility of mt-aaRS with respect to its bacterial homolog albeit a common architecture. This gain in plasticity may represent an evolutionary process that allows the nuclear-encoded proteins to adapt to the structurally degenerated RNAs from organelles. Evolutionary induced changes in intrinsic properties of proteins, may thus represent an alternative to other strategies, such as those reported for the mitochondrial ribosome, where the strong restriction in RNA sizes is compensated by extension of the number and size of the nuclear encoded proteins (Willkomm and Hartmann 2006). If mt-aaRS do have partner proteins that might also contribute to improved recognition of degenerated tRNAs remains an open question.

3.5 Human Mitochondrial tRNA and Synthetases in Pathologies

3.5.1 Mitochondrial tRNAs and Human Pathologies

In the last two decades, a large number of human neuromuscular and neurodegenerative disorders have been reported as correlated to point mutations in the mt-DNA encoded genes, with a large prevalence of mutations in tRNA genes

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **69/83**

Editor Proof

Table 3.1 Major structural and functional differences between dimeric mt and bacterial aspartyl-tRNA synthetases in favor of a greater plasticity of the organelle enzyme

Aspartyl tRNA synthetase			
Dimer surface	20 additional basic residues		(18 Lys + 2 Arg)
	Electrostatic potential more positive		
Dimer interface	70 versus 60 H bonds and 28 versus 20 salt bridges		
	About 25 % less specific interactions per $Å^2$		
Thermal stability (T_m)	Alone		37 °C versus 50 °C
	Bound to cognate tRNA		40 °C versus 50 °C
	Bound to AspAMS		45 °C versus 55 °C
Thermodynamic parameters			
Cognate $tRNA^{Asp}$	Kd	0.26 versus 3.1	μM
	ΔH	−20.3 versus −13.0	kcal/mol
	ΔT	+ 11.2 versus + 5.5	kcal/mol
	ΔG	−9.1 versus −7.5	kcal/mol
Noncognate $tRNA^{Asp}$	Kd	0.24 versus 22	μM
	ΔH	−14.0 versus −30.8	kcal/mol
	ΔT	+ 5.0 versus +24.4	kcal/mol
	ΔG	−9.0 versus −6.4	kcal/mol
AspAMS (monomer 1)	Kd	129 versus 29	nM
	ΔH	−13.2 versus −5.5	kcal/mol
	ΔT	+ 3.9 versus −4.8	kcal/mol
	ΔG	−9.4 versus −10.2	kcal/mol
AspAMS (monomer 2)	Kd	17 versus 3	nM
	ΔH	−21.8 versus −8.2	kcal/mol
	ΔT	+ 10.6 versus −3.5	kcal/mol
	ΔG	−10.6 versus −11.6	kcal/mol
$tRNA^{Asp}$			
Nucleotide content	90 versus 66 % A, U, and C		
	16 versus 7 out of 21 A–U and G–U base pairs		
	D and T loops are not classical		
Structural stability	ΔG	−22 versus −42	kcal/mol

Both polypeptide chains share 43 % identity between their amino acid sequences
Data compiled from (Fender et al. 2012) and (Neuenfeldt et al. 2013)

UNCORRECTED PROOF

(reviewed for example in Florentz and Sissler 2003; Suzuki et al. 2011; Yarham et al. 2010; Ylikallio and Suomalainen 2012). Among the mutations leading to "mitochondrial disorders", 232 are distributed all over the 22 tRNAs (Fig. 3.4; data from MITOMAP, a human mitochondrial genome database http://www.mitomap.org/MITOMAP). Most striking cases concern $tRNA^{Lys}$, $tRNA^{Leu(UUR)}$, and $tRNA^{Ile}$, which form "hot spots" for mutations. Mutations in these tRNAs are most frequently correlated with Myoclonus Epilepsy with Ragged Red Fibers (MERRF) (Shoffner et al. 1990) and Mitochondrial Encephalomyopathy with Lactic Acidosis and Stroke-like episodes (MELAS) (Goto et al. 1990), respectively. However, this does not indicate a peculiar mutational susceptibility of the three mitochondrial genes, but is more likely due to systematic and intensive

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **70/83**

Editor Proof

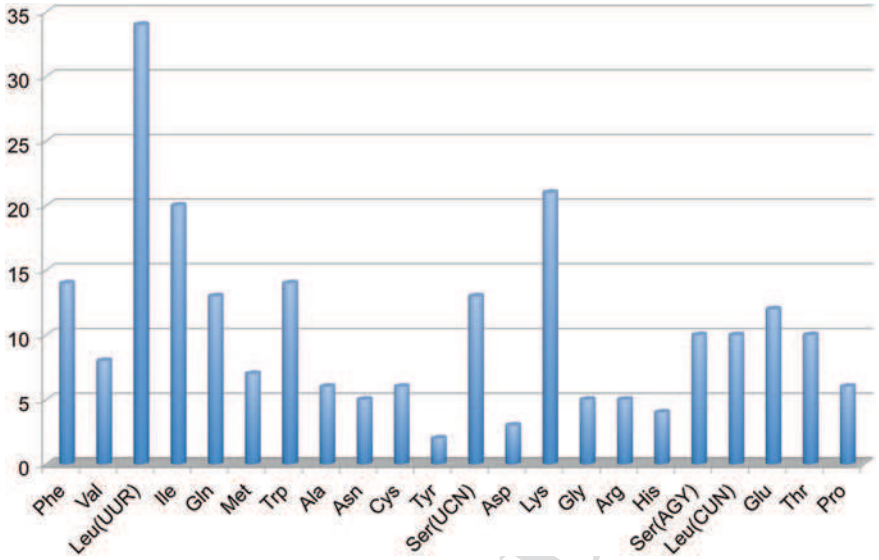


Fig. 3.4 Pathology-related mutations in the 22 mt- tRNA genes. tRNA genes are indicated by the three-letter code of the corresponding amino acid and are sorted according to their location on mt-DNA. Each column corresponds to the number of different mutations reported so far. To be noticed, three additional mutations are reported in the precursor of $tRNA^{Ser(UCN)}$ and one additional mutation at the junction between $tRNA^{Gln}$ and $tRNA^{Met}$. Data are from MITOMAP, a database for mitochondrial genome mutations (http://www.mitomap.org/MITOMAP)

UNCORRECTED

investigations of theses firstly reported examples of mt-tRNA genes correlated with human mitochondrial disorders.

The exponential rate of discovery of mutations in tRNA genes (3 mutations reported in 1990 and 232 mutations in 2012), together with the key role of tRNA in mitochondrial protein synthesis (linked to its global biology including gene expression, tRNA maturation, specific amino acid transfer, and regulation of these different functions), called for clarification of the molecular mechanisms for their pathogenicity. However, the relationships between genotypes and phenotypes appear very complicated since a given mutation can lead to a large variety of disorders of different severities (ranging from, e.g., limb weakness and exercise intolerance, to diabetes, leukoencephalopathy, encephalomyopathy, or fatal infantile cardiomyopathy, etc.). At the opposite, a given disorder can be linked to a variety of single point mutations in different tRNA genes. Also, each cell may contain hundreds to thousands copies of the mitochondrial genome, in a mixture of wild-type and mutated versions (heteroplasmic status). The variable distribution of affected tissues and the variable heteroplasmy levels lead to remarkable erratic and heterogenous clinical manifestations. Therefore, establishment of a mitochondrial disorder diagnosis can be difficult. It requires an evaluation of the family pedigree, in conjunction with a thorough assessment of clinical, imaging, and muscle biopsy analyses (McFarland et al. 2004). Also, a theoretical comparison of polymorphic

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **71/83**

Editor Proof

(neutral mutations with no pathogenic manifestations) versus pathogenic mutations remains unsuccessful to identify simple basic features (at the levels of primary and secondary tRNA structures) that would make possible the prediction of pathogenicity of new mutations (Florentz and Sissler 2001; Yarham et al. 2011).

Numerous studies have attempted to unravel the molecular impacts of the mutations on the various properties of the affected tRNAs and lead so far to a mosaïcity of impacts. It is now clear that mutations can affect any step of the tRNA life cycle, either along tRNA biogenesis (maturation of 3′- or 5′- ends within the initial primary transcript, synthesis of the non coded CCA end, post-transcriptional modifications, folding and structure, stability), or tRNA function (aminoacylation, interaction with translation factors). Several reviews summarize the present view (Florentz and Sissler 2003; Rötig 2011; Ylikallio and Suomalainen 2012). In most cases, the effects of mutations are mild and affect either a single step of the tRNA life cycle or a combination of several of them. However, an initial impact is frequently observed on structural properties of affected tRNAs, followed by subsequent cascade effects on downstream functions (Florentz et al. 2003; Levinger et al. 2004; Wittenhagen and Kelley 2003). Therefore, any insight on the precise rules governing secondary and tertiary folding of the full set of human mt-tRNAs remains of high importance in order to comprehend the high sensitivity of these tRNAs to mutations perturbing their structure. Along these lines, pioneered experiments (Helm et al. 1998; Messmer et al. 2009), combined with the implementation of dedicated database (Pütz et al. 2007), and the development of bioinformatics tools (Bernt et al. 2012; Jühling et al. 2012a) are opening the path toward a solid knowledge on mt-tRNA 3D structures.

Finally, the housekeeping function of tRNAs, namely their capacity to become esterified by an amino acid, is not systematically affected in mutated variants, so that alternative functions of mt-tRNAs (Hou and Yang 2013; Mei et al. 2010) or alternative partnerships have to be considered (Jacobs and Holt 2000; Giegé et al. 2012).

3.5.2 Mitochondrial aaRSs and Pathologies

Lately, case-by-case reports linking mutations in nuclear genes coding for mitochondrial translation machinery proteins to pathologies (such as mutations in genes for elongation factor, tRNA modification enzymes, and ribosomal proteins), opened the way to “mitochondrial translation disorders” (Jacobs 2003). A new breakthrough took place in 2007 with the discovery in patients with cerebral white matter abnormalities of unknown origin of a first set of mutations present in *DARS2*, the nuclear gene coding for mt-AspRS (Scheper et al. 2007). These abnormalities were part of childhood-onset disorder called Leukoencephalopathy with Brain stem and Spinal cord involvement and Lactate elevation (LBSL; van der Knaap et al. 2003). Since this first discovery, mutations in eight additional mt-aaRS-encoding genes have been reported. They hit mt-ArgRS (Edvardson et al. 2007), mt-TyrRS (Riley et al. 2010), mt-SerRS (Belostotsky et al. 2011), mt-HisRS (Pierce et al. 2011), mt-AlaRS

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **72/83**

Editor Proof

(Götz et al. 2011), mt-MetRS (Bayat et al. 2012), mt-GluRS (Steenweg et al. 2012) and mt-PheRS (Elo et al. 2012) (Table 3.2). These recent correlations with human pathologies and the exponential description of reported cases, suggest as evidence that all mt-aaRS genes are likely affected by pathology-related mutations (that remain yet unveil), leading to a new family of disorders named according to the incriminated proteins namely "mt-aaRS disorders".

A detailed description and analysis of the full set of mutations in human mt-aaRS genes and their molecular and phenotypic implications has been reviewed (Konovalova and Tyynismaa 2013; Schwenzer et al. 2013). Here, the general outcomes are summarized. Table 3.2 recalls the main features characterizing the 65 nowadays-reported mutations in mt-aaRS genes. These include the type of pathogenic manifestation, familial pedigree, and affected tissues, as well as the number and types of mutations in each gene, their heterozygous versus homozygous, as well as the molecular impact on the synthetase and the final molecular impact on respiratory chain complexes. As a major outcome, it appears that whatever the mutation, no common combination of molecular steps correlates the mutations with the phenotypic expressions. Interestingly, the molecular impact of the mutations is not necessarily at the level of the housekeeping function of the synthetase, namely aminoacylation. Pathology-related mutations may have either a direct effect on the mitochondrial translation machinery by impacting one or several steps of mt-aaRS biogenesis and/or functioning. They may alternatively have an indirect effect by impacting ensuing steps and/or subsequent products activities [translation of the 13 mt-DNA-encoded subunits of respiratory chain complexes, respiratory chain complexes activities, and ATP synthesis]. Also, despite a dominant effect on brain and neuronal system is observed, sporadic manifestations are as well occurring in skeletal muscle, kidney, lung and/or heart. Along these lines, the selective vulnerability of tracts within the nervous system in case of mutations leading to splicing defects, for instance, is explained by tissue-specific differences in the concentration of the splicing factors (reduced in neural cell) (Edvardson et al. 2007; van Berge et al. 2012). However, the tissue-specificity of disorders remains an intriguing question. It is worth to establish the steady-state levels of various components of the mitochondrial translation machinery in different tissues, and correlate these levels with mitochondrial activity. This approach has been initiated by the evaluation of mRNA levels of the full set of human mt-aaRSs in 20 different human tissues (Fig. 3.5). A striking landscape of mRNA levels is observed highlighting tissue-specific differences by several orders of magnitude. There is no correlation between the various levels of mRNA and the amino acid content of the 13 mt-encoded proteins: leucine content is highest (14.4 %) followed by isoleucine, serine, and threonine (7 %), while arginine, aspartate, cysteine, glutamine, glutamate, and lysine contents is below 3 % (Schwenzer et al. 2013). We suggest that the low levels of aaRS mRNAs in brain, muscle and heart, lead to limiting mt translation activity in these tissues. Even a subtle change in mitochondrial translation efficiency may be detrimental in these tissues of high-energy demand. These data also suggest that mt-aaRS expressed to high levels may be involved in other functions than exclusively translation as is the case for cytosolic synthetases (Park et al. 2008).

UNCORRECTED PROOF

Editor Proof

Layout: T1 Standard SC Book ID: 307498_1_En Book ISBN: 978-3-642-39426-3
Chapter No.: 3 Date: 23-06-2013 Page: 73/83

Table 3.2 Human mt-aaRSs involved in mitochondrial disorders

	mt-AlaRS	mt-ArgRS	mt-AspRS	mt-GluRS	mt-HisRS	mt-MetRS	mt-pheRS	mt-SerRS	mt-TyrRS
Pathogenic manifestation[a]	CMP	PCH	LBSL	LBSL	PS	ARSAL	Encephatopathy	HUPRA	MLASA
Consanguinity of the parents	no	yes/no	yes/no	yes/no	no	no	yes/no	yes/no	yes/no
Affected tissues	Heart, brain, skeletal muscle	Brain	Brain, spinal cord	Brain	Ovarian sensori-neural system	Brain	Brain ± liver	Kidney, lung,	Blood, skelatal muscle
Number of mutations	2	10	27	15	3	?	4	1	2
Non sense (frameshift/stop)	0	1	8	3	0	0	0	0	0
Missense	2	7	14	11	2	0	4	1	2
Deletion/insertion	0	2	5	1	1	Large insertion	0	0	0
Other	0	0	0	0	0	Gene duplication	0	0	0
Genetics compound	Heterozygous	Heterozygous/ homozygous	Heterozygous/ homozygous	Heterozygous/ homozygous	Heterozygous	Heterozygous/ homozygous	Heterozygous/ homozygous	Homozygous	Homozygous
Negativly affected molecular event									
AaRS encoding mRNA expression/ processing	nd	yes/no	yes/no	nd	no	no	nd	nd	nd
AaRS expression/ stability	nd	yes/no	yes/no	nd	yes/no	yes	nd	nd	yes/no
AaRS import/ oligomerziation/ structure	nd	nd	yes/no	nd	yes/no	nd	yes/no	nd	no

(continued)

UNCORRECTED PROOF

Editor Proof

Layout: T1 Standard SC
Chapter No.: 3
Book ID: 307498_1_En
Date: 23-06-2013
Book ISBN: 978-3-642-39426-3
Page: 74/83

Table 3.2 (continued)

	mt-AlaRS	mt-ArgRS	mt-AspRS	mt-GluRS	mt-HisRS	mt-MetRS	mt-pheRS	mt-SerRS	mt-TyrRS
Aminoacylation activity	yes(predicated)	yes/no	yes/no	nd	yes	no	yes	yes	yes/no
Impact on the respiratory chain complex									
Global impact on translation	no	nd	no	yes	nd	yes	yes	nd	yes
Impact on RC activity[b]	yes	yes	yes	yes	nd	yes	yes	yes	yes

[a] Pathogenic manifestations are presented under the following acronyms. *CMP* Infantile Mitochondrial Cardimyopathy, *PCH* Ponto Cerebellar Hypoplasia, *LBSL* Leukoencephalopathy with Brain stem and Spinal cord involvement and Lactate elevation, *PS* Perrault Syndrome, *ARSAL* Autosomal Recessive Spastic Ataxia with Leukoencephalopathy, *HUPRA* HyperUricemia, Pulmonary hypertensions and Renal failure in infancy and Alkalosis, *MLASA* Myopathy, Lactic Acidosis and Sideroblastic Anemia. Consanguineous state of the parents and affected tissues are recalled. The number and type of mutations as well as their genetic compound are given. Molecular effects and impacts on the respiratory chain complexes are displayed

[b] All possible molecular effects on either aaRS biogenesis and/or function, or on translation and/or activity of the respiratory chain complexes, have not necessarily been investigated for all reported cases. For a more detailed view, please refer to (Schwenzer et al. 2013). For negatively affected molecular events, "yes" means that all tested mutations show this defect; "no" means that all tested mutations show no defect; "yes/no" means that some show a defect some do not. The table displays a mean picture: (i) a defect in mRNA expression and in aminoacylation does not necessarily refer to a same mutation; (ii) impacts on translation and aminoacylation activity can originate from separate mutations; and (iii) RC defects do not refer to tissue specificity: translation defect may correspond to, e.g., fibroblasts while aminoacylation activity was measured in muscle cells. *nd* stands for not determined

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **75/83**

Editor Proof

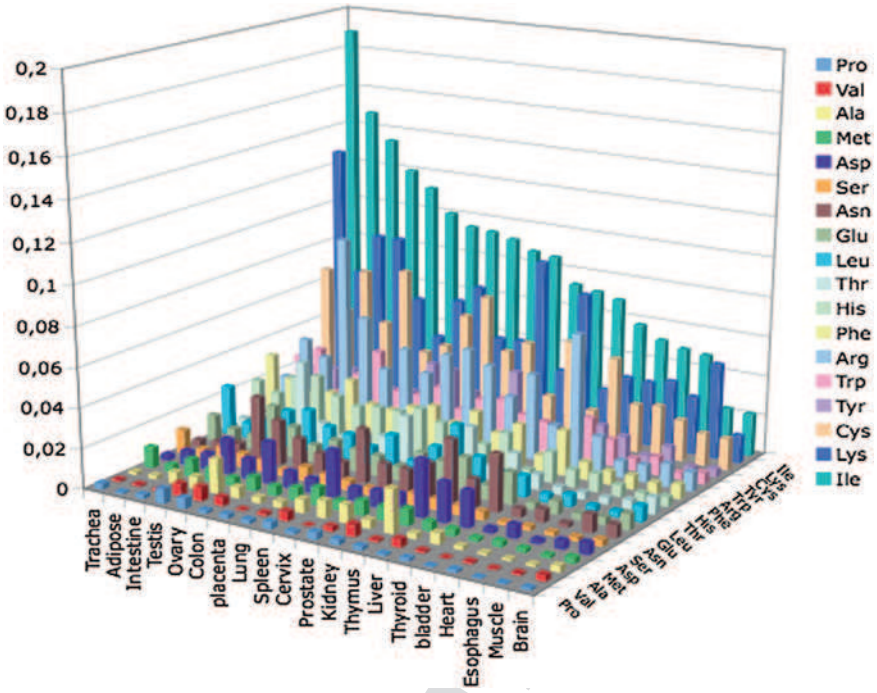


Fig. 3.5 Dosage of mt-aaRS-encoding mRNAs in different human tissues. The amount of mRNAs coding for 18 mt-aaRSs is determined by quantitative PCR on cDNAs prepared from total mRNA extracted from 20 human tissues. The mRNA for LysRS codes both for the cytosolic and mt enzyme. Results obtained for GlyRS, highly expressed, are missing from the graph (cytosolic and mt GlyRSs are both encoded by a single nuclear gene). No gene for mt-GlnRS has been reported so far. Values are normalized against the standard expression of GAPDH. They are mean values out of at least three independent experiments, from which the standard deviation is close to 50 %

UNCORRECTED

To conclude, links between the activity of a given mt-aaRS along mitochondrial translation on one hand and ATP production on the other hand, involved a number of issues that need to be further explored. Those issues should take also into account the possibility that aminoacylation may turn out to be not the sole function of mt-aaRSs in a living cell and that these enzymes may also participate in other processes and/or be implicated in various fine-tuning mechanisms as is the case for cytosolic aaRSs. Indeed, various bacterial and eukaryal aaRSs were found to have many additional functions (e.g., Guo and Schimmel 2013). It becomes thus necessary to determine all potential interacting components of mt-aaRSs and to study their dynamic location within the organelle. In other works, the functional network of mt-aaRSs and its regulation needs to be tackled. New routes toward understanding of the molecular impacts of point mutations in nuclear mt-aaRS genes outside the frame of mitochondrial translation should become opened along these lines.

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **76/83**

Editor Proof

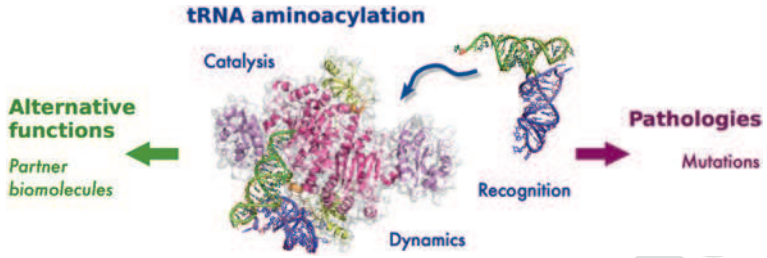


Fig. 3.6 Mammalian mitochondrial tRNAs and aminoacyl-tRNA synthetases are key macromolecules in mitochondrial translation. They are also hot spots for a growing number of human pathologies, some of which being related to defects in the housekeeping functions, some not. Numerous questions both on fundamental knowledge ("order") and toward understanding of the molecular events underlying pathologies ("disorder") remain open. Both stimulate the development of new research lines outside the strict frame of mitochondrial translation

3.6 Conclusion and Perspectives

Mammalian mt-tRNAs and mt-aaRSs are known as key actors of mitochondrial translation, leading to the synthesis of 13 essential mitochondrial inner membrane proteins and subunits of respiratory chain complexes. The fact that these two families of partner macromolecules are coded by either of the two cellular genomes, the nuclear genome coding for the mt-aaRSs, and the mitochondrial genome coding for the mt-tRNAs, highlights different, but however connected, evolutionary pathways between RNAs and partner proteins allowing for tRNA aminoacylation. Despite important structural degeneration of mt-tRNAs linked to the high mutation rate of the mitochondrial genome, the conservation of "canonical" structural properties of mt-aaRSs include subtle but strong molecular adaptation so that the partnership between both macromolecules is maintained for accurate mitochondrial translation (Fig. 3.6). Detailed analysis of mammalian mt-tRNAs revealed new structural rules for RNAs, and bioinformatics compilations confirmed and enlarged the structural diversity of mt-tRNAs to the shortest versions ever discovered, calling for additional structural adaptation of functional RNAs. Investigation of mammalian mt-tRNAs and aaRSs are however still in an initial stage. Only a few systems have been characterized along a limited number of aspects. Open questions include for example, the common structural properties of mt-tRNAs and of synthetases all along metazoan mitochondria. Which are the identity elements in mt-tRNAs for specific aminoacylation by cognate mt-aaRS? Are these conserved or idiosynchratic? How far can a given set be degenerated and still allow for specificity? Post-transcriptional modification patterns of tRNAs and possible post-translational modifications of synthetases remain to become determined on full scale. The importance of these modifications is of crucial interest not only in structural stabilization and in the housekeeping codon reading function of mt-tRNAs, but also in alternative functions of mt-aaRSs. Post-translational modifications are indeed triggers to alternative functions of cytosolic aaRSs (e.g., Kim et al.

UNCORRECTED PRO

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **77/83**

Editor Proof

2012; Ofir-Birin et al. 2013). The related question of alternate function of mt-aaRSs is of key importance, especially in regard of understanding the molecular mechanisms of related disorders. Such functions, unrelated to aminoacylation, correspond to an emerging field of discoveries for cytosolic aaRSs. Examples include TyrRS involved in receptor-mediated signaling pathways associated with angiogenesis (Wakasugi et al. 2002a), TrpRS activated as an angiostatic factor (Wakasugi et al. 2002b), Glu-ProRS involved in the inflammatory response (Jia et al. 2008), LysRS plays a role in HIV-I packaging (Kleiman and Cen 2004) of GlyRS as anti-tumorigenic agent (Park et al. 2012).

The existence of an always growing panel of human disorders correlated to point mutations in either mt-tRNA genes or mt-aaRS genes, leads to investigations as to the molecular impacts of the mutations (Fig. 3.6). While a number of approaches pinpointed a mosaicity of impacts on human mt-tRNA structure and function, and more recently on biophysical and functional properties of a small set of mt-aaRS, all linked to the housekeeping activity of both macromolecules in mitochondrial translation, a new field of investigation is emerging. Several of the pathology-related mutations are not located in the catalytic site of the protein, do not affect protein synthesis, and are thus indicative of new properties of mt-tRNAs and mt-aaRSs outside translation. Some hints on alternative functions have already been reported. Any route along this new topic deserves interest and should now be considered, as is currently the case for cytosolic tRNAs and synthetases (Guo and Schimmel 2013; Guo et al. 2010a, 2010b). Importantly, many other actors of the mammalian mitochondrial translation machinery also remain to be explored more systematically, not only for fundamental and evolutionary understanding but also because of their growing implication in human pathologies (e.g., Rötig 2011; Taylor and Turnbull 2005; Watanabe 2010).

Acknowledgments We thank Richard Giegé for critical reading of the manuscript and Gert Scheper and Koen de Groot for help in the qPCR experiments. Numerous contributions on mt-tRNAs and aaRS could not be mentioned because of space limitation and we apologize for this. Financial support came from Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg, ANR MITOMOT (ANR-09-BLAN-0091-01/03), French National Program 'Investissements d'Avenir' (Labex MitoCross) administered by the "Agence National de la Recherche", and referenced ANR-10-IDEX-002-02; French-German PROCOPE program (DAAD D/0628236, EGIDE PHC 14770PJ), and German Academic Exchange Service (DAAD D/10/43622) for a doctoral fellowship.

References

Alfonzo JD, Söll D (2009) Mitochondrial tRNA import—the challenge to understand has just begun. Biol Chem 390:717–722

Anderson S, Bankier AT, Barrel BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon JC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Arcari P, Brownlee GG (1980) The nucleotide sequence of a small (3S) seryl-tRNA (anticodon GCU) from beef heart mitochondria. Nucleic Acids Res 8:5207–5212

Bayat V, Thiffault I, Jaiswal M, Tétreault M, Donti T, Sasarman F, Bernard G, Demers-Lamarche J, Dicaire MJ, Mathieu J, Vanasse M, Bouchard JP, Rioux MF, Lourenco CM, Li Z, Haueter

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **78/83**

Editor Proof

C, Shoubridge EA, Graham BH, Brais B, Bellen HJ (2012) Mutations in the mitochondrial methionyl-tRNA synthetase cause a neurodegenerative phenotype in flies and a recessive ataxia (ARSAL) in humans. PLoS Biol 10:e1001288

Belostotsky R, Ben-Shalom E, Rinat C, Becker-Cohen R, Feinstein S, Zeligson S, Segel R, Elpeleg O, Nassar S, Frishberg Y (2011) Mutations in the mitochondrial seryl-tRNA synthetase cause hyperuricemia, pulmonary hypertension, renal failure in infancy and alkalosis, HUPRA syndrome. Am J Hum Genet 88:193–200

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2012) MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol [Epub ahead of print]

Bonnefond L, Fender A, Rudinger-Thirion J, Giegé R, Florentz C, Sissler M (2005a) Toward the full set of human mitochondrial aminoacyl-tRNA synthetases: characterization of AspRS and TyrRS. Biochemistry 44:4805–4816

Bonnefond L, Frugier M, Giegé R, Rudinger-Thirion J (2005b) Human mitochondrial TyrRS disobeys the tyrosine idenity rules. RNA 11:558–562

Bonnefond L, Frugier M, Touzé E, Lorber B, Florentz C, Giegé R, Sauter C, Rudinger-Thirion J (2007) Crystal structure of human mitochondrial tyrosyl-tRNA synthetase reveals common and idiosyncratic features. Structure 15:1505–1516

Braband A, Cameron SL, Podsiadlowski L, Daniels SR, Mayer G (2010) The mitochondrial genome of the onychophoran Opisthopatus cinctipes (Peripatopsidae) reflects the ancestral mitochondrial gene arrangement of Panarthropoda and Ecdysozoa. Mol Phylogenet Evol 57:285–292

Brindefalk B, Viklund J, Larsson D, Thollesson M, Andersson SG (2007) Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases. Mol Biol Evol 24:743–756

Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 76:1967–1971

Bullard J, Cai Y-C, Spremulli L (2000) Expression and characterization of the human mitochondrial leucyl-tRNA synthetase. Biochem Biophys Acta 1490:245–258

Cader MZ, Ren J, James PA, Bird LE, Talbot K, Stammers DK (2007) Crystal structure of human wildtype and S581L-mutant glycyl-tRNA synthetase, an enzyme underlying distal spinal muscular atrophy. FEBS Lett 581:2959–2964

Castellana S, Vicario S, Saccone C (2011) Evolutionary patterns of the mitochondrial genome in Metazoa: exploring the role of mutation and selection in mitochondrial protein coding genes. Genome Biol Evol 3:1067–1079

Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N (2009) Importing mitochondrial proteins: machineries and mechanisms. Cell 1387:628–644

Chimnaronk S, Gravers Jeppesen M, Suzuki T, Nyborg J, Watanabe K (2005) Dual-mode recognition of noncanonical tRNAs(Ser) by seryl-tRNA synthetase in mammalian mitochondria. EMBO J 24:3369–3379

de Bruijn MH, Schreier PH, Eperon IC, Barrell BG, Chen EY, Armstrong PW, Wong JF, Roe BA (1980) A mammalian mitochondrial serine transfer RNA lacking the "dihydrouridine" loop and stem. Nucleic Acids Res 8:5213–5222

Dörner M, Altmann M, Pääbo S, Mörl M (2001) Evidence for import of a lysyl-tRNA into marsupial mitochondria. Mol Biol Cell 12:2688–2698

Duchêne AM, Pujol C, Maréchal-Drouard L (2009) Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet 55:1–18

Edvardson S, Shaag A, Kolesnikova O, Gomori JM, Tarassov I, Einbinder T, Saada A, Elpeleg O (2007) Deleterious mutation in the mitochondrial arginyl-transfer RNA synthetase gene is associated with pontocerebellar hypoplasia. Am J Hum Genet 81:857–862

Elo JM, Yadavalli SS, Euro L, Isohanni P, Götz A, Carroll CJ, Valanne L, Alkuraya FS, Uusimaa J, Paetau A, Caruso EM, Pihko H, Ibba M, Tyynismaa H, Suomalainen A (2012) Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy. Hum Mol Genet 21:4521–4529

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **79/83**

Editor Proof

Enriquez JA, Attardi G (1996) Analysis of aminoacylation of human mitochondrial tRNAs. Methods Enzymol 264:183–196

Fender A, Gaudry A, Jühling F, Sissler M, Florentz C (2012) Adaptation of aminoacylation rules to mammalian mitochondria. Biochimie 94:1090–1097

Fender A, Sauter C, Messmer M, Pütz J, Giegé R, Florentz C, Sissler M (2006) Loss of a primordial identity element for a mammalian mitochondrial aminoacylation system. J Biol Chem 281:15980–15986

Florentz C, Sissler M (2001) Disease-related *versus* polymorphic mutations in human mitochondrial tRNAs: where is the difference? EMBO Rep 2(6):481–486

Florentz C, Sissler M (2003) Mitochondrial tRNA aminoacylation and human diseases. In: Lapointe J, Brakier-Gingras L (eds) Translation mechanisms. Landes Bioscience, Georgetown, pp 129–143

Florentz C, Sohm B, Tryoen-Tóth P, Pütz J, Sissler M (2003) Human mitochondrial tRNAs in health and disease. Cell Mol Life Sci 60:1356–1375

Frechin M, Duchêne A-M, Becker HD (2009a) Translating organellar glutamine codons: A case by case scenario? RNA Biol 6:31–34

Frechin M, Senger B, Brayé M, Kern D, Martin RP, Becker HD (2009b) Yeast mitochondrial Gln-tRNA(Gln) is generated by a GatFAB-mediated transamidation pathway involving Arc1p-controlled subcellular sorting of cytosolic GluRS. Genes Dev 23:1119–1130

Friederich MW, Hagerman PJ (1997) The angle between the anticodon and aminoacyl acceptor stems of yeast tRNA(Phe) is strongly modulated by magnesium ions. Biochemistry 36:6090–6099

Gaudry A, Lorber B, Messmer M, Neuenfeldt A, Sauter C, Florentz C, Sissler M (2012) Redesigned N-terminus enhances expression, solubility, and crystallisability of mitochondrial enzyme. Protein Eng Des Sel 25:473–481

Giegé R (2008) Toward a more complete view of tRNA biology. Nat Struct Mol Biol 15:1007–1014

Giegé R, Florentz C, Kern D, Gangloff J, Eriani G, Moras D (1996) Aspartate identity of transfer RNAs. Biochimie 78:605–623

Giegé R, Jühling F, Pütz J, Stadler P, Sauter C, Florentz C (2012) Structure of transfer RNAs: similarity and variability. Wiley Interdiscip Rev RNA 3:37–61

Giegé R, Sissler M, Florentz C (1998) Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Res 26:5017–5035

Goto Y, Nonaka I, Horai S (1990) A mutation in the $tRNA^{Leu(UUR)}$ gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. Nature 348:651–653

Götz A, Tyynismaa H, Euro L, Ellonen P, Hyötyläinen T, Ojala T, Hämäläinen RH, Tommiska J, Raivio T, Oresic M, Karikoski R, Tammela O, Simola KO, Paetau A, Tyni T, Suomalainen A (2011) Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. Am J Hum Genet 88:635–642

Guo M, Ignatov M, Musier-Forsyth K, Schimmel P, Yang XL (2008) Crystal structure of tetrameric form of human lysyl-tRNA synthetase: Implications for multisynthetase complex formation. Proc Natl Acad Sci USA 105:2331–2336

Guo M, Schimmel P (2013) Essential nontranslational functions of tRNA synthetases. Nat Chem Biol 9:145–153

Guo M, Schimmel P, Yang X-L (2010a) Functional expansion of human tRNA synthetases achieved by structural inventions. FEBS Lett 584:434–442

Guo M, Yang XL, Schimmel P (2010b) New functions of aminoacyl-tRNA synthetases beyond translation. Nat Rev Mol Cell Biol 11:668–674

Haen KM, Pett W, Lavrov DV (2010) Parallel loss of nuclear-encoded mitochondrial aminoacyl-tRNA synthetases and mtDNA-encoded tRNAs in Cnidaria. Mol Biol Evol 27:2216–2219

Helm M, Attardi G (2004) Nuclear control of cloverleaf structure of human mitochondrial tRNA(Lys). J Mol Biol 337:545–560

Helm M, Brulé H, Degoul F, Cepanec C, Leroux J-P, Giegé R, Florentz C (1998) The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. Nucleic Acids Res 26:1636–1643

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **80/83**

Editor Proof

Helm M, Brulé H, Friede D, Giegé R, Pütz J, Florentz C (2000) Search for characteristic structural features of mammalian mitochondrial tRNAs. RNA 6:1356–1379

Helm M, Florentz C, Chomyn A, Attardi G (1999) Search for differences in post-transcriptional modification patterns of mitochondrial DNA-encoded wild-type and mutant human $tRNA^{Lys}$ and $tRNA^{Leu(UUR)}$. Nucleic Acids Res 27:756–763

Hou YM, Yang X (2013) Regulation of cell death by transfer RNA. Antioxid Redox Signal [Epub ahead of print]

Ibba M, Francklyn C, Cusack S (2005) The aminoacyl-tRNA synthetases. Landes Biosciences, Georgetown

Jacobs HT, Holt IJ (2000) The np 3243 MELAS mutation: damned if you aminoacylate, damned if you don't. Hum Mol Genet 1:463–465

Jacobs HT (2003) Disorders of mitochondrial protein synthesis. Hum Mol Genet 12:R293–301

Jia J, Arif A, Ray PS, Fox PL (2008) WHEP domains direct noncanonical function of glutamyl-Prolyl-tRNA synthetase in translational control of gene expression. Mol Cell 29:679–690

Jühling F, Pütz J, Bernt M, Donath A, Middendorf M, Florentz C, Stadler PF (2012a) Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. Nucleic Acids Res 40:2833–2845

Jühling F, Pütz J, Florentz C, Stadler PF (2012b) Armless mitochondrial tRNAs in Enoplea (Nematoda). RNA Biol 9:1161–1166

Kim DG, Choi JW, Lee JY, Kim H, Oh YS, Lee JW, Tak YK, Song JM, Razin E, Yun SH, Kim S (2012) Interaction of two translational components, lysyl-tRNA synthetase and p40/37LRP, in plasma membrane promotes laminin-dependent cell migration. FASEB J 26:4142–4159

Kirino Y, Goto Y, Campos Y, Arenas J, Suzuki T (2005) Specific correlation between the wobble modification deficiency in mutant tRNAs and the clinical features of a human mitochondrial disease. Proc Natl Acad Sci USA 102:7127–7132

Kleiman L, Cen S (2004) The tRNALys packaging complex in HIV-1. Int J Biochem Cell Biol 36:1776–1786

Klipcan L, Levin I, Kessler N, Moor N, Finarov I, Safro M (2008) The tRNA-induced conformational activation of human mitochondrial phenylalanyl-tRNA synthetase. Structure 16:1095–1104

Klipcan L, Moor N, Finarov I, Kessler N, Sukhanova M, Safro MG (2012) Crystal structure of human mitochondrial PheRS complexed with tRNA(Phe) in the active "open" state. J Mol Biol 415:527–537

Konovalova S, Tyynismaa H (2013) Mitochondrial aminoacyl-tRNA synthetases in human disease. Mol Genet Metab [Epub ahead of print]

Kumazawa Y, Himeno H, Miura K, Watanabe K (1991) Unilateral aminoacylation specificity between bovine mitochondria and eubacteria. J Biochem 109:421–427

LaRiviere FJ, Wolfson AD, Uhlenbeck OC (2001) Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. Science 294:165–168

Laslett D, Canbäck B (2008) ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24:172–175

Leontis NB, Stombaugh J, Westhof E (2002) The non-Watson-Crick base pairs and their associated isostericity matrices. Nucleic Acids Res 30:3497–3531

Levinger L, Mörl M, Florentz C (2004) Mitochondrial tRNA 3′ end metabolism and human disease. Nucleic Acids Res 32:5430–5441

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

Macey JR, Larson A, Ananjeva NB, Papenfuss TJ (1997) Replication slippage may cause parallel evolution in the secondary structures of mitochondrial transfer RNAs. Mol Biol Evol 14:30–39

McFarland R, Elson JL, Taylor RW, Howell N, Turnbull DM (2004) Assigning pathogenicity to mitochondrial tRNA mutations: when 'definitely maybe' is not good enough. Trends Genet 20:591–596

Mei Y, Yong J, Liu H, Shi Y, Meinkoth J, Dreyfuss G, Yang X (2010) tRNA binds to cytochrome c and inhibits caspase activation. Mol Cell 37:688–698

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **81/83**

Editor Proof

Messmer M, Pütz J, Suzuki T, Suzuki T, Sauter C, Sissler M, Florentz C (2009) Tertiary network in mammalian mitochondrial tRNAAsp revealed by solution probing and phylogeny. Nucleic Acids Res 37:6881–6895

Miyamoto H, Machida RJ, Nishida S (2010) Complete mitochondrial genome sequences of the three pelagic chaetognaths Sagitta nagae, Sagitta decipiens and Sagitta enflata. Comp Biochem Physiol Part D Genomics Proteomics 5:65–72

Motorin Y, Helm M (2010) tRNA stabilization by modified nucleotides. Biochemistry 49:4934–4944

Mudge SJ, Williams JH, Eyre HJ, Sutherland GR, Cowan PJ, Power DA (1998) Complex organisation of the 5′-end of the human glycine tRNA synthetase gene. Gene 209:45–50

Nagao A, Suzuki T, Katoh T, Sakaguchi Y, Suzuki T (2009) Biogenesis of glutaminyl-mt tRNAGln in human mitochondria. Proc Natl Acad Sci USA 106:16209–16214

Nagao A, Suzuki T, Suzuki T (2007) Aminoacyl-tRNA surveillance by EF-Tu in mammalian mitochondria. Nucleic Acids Symp Ser (Oxf) 51:41–42

Nawroki EP, Kolbe DL, Eddy SR (2009) Infernal 1.0: Inference of RNA Alignments. Bioinformatics 25:1335–1337

Neuenfeldt A, Lorber B, Ennifar E, Gaudry A, Sauter C, Sissler M, Florentz C (2013) Thermodynamic properties distinguish human mitochondrial aspartyl-tRNA synthetase from bacterial homolog with same 3D architecture. Nucleic Acids Res 41:2698–2708

Ofir-Birin Y, Fang P, Bennett SP, Zhang HM, Wang J, Rachmin I, Shapiro R, Song J, Dagan A, Pozo J, Kim S, Marshall AG, Schimmel P, Yang XL, Nechushtan H, Razin E, Guo M (2013) Structural switch of lysyl-tRNA synthetase between translation and transcription. Mol Cell 49:30–42

Park MC, Kang T, Jin D, Han JM, Kim SB, Park YJ, Cho K, Park YW, Guo M, He W, Yang XL, Schimmel P, Kim S (2012) Secreted human glycyl-tRNA synthetase implicated in defense against ERK-activated tumorigenesis. Proc Natl Acad Sci USA 109:E640–E647

Park SG, Schimmel P, Kim S (2008) Aminoacyl tRNA synthetases and their connections to disease. Proc Natl Acad Sci USA 105:11043–11049

Pierce SB, Chisholm KM, Lynch ED, Lee MK, Walsh T, Opitz JM, Li W, Klevit RE, King MC (2011) Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. Proc Natl Acad Sci USA 108:6543–6548

Pruitt KD, Tatusova T, Maglott DR (2007) NCBI reference sequences (RefSeq): a curated nonredundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 5(Database issue):D61–D65

Pütz J, Dupuis B, Sissler M, Florentz C (2007) Mamit-tRNA, a database of mammalian mitochondrial tRNA primary and secondary structures. RNA 13:1184–1190

Riley LG, Cooper S, Hickey P, Rudinger-Thirion J, McKenzie M, Compton A, Lim SC, Thorburn D, Ryan MT, Giegé R, Bahlo M, Christodoulou J (2010) Mutation of the mitochondrial tyrosyl-tRNA synthetase gene, YARS2, causes myopathy, lactic acidosis, and sideroblastic anemia–MLASA syndrome. Am J Hum Genet 87:52–59

Rinehart J, Krett B, Rubio M-AT, Alfonzo JD, Söll D (2005) *Saccharomyces cerevisiae* imports the cytosolic pathway for Gln-tRNA synthesis into the mitochondion. Genes Dev 19:583–592

Rötig A (2011) Human diseases with impaired mitochondrial protein synthesis. Biochim Biophys Acta 1807:1198–1205

Scheper GC, van der Klok T, van Andel RJ, van Berkel CG, Sissler M, Smet J, Muravina TI, Serkov SV, Uziel G, Bugiani M, Schiffmann R, Krageloh-Mann I, Smeitink JA, Florentz C, Coster RV, Pronk JC, van der Knaap MS (2007) Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. Nat Genet 39:534–539

Schwenzer H, Zoll J, Florentz C, Sissler M (2013) Pathogenic implications of human mitochondrial aminoacyl-tRNA synthetases. In: KIM (ed) Topics in current chemistry-aminoacyl-tRNA synthetases: Applications in chemistry, Biology and Medicine. Springer (in press)

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **82/83**

Editor Proof

878 Segovia R, Pett W, Trewick S, Lavrov DV (2011) Extensive and evolutionarily persistent mitochon-
879 drial tRNA editing in Velvet Worms (phylum Onychophora). Mol Biol Evol 28:2873–2881
880 Seutin G, Lang BF, Mindell DP, Morais R (1994) Evolution of the WANCY region in amniote
881 mitochondrial DNA. Mol Biol Evol 11:329–340
882 Shiba K, Schimmel P, Motegi H, Noda T (1994) Human glycyl-tRNA synthetase. Wide diver-
883 gence of primary structure from bacterial counterpart and species-specific aminoacylation. J
884 Biol Chem 269:30049–30055
885 Shoffner J, Lott M, Lezza AMS, Seibel P, Ballinger SW, Wallace DC (1990) Myoclonic epilepsy
886 and ragged red fiber disease (MERRF) is associated with mitochondrial DNA $tRNA^{Lys}$
887 mutation. Cell 61:931–937
888 Sissler M, Pütz J, Fasiolo F, Florentz C (2005) Mitochondrial aminoacyl-tRNA synthetases. In:
889 Ibba M, Francklyn C, Cusack S (eds), Aminoacyl-tRNA synthetases, chapter 24, pp 271–
890 284. Landes Biosciences, Georgetown
891 Sohm B, Frugier M, Brulé H, Olszak K, Przykorska A, Florentz C (2003) Towards understanding
892 human mitochondrial leucine aminoacylation identity. J Mol Biol 328:995–1010
893 Steenweg ME, Ghezzi D, Haack T, Abbink TE, Martinelli D, van Berkel CG, Bley A, Diogo L,
894 Grillo E, Te Water Naudé J, Strom TM, Bertini E, Prokisch H, van der Knaap MS, Zeviani
895 M (2012) Leukoencephalopathy with thalamus and brainstem involvement and high lactate
896 'LTBL' caused by EARS2 mutations. Brain 135:1387–1394
897 Suga K, Mark Welch DB, Tanaka Y, Sakakura Y, Hagiwara A (2008) Two circular chromosomes
898 of unequal copy number make up the mitochondrial genome of the rotifer Brachionus plica-
899 tilis. Mol Biol Evol 25:1129–1137
900 Suzuki T, Nagao A, Suzuki T (2011) Human mitochondrial tRNAs: biogenesis, function, struc-
901 tural aspects, and diseases. Annu Rev Genet 45:299–329
902 Taylor RW, Turnbull DM (2005) Mitochondrial DNA mutations in human disease. Nat Rev
903 Genet 6:389–402
904 Tolkunova E, Park H, Xia J, King MP, Davidson E (2000) The human lysyl-tRNA synthetase
905 gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual
906 splicing of the primary transcript. J Biol Chem 275:35063–35069
907 van Berge L, Dooves S, van Berkel CG, Polder E, van der Knaap MS, Scheper GC (2012)
908 Leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation is
909 associated with cell-type-dependent splicing of mtAspRS mRNA. Biochem J 441:955–962
910 van der Knaap MS, van der Voorn P, Barkhof F, Van Coster R, Krägeloh-Mann I, Feigenbaum A,
911 Blaser S, Vles JS, Rieckmann P, Pouwels PJ (2003) A new leukoencephalopathy with brain-
912 stem and spinal cord involvement and high lactate. Ann Neurol 53:252–258
913 Wakasugi K, Slike BM, Hood J, Ewalt KL, Cheresh DA, Schimmel P (2002a) Induction of angi-
914 ogenesis by a fragment of human tyrosyl-tRNA synthetase. J Biol Chem 277:20124–20126
915 Wakasugi K, Slike BM, Hood J, Otani A, Ewalt KL, Friedlander M, Cheresh DA, Schimmel P
916 (2002b) A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. Proc Natl
917 Acad Sci USA 99:173–177
918 Wakita K, Watanabe Y-I, Yokogawa T, Kumazawa Y, Nakamura S, Ueda T, Watanabe K,
919 Nishikawa K (1994) Higher-order structure of bovine mitochondrial $tRNA^{Phe}$ lacking the
920 'conserved' GG and TYCG sequences as inferred by enzymatic and chemical probing.
921 Nucleic Acids Res 22:347–353
922 Wang X, Lavrov DV (2008) Seventeen new complete mtDNA sequences reveal extensive mito-
923 chondrial genome evolution within the Demospongiae. PLoS ONE 3:e2723
924 Watanabe K (2010) Unique features of animal mitochondrial translation systems. The non-uni-
925 versal genetic code, unusual features of the translational apparatus and their relevance to
926 human mitochondrial diseases. Proc Jpn Acad Ser B Phys Biol Sci 86:11–36
927 Willkomm DK, Hartmann RK (2006) Intricacies and surprises of nuclear-mitochondrial co-evo-
928 lution. Biochem J 399:e7–e9
929 Wittenhagen LM, Kelley SO (2003) Impact of disease-related mitochondrial mutations on tRNA
930 structure and function. Trends Biochem Sci 28:605–611

UNCORRECTED PROOF

Layout: **T1 Standard SC** | Book ID: **307498_1_En** | Book ISBN: **978-3-642-39426-3**
Chapter No.: **3** | Date: **23-06-2013** | Page: **83/83**

Editor Proof

931 Woese CR, Olsen GJ, Ibba M, Söll D (2000) Aminoacyl-tRNA synthetases, the genetic code, and
932 the evolutionary process. Microbiol and Mol Biol Reviews 64:202–236
933 Wolstenholme DR, Okimoto R, Mcfarlane JL (1994) Nucleotide correlations that suggest tertiary
934 interactions in the TV-replacement loop-containing mitochondrial tRNAs of the nematodes,
935 *Caenorhabditis elegans* and *Ascaris suum*. Nucleic Acids Res 22:4300–4306
936 Xie W, Schimmel P, Yang XL (2006) Crystallization and preliminary X-ray analysis of a native
937 human tRNA synthetase whose allelic variants are associated with Charcot-Marie-Tooth
938 disease. Acta Crystallograph Sect F Struct Biol Cryst Commun 62:1243–1246
939 Yadavalli SS, Klipcan L, Zozulya A, Banerjee R, Svergun D, Safro M, Ibba M (2009) Large-
940 scale movement of functional domains facilitates aminoacylation by human mitochondrial
941 phenylalanyl-tRNA synthetase. FEBS Lett 583:3204–3208
942 Yao YN, Wang L, Wu XF, Wang ED (2003) The processing of human mitochondrial leucyl-tRNA
943 synthetase in the insect cells. FEBS Lett 534:139–142
944 Yarham JW, Al-Dosary M, Blakely EL, Alston CL, Taylor RW, Elson JL, McFarland R (2011)
945 A comparative analysis approach to determining the pathogenicity of mitochondrial tRNA
946 mutations. Hum Mutat 32:1319–1325
947 Yarham JW, Elson JL, Blakely EL, McFarland R, Taylor RW (2010) Mitochondrial tRNA muta-
948 tions and disease. Wiley Interdiscip Rev RNA 1:304–324
949 Ylikallio E, Suomalainen A (2012) Mechanisms of mitochondrial diseases. Ann Med 44:41–59

UNCORRECTED PROOF

Book Chapter #2

Pathogenic Implications of Human Mitochondrial Aminoacyl-tRNA Synthetases

Hagen Schwenzer, Joffrey Zoll, Catherine Florentz, and Marie Sissler

2013

Published in: Top Curr Chem Springer-Verlag Berlin Heidelberg 2013
DOI: 10.1007/128_2013_457 #
(*in press*)

Top Curr Chem (2013)
DOI: 10.1007/128_2013_457
© Springer-Verlag Berlin Heidelberg 2013

Pathogenic Implications of Human Mitochondrial Aminoacyl-tRNA Synthetases

Hagen Schwenzer, Joffrey Zoll, Catherine Florentz, and Marie Sissler

Abstract Mitochondria are considered as the powerhouse of eukaryotic cells. They host several central metabolic processes fueling the oxidative phosphorylation pathway (OXPHOS) that produces ATP from its precursors ADP and inorganic phosphate Pi (PPi). The respiratory chain complexes responsible for the OXPHOS pathway are formed from complementary sets of protein subunits encoded by the nuclear genome and the mitochondrial genome, respectively. The expression of the mitochondrial genome requires a specific and fully active translation machinery from which aminoacyl-tRNA synthetases (aaRSs) are key actors. Whilst the macromolecules involved in mammalian mitochondrial translation have been under investigation for many years, there has been an explosion of interest in human mitochondrial aaRSs (mt-aaRSs) since the discovery of a large (and growing) number of mutations in these genes that are linked to a variety of neurodegenerative disorders. Herein we will review the present knowledge on mt-aaRSs in terms of their biogenesis, their connection to mitochondrial respiration, i.e., the respiratory chain (RC) complexes, and to the mitochondrial translation machinery. The pathology-related mutations detected so far are described, with special attention given to their impact on mt-aaRSs biogenesis, functioning, and/or subsequent

Note: Rigorously, amino acid conversion of a given mutation should be preceded by the letter "p." to indicate that the protein level is considered. For example, the 172C > G nucleotide change engenders the p.R58G mutation in *DARS2* (referencing the gene) or mt-AspRS (referencing the protein). For sake of simplicity, the "p." is omitted throughout the chapter.

H. Schwenzer, C. Florentz, and M. Sissler (✉)
Architecture et Réactivité de l'ARN, CNRS, Université de Strasbourg, IBMC,
15 rue René Descartes, 67084 Strasbourg Cedex, France
e-mail: H.Schwenzer@ibmc-cnrs.unistra.fr; C.Florentz@ibmc-cnrs.unistra.fr;
M.Sissler@ibmc-cnrs.unistra.fr

J. Zoll
EA 3072, Université de Strasbourg, Faculté de médecine, 4 rue Kirschleger,
67085 Strasbourg, France
e-mail: Joffrey.zoll@unistra.fr

activities. The collected data to date shed light on the diverse routes that are linking primary molecular possible impact of a mutation to its phenotypic expression. It is envisioned that a variety of mechanisms, inside and outside the translation machinery, would play a role on the heterogeneous manifestations of mitochondrial disorders.

Keywords Aminoacyl-tRNA synthetase · Human mitochondrial disorders · Pathology-related mutations · Respiratory chain defects

Contents

1 Mt-aaRSs and Mitochondrial ATP Synthesis
  1.1 Mitochondrial Respiratory Chain Complexes
  1.2 The Human Mitochondrial Translation Machinery
  1.3 Link Between Mitochondrial Translation, Mitochondrial Respiration, and Mitochondrial Disorders
  1.4 Biochemical Analysis of Mitochondrial Respiration: A Potential Diagnostic Tool for the Detection of Mitochondrial Translation Defects
2 Mt-aaRSs in Mitochondrial Translation
  2.1 Nuclear-Encoded aaRSs of Mitochondrial Location and Evolutionary Considerations
  2.2 Mt-aaRSs are Imported Proteins
  2.3 Structural Insights of mt-aaRSs
  2.4 Some Functional Peculiarities
3 Mt-aaRSs in Human Disorders
  3.1 Discovery of aaRS-Related Disorders
  3.2 The Present-Day List of Pathology-Related Mutations Within mt-aaRS Encoding Genes
  3.3 Compound Heterozygous vs. Homozygous States
4 Diverse Molecular Impacts
  4.1 Impact of Mutations on mt-aaRSs Biogenesis
  4.2 Impact of Mutations on mt-aaRSs Function
  4.3 Impact of Mutations on Mitochondrial Translation and Activity of the Respiratory Chain Complexes
5 Outlook
References

Abbreviations

AaRS	Aminoacyl-tRNA synthetase (specificity is indicated by the name of the amino acid (abbreviated in a three-letter code) transferred to the cognate tRNA. As an example, AspRS stands for aspartyl-tRNA synthetase)
mt	Mitochondrial
MTS	Mitochondrial targeting sequence
RC	Respiratory chain

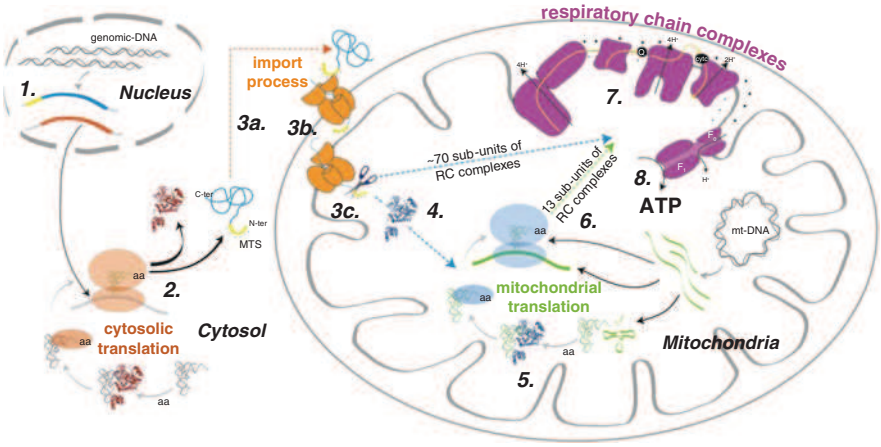


Fig. 1 From mitochondrial aminoacyl-tRNA synthetases to mitochondrial ATP synthesis. The route from the place of encoding of mt-aaRSs (the nucleus) to their place of biosynthesis (the cytosol) and their place of use (the mitochondria) is schematized. Mt-aaRSs biogenesis comprises mRNAs expression and processing (*1*), mt-aaRSs synthesis (*2*), import process into mitochondria (*3a* addressing; *3b* translocation; *3c* processing), and proper folding, oligomerization, and stability upon entry to mitochondria (*4*). Mt-aaRSs functioning includes amino acid activation, tRNA recognition, tRNA charging (*5*). Mt-aaRSs are devoted to the mitochondrial translation, and thus, to the synthesis of the 13 mt-DNA-encoded respiratory chain (RC) complexes (*6*), for which the activity (*7*) ultimately lead to ATP production (*8*). Of note, all other sub-units of the RC complexes (~70) are also imported from the cytosol

1 Mt-aaRSs and Mitochondrial ATP Synthesis

In order to facilitate the understanding of the connection between human mt-aaRSs and ATP synthesis, the route from their place of encoding (the nucleus) to their place of biogenesis (the cytosol) and their place of use (the mitochondria) is schematized in Fig. 1.

1.1 Mitochondrial Respiratory Chain Complexes

One of the most prominent functions of mitochondria is the production of cellular free energy in the form of ATP, in a process known as oxidative phosphorylation (OXPHOS). This process takes place in five large multi-subunit complexes (the respiratory chain complexes) that are located in the inner mitochondrial membrane (Fig. 2a). Complexes I to IV, accompanied by the mobile elements Coenzyme Q and cytochrome *c*, allow for the activity of Complex V, the ATP synthase. In mitochondria, the final oxidation of nutrients releases CO_2 (mainly through the Krebs cycle) concomitantly with the reduction of NAD^+ into

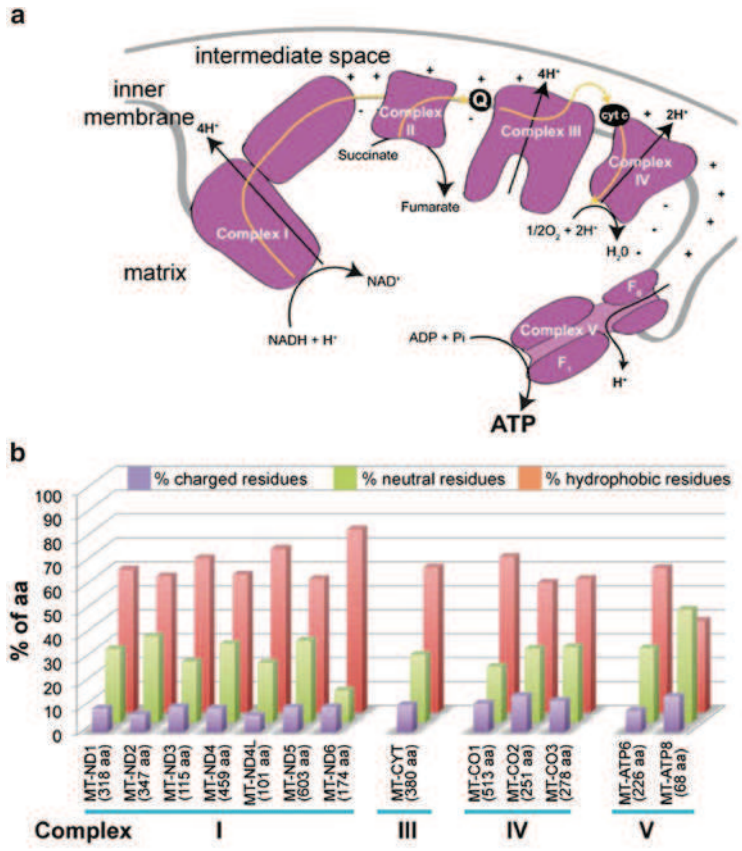


Fig. 2 The respiratory chain complexes. (**a**) Organization of the complexes along the mitochondrial inner membrane. Complex I: NADH: ubiquinone reductase; Complex II: Succinate-coenzyme Q reductase; Complex III: Coenzyme Q: cytochrome *c* oxidoreductase; Complex IV: Cytochrome *c* oxidase; Complex V: ATP synthase. Q stands for Coenzyme Q, and cyt *c* for cytochrome *c*. Sub-units composition of the five complexes in given is Table 1. (**b**) Composition of the 13 mt-DNA-encoded sub-units

$NADH + H^+$ and FAD into $FADH_2$. Oxidation of these hydrogen carriers involves the transfer of protons and electrons to the respiratory chain Complexes I and II respectively, followed by the channeling of these electrons through Complexes III and IV, where the electron is finally accepted by oxygen to form metabolic water. During electron transport, proton pumps in Complexes I, III, and IV become activated, leading to the expulsion of protons from the mitochondrial matrix to the mitochondrial intra-membrane space. The generated proton gradient (chemical potential) combined with the electron's movement (electrical potential) lead to return of protons to the matrix, thus activating ATP synthase by a proton-motive force. This enzyme binds ADP to inorganic phosphate and generates ATP. In summary, final oxidation of nutrients into CO_2 and H_2O takes place inside mitochondria and is directly correlated to oxygen consumption and ATP synthesis by the respiratory chain complexes [1].

Table 1 Name and composition of sub-units or the respiratory chain complexes

Complex	Name	Mitochondrial-encoded sub-units	Nuclear-encoded sub-units
I	*NADH:ubiquinone reductase*	7 (MT-ND1, MT-ND2, MT-ND3, MT-ND4, MT-ND4L, MT-ND5, MT-ND6)	38 (NDUFA1, NDUFA2, NDUFA3, NDUFA4, NDUFA5, NDUFA6, NDUFA7, NDUFA8, NDUFA9, NDUFA10, NDUFA11, NDUFA12, NDUFA13, NDUFB1, NDUFAB1, NDUFB2, NDUFB3, NDUFB4, NDUFB5, NDUFB6, NDUFB7, NDUFB8, NDUFB9, NDUFB10, NDUFB11, NDUFC1, NDUFC2, NDUFS1, NDUFS2, NDUFS3, NDUFS4, NDUFS5, NDUFS6, NDUFS7, NDUFS8, NDUFV1, NDUFV2, NDUFV3)
II	*Succinate: coenzyme Q reductase*	None	4 (SDHA, SDHB, SDHC, SDHD)
III	*Coenzyme Q: cytochrome c oxidoreductase*	1 (MT-CYB)	9 (CYC1, UQCR10, UQCR11, UQCRB, UQCRC1, UQCRC2, UQCRFS1, UQCRH, UQCRQ)
IV	*Cytochrome c oxidase*	3 (MT-CO1, MT-CO2, MT-CO3)	16 (COX4I1, COX4I2, COX5A, COX5B, COX6A1, COX6A2, COX6B1, COX6B2,COX6C, COX7A1, COX7A2, COX7B, COS7B2, COX7C, COX8A, COX8C)
V	*ATP Synthase*	2 (MT-ATP6, MT-ATP8)	17 (ATP5A1, ATP5B, ATP5C1, ATP5D, ATP5E, ATP5F1, ATP5G1, ATP5G2, ATP5G3, ATP5H, ATP5I, ATP5J, ATP5J2, ATP5L, ATP5L2, ATP5O, ATP5IF1)
	Mitochondrial respiratory chain complex assembly factors	None	28 (ATPAF1, ATPAF2, BCS1L, COA1, COA3, COA4, COA5, COA6, COX10, COX11, COX14, COX15, COX17, COX18, COX19, NDUFAF1, NDUFAF2, NDUFAF3, NDUFAF4, NDUFAF5, NDUFAF6, NDUFAF7, NUBPL, SCO1, SOC2, SDHAF1, SDHAF2, SURF1)

Respiratory chain complexes are large multi-protein complexes. Interestingly, the sets of proteins involved are of dual genetic origin. In humans, a total of 84 subunits and an additional 28 assembly factors are nuclear-encoded while 13 subunits are encoded by the mitochondrial genome (Table 1). These correspond to seven subunits of Complex I (NADH:ubiquinone reductase), one subunit of Complex III (Coenzyme Q: cytochrome *c* oxidoreductase), three subunits of Complex IV (cytochrome *c* oxidase), and two subunits of Complex V (ATP synthase). Complex II (succinate:coenzyme Q reductase) is the only complex formed exclusively by nuclear-encoded subunits. The size of the 13 mt-DNA-encoded proteins ranges from 68 amino acids (aa) (ATP8) to 603 aa (ND5) (Fig. 2b). The proteins are rather hydrophobic with 59.4% ± 8.5% aliphatic and aromatic residues, 29.8% ± 7.8% neutral residues, and 10.8% ± 2.6% charged residues. Leucine residues are present at the highest levels (14.4%); isoleucine, serine, and threonine are present at more than 7%, while some residues represent less than 3% (arginine, aspartate, cysteine, glutamine, glutamate, and lysine) of the protein compositions. It should be noted that a full and active set of mitochondrial translation components has been maintained for the synthesis of solely these 13 mt-DNA-encoded subunits.

1.2 The Human Mitochondrial Translation Machinery

The human mitochondrial genome is a circular double-stranded DNA of 16,569 bp [2]. This genome is tightly packed (with a single non-coding domain, the D-loop) and codes for the 13 respiratory chain subunits, in addition to 2 ribosomal RNAs (rRNAs) and 22 transfer RNAs (tRNAs). All three families of RNAs – mRNAs, rRNAs, and tRNAs – are processed from large primary transcripts according to the tRNA punctuation model [3]. The mitochondrial translation apparatus (Fig. 3) further involves a large number of proteins that are all nuclear-encoded, and are synthesized in the cytosol, before being imported into mitochondria for maturation. These include the full set of ribosomal proteins and ribosomal assembly proteins (translation initiation, elongation, termination factors) and tRNA maturation and modifying enzymes (enzymes cleaving the 5′- and 3′- ends of tRNAs primary transcripts, enzymes fixing the non-coded CCA 3′-end, enzymes of post-transcriptional modification). More than 100 proteins have been reported so far as being actors of the mitochondrial translation machinery [4–7] and will not be further discussed herein. Last but not least, a full set of nuclear-encoded aaRSs is required. These enzymes catalyze the specific esterification of tRNA 3′-ends with the corresponding amino acid so that the aminoacyl-tRNA (aa-tRNA) can be taken up by the translation factors and brought to the ribosome where the nascent protein is synthesized [8, 9]. The present knowledge of human mt-aaRSs will be discussed extensively below.

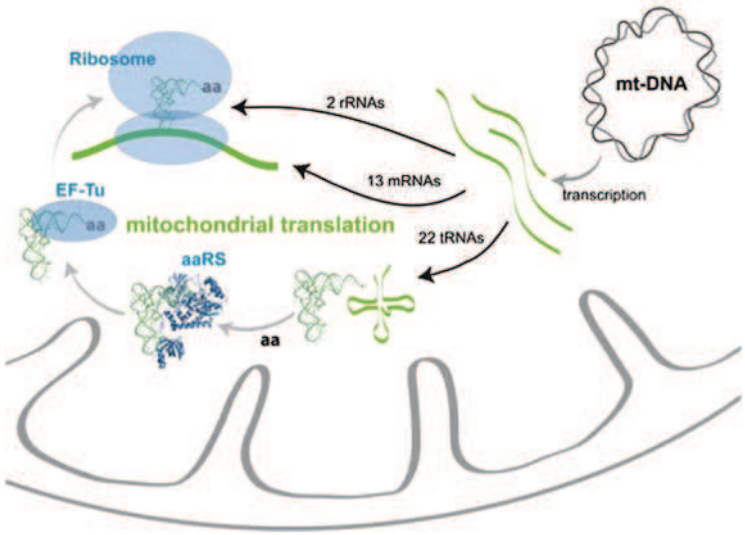


Fig. 3 The mitochondrial translation machinery. The human mt-DNA codes for 13 respiratory chain subunits (mRNAs), 2 ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs). All other requested proteins, such as, e.g., ribosomal proteins and ribosomal assembly proteins, translation initiation, elongation, and termination factors, tRNA maturation and modifying enzymes, mt-aaRSs, are encoded by the nucleus, synthesized in the cytosol, and imported into the mitochondria

1.3 Link Between Mitochondrial Translation, Mitochondrial Respiration, and Mitochondrial Disorders

As mentioned above, ATP synthesis by mitochondria is dependent on the coordinated expression of nuclear and mitochondrial genes. First, there is a need to coordinate the biogenesis of the respiratory chain complexes along the inner mitochondrial membrane so that partner proteins find each other to form the individual multiprotein complexes. Second, there is also a need to coordinate the setup and maintenance of mitochondrial translation machinery. Above all, this involves efficient partnerships between mt-DNA-encoded RNAs and nuclear-encoded proteins, especially between rRNAs and ribosomal proteins to form active ribosomes and between tRNAs and aaRSs to allow for accurate synthesis of aa-tRNAs. Accordingly, there are key links between the aminoacylation activity of aaRSs in charge of the synthesis of the 13 mt-DNA-encoded proteins, and the activity of respiratory chain complexes. It can be anticipated that any dysfunction of a single macromolecule of the translation machinery may have severe impacts on the activity of the respiratory chain complexes.

Mitochondrial disorders were defined as pathologies with aberrant oxidative phosphorylation (OXPHOS). Potential causes include an aberrant ROS production, elevation of NADH/NAD^+ ratio and lactate production, and/or ATP production deficiency. Defects were observed in a large variety of organs, and could manifest at any stage of life [10]. They were described in the late 1980s as exclusively

related to mutations within the mt-DNA and thus maternally inherited. Additional disorders were subsequently associated with mutations within nuclear genes coding for proteins of mitochondrial location, and thus followed Mendelian inheritance. Mitochondrial disorders are nowadays classified according to the genetic origins of the involved-mutations. (1) The first category, actually the firstly reported, concerns mt-DNA-encoded RNAs. All the 22 mt-DNA-encoded tRNAs have been linked to pathology-related mutations. The most striking examples concern the $tRNA^{Lys}$ and $tRNA^{Leu}$, currently described as "hot spots" for mutations, and correlated respectively with Myoclonus Epilepsy with Ragged Red Fibers (MERRF [11]) and Mitochondrial Encephalomyopathy with Lactic Acidosis and Stroke-like episodes (MELAS [12]). More than 230 mutations in tRNA genes are presently referenced in the 2012 MITOMAP Human Mitochondrial Genome Database (http://www.mitomap.org). Mt-DNA also codes for 2 rRNAs, with around 50 disease-related mutations described that are most frequently connected with aminoglycoside-induced deafness or non-syndromic sensorineural deafness (DEAF). (2) The second category concerns mt-DNA-encoded proteins. They are all sub-units of the respiratory chain complexes, making those complexes sensitive to mt-DNA mutations. As an example, mutations in ND5 (subunit 5 of complex I) can lead to MELAS, Leigh syndrome, MERRF, or Leber's Hereditary Optic Neuropathy (LHON) defects (reviewed in, e.g., [13]). (3) The third category, the most diverse, concerns nuclear-encoded proteins of mitochondrial location. On the one hand, mutations can affect proteins of the RC, which are directly contributing to OXPHOS (e.g., mutation in cytochrome oxidase subunits which are linked with Leigh syndrome, reviewed in [14]). On the other hand, mutations can affect proteins involved in the mt-DNA maintenance (e.g., DNA polymerase gamma, reviewed in [15]) and/or translation (e.g., mitochondrial elongation factor, reviewed in, e.g., [16]) interfering indirectly with OXPHOS. AaRSs of mitochondrial location belong to the last category. The recent discovery of mutations within mt-aaRS genes and the growing number of reported cases is opening a path to an emerging field of investigation (reviewed in, e.g., [6, 17]), and is the reason for a strong interest in the understanding of fundamental function of the mt-aaRSs and their implication in mitochondrial disorders.

Before reviewing in detail the molecular aspects of point mutations on mt-aaRSs properties, the various approaches available for the evaluation of respiratory chain complex activities used as tools either for diagnosis or for molecular investigation of this links between translation and respiration will be described.

1.4 Biochemical Analysis of Mitochondrial Respiration: A Potential Diagnostic Tool for the Detection of Mitochondrial Translation Defects

Because mitochondria provide much of the cellular energy, mitochondrial disorders preferentially affect tissues with high energy demands, including brain, muscle, heart, and endocrine systems. Consequently, mitochondrial defects play a central role in hereditary mitochondrial diseases, ischemia reperfusion injury, heart failure,

Table 2 Some of the various experimental procedures used to evaluate the impact of a mutation on the different steps from mt-aaRS expression to mitochondrial ATP synthesis

Steps of mt-aaRSs life cycle		Methods
1	aaRS mRNAs expression/ processing	RT-PCR, qPCR, northern blot
2	aaRS synthesis/stability	Western blot, inhibition of cytosolic translation
3	aaRS import	GFP-fusion protein, immuno-cytochemistry, in vitro import assay, in vitro maturation assay
4	aaRS folding/oligomerization/ stability	Protein refolding, coexpression of diffently tagged proteins
5	amino acid activation	[P^{32}]/colorimetric-based ATP-PPi exchange assay
	tRNA recognition	Aminoacylation assay
	tRNA charging	Northern blot, aminoacylation assay
6	Synthesis of mt-DNA encoded RC sub-units	BN-PAGE, pulse-chase experiment
7/8	RC complex activity	Histochemical and immunohistochemical methods, polarography

Numbers on the left recall the steps displayed in Fig. 1

metabolic syndrome, neurodegenerative diseases, and cancer [18, 19]. There are remarkably diverse causes for mitochondrial disorders. These may be linked to the dual genetic systems encoding components of the respiratory chain complexes, to the need for mitochondrial translation machinery, but also to mechanisms required for the biosynthesis and maintenance of mt-DNA and to the biogenesis of the organelle itself. Moreover, each cell may contain hundreds to thousands of copies of the mitochondrial genome. The distribution of the affected tissues and the proportion of mutant to wild-type mt-DNA (termed heteroplasmy) lead to clinical manifestations, which are remarkably variable and heterogeneous. An additional breakthrough came from the later discovery of mutations within nuclear genes as the causes for similar diseases. Therefore, the establishment of a mitochondrial disorder diagnosis can be very difficult. It requires an evaluation of the family pedigree, in conjunction with a thorough assessment of the clinical, imaging, and muscle biopsy analysis [20]. Isolated OXPHOS deficiencies are generally caused by mutations in genes encoding subunits of the OXPHOS system. Combined deficiencies in the respiratory chain complexes may reflect the consequence of mutations in mt-DNA-encoded tRNAs or rRNAs, or are due to rearrangements or depletion of mt-DNA [21]. They may also reflect a dysfunction of the mitochondrial translation machinery.

Table 2 summarizes some of the various experimental procedures used to evaluate the impact of a mutation on the different steps from mt-aaRS expression to mitochondrial ATP synthesis. It includes the initial screening procedure of muscle biopsy analysis, i.e., histochemistry and immunohistochemistry. The activity of the five multiprotein enzymatic complexes can be assayed globally by measuring the mitochondrial inner membrane electrochemical potential, oxygen consumption, or ATP synthesis, or assayed individually by measurement of their enzymatic activities. This is performed by polarography. The structural integrity of the multiprotein complexes, evaluated by blue native gel electrophoresis is also increasingly used

for diagnostics [22]. Finally, once the diagnosis is performed, it is possible to evaluate the impact of mutations on the rate of mitochondrial protein synthesis using chase experiments on specific recombinant cell-lines [23].

Several *histochemical and immunohistochemical methods* can be used as reliable morphological tools in order to visualize respiratory chain abnormalities on tissue sections. Classical histochemistry techniques allow visualizing succinate dehydrogenase (SDH) and cytochrome *c* oxidase (COX) activity. Indeed, the most informative histochemical impairment of mitochondria in skeletal muscle is ragged red fibers (RRF), observed on frozen sections traditionally with the modified Gomori trichrome method [24]. Since the accumulation of material other than mitochondria sometimes simulates a RRF appearance, the identification of deposits suspected of being mitochondrial proliferation should be confirmed by histochemical staining of oxidative enzymes as SDH and COX. Muscle from mitochondrial myopathy associated with mt-DNA mutations tends to show a mosaic expression of COX consisting of a variable number of COX-deficient and COX-positive fibers, and RRFs can be COX negative or COX positive. The mosaic pattern of COX expression could be considered as the histochemical signature of a heteroplasmic mt-DNA mutation affecting the expression of mt-DNA-encoded genes in skeletal muscle [25]. Immunohistochemical methods allow the visualization of the expression of several mt-DNA and nuclear-encoded subunits of the respiratory chain. An antibody against COX IV is routinely used as a probe for nuclear-encoded mitochondrial protein and an antibody against COX II as a probe for mt-DNA-encoded protein. However, any other combination of antibodies can also be used.

Polarography (spectrophotometric assays) can be applied to both tissue samples and cultured cells and is designed to assess the enzymatic activity of the individual OXPHOS Complexes I–V, along with the Krebs cycle enzyme citrate synthase as a mitochondrial control. Determining the enzymatic activities can be valuable in defining isolated or multicomplex disorders and may be relevant to the design of future molecular investigations [26]. Different assays have to be performed in order to analyze each complex separately. Assays for Complexes I, II, II + III, III, and IV are routinely performed when there is a suspicion of mitochondrial defects. The principle of polarographic approach is based on the fact that mitochondria require oxygen to produce ATP. The rate of oxygen consumption from isolated mitochondria or directly in skinned fibers is a useful and valuable technique in the research and evaluation of mitochondrial dysfunction and disease, because ADP-dependent oxygen consumption directly reflects OXPHOS efficiency [27–31]. In the presence of oxidizable substrates, freshly isolated mitochondria are introduced into the polarograph and oxygen consumption is measured in the presence of exogenously added ADP as well as several inhibitors. The typical parameters determined from mitochondrial polarography include state III rate (maximal mitochondrial respiration with ADP), state IV rate (basal mitochondrial respiration), and RCR (respiratory control ratio or state III rate/state IV rate). RCR is a good indicator of the integrity of the inner membrane of the isolated mitochondria and is sensitive for indicating OXPHOS defects. The ratio of ADP consumed/oxygen consumed during the experiment is a direct reflection of phosphorylation efficiency

and can indicate abnormalities of the ATP synthase activity or uncoupling between the activities of Complexes I to IV and complex V (ATP synthase). It is possible to explore the respiratory parameters of skeletal muscle with permeabilized muscle fibers, thus skipping a mitochondria purification step [32]. Muscle fibers are permeabilized by saponin, allowing respiratory substrates and inhibitors to reach the mitochondria [31, 33].

Blue native polyacrylamide (or agarose) gel electrophoresis (BN-PAGE) allows for the isolation of intact respiratory chain complexes and analysis of their subunit content (reviewed in [34]). Briefly, after solubilization of mitochondria in the presence of dodecylmaltoside, large complexes are first separated by native gels electrophoresis on low percentage polyacrylamide or agarose gel. A second dimension electrophoresis is then performed under non-native conditions in the presence of SDS and β-mercaptoethanol using a 10% polyacrylamide gel. Individual subunits are detected by western blotting or mass spectrometry. The two-dimensional separation approach can also be adapted to perform in gel activity assays to address the dynamics of protein synthesis and complex assembly (in combination with pulse-chase labeling of proteins in cultured cells).

In conclusion, a variety of biochemical approaches are available to evaluate mitochondrial function. These form a powerful toolkit that can be used to diagnose the mitochondrial origin of a disorder. However, the gap is large between the dysfunction of the respiratory chain (as a whole, or as individual complexes) and the understanding of the dysfunction at the molecular level. Indeed, considering the translation rate of each of the 13 mt-DNA-encoded proteins as the sole outcome resulting from defects in tRNAs and/or aaRSs is too simplistic. The links between the mitochondrial translation of a given aaRS on the one hand, and ATP production on the other, involve a number of issues that need to be explored. Some of these will be discussed in what follows. One must also take into account the possibility of alternative functions of mt-aaRSs, outside their strict housekeeping role in translation.

2 Mt-aaRSs in Mitochondrial Translation

The housekeeping function of aaRSs is to provide aminoacylated-tRNAs (aa-tRNAs) for translation. Enzymes connect tRNAs with their corresponding amino acid in an efficient and specific way through a two-step reaction named aminoacylation. In the first step, the amino acid is activated by ATP into an aminoacyl-adenylate (followed by the release of PPi). In the second step, the activated amino acid is transferred onto the cognate tRNA, releasing AMP. The formation of the 20 canonical aa-tRNA species in human mitochondria requires the import of a complete set of mt-targeted aaRSs encoded by the nuclear genome. The faster evolution rate of mt-DNA than the nuclear genome [35, 36] leads to abnormal RNAs, shrunken in size and often lacking important signals. For instance, most mt-tRNAs have shortened sizes, miss crucial folding and recognition nucleotides as compared to “classical” tRNAs, and are more flexible [6, 7, 37, 38]. Mechanisms compensating for the degeneration of mt-tRNAs

remain mainly unsolved and raise the question of molecular adaptation of partner proteins, especially mt-aaRSs. The structural and functional deciphering of the set of aaRSs of human mitochondrial location is at the early stages.

2.1 *Nuclear-Encoded aaRSs of Mitochondrial Location and Evolutionary Considerations*

The nuclear gene annotation of the human aaRSs of mitochondrial location was completed a decade ago [8]. Access to gene sequences highlighted that the set of aaRSs dedicated to translation in human mitochondria is mainly different from the set acting in the cytosolic translation (Fig. 4). This concerns 17 out of the 19 pairs of aaRS. The two exceptions concern the GlyRSs and LysRSs. GlyRSs are generated from two translation initiation sites on the same gene, leading to one enzyme with a mitochondrial targeting sequence (MTS), mt-GlyRS, and a second without, cytosolic GlyRS [39, 40]. With the LysRSs, an alternative mRNA splicing pathway allows for the insertion – or not – of the nucleotide sequence coding for the MTS, leading to two mature mt- and cytosolic LysRSs differing only by a few residues at their N-terminus [41]. No gene coding for mt-GlnRS, the 20th synthetase, has been found so far, leaving open the question about how glutaminylation in human mitochondria is performed. Among possible explanations, it is proposed that either the sequence of human mt-GlnRS has evolved so much that it has become unrelated to any of the known GlnRSs, or its function is fulfilled by an mt-addressed version of cyt-GlnRS (as proposed in yeast [42]), or that the synthesis of mt Gln-tRNA^{Gln} occurs via an indirect pathway (the transamidation pathway) involving misacylation of tRNA^{Gln} by a non-discriminative GluRS followed by Glu-amidation [43]. The existence of an indirect pathway in mitochondria was demonstrated in the cases of plants [44, 45], yeast [46], and more recently humans [47]. However, the coexistence of direct and indirect pathways for Gln-tRNA^{Gln} synthesis in yeast and mammalian mitochondria is still under consideration [48, 49].

Despite the conventional view of the endosymbiotic origin of mitochondria [50], the source of nuclear genes for mt-addressed aaRSs is diverse and not necessarily easy to trace back. Some of the mt-addressed aaRSs originate from the bacterial domain, but none specifically from the alpha-proteobacteria, although the alpha-proteobacterial contribution to the mitochondrial genome is well established. This favors the hypothesis that mt-aaRSs have been acquired by numerous post-endosymbiotic and/or lateral gene transfer events, from sources representative of all kingdoms of life [51]. The precise knowledge on the origin of all human mt-aaRS genes is necessary for the investigation of pathology-related mutations (see below) and will be of help in building up homology models in cases where crystallographic structures for mt-aaRS are not available. This global view on the origin of all human mt-aaRS genes will be established soon (Sissler et al., in preparation).

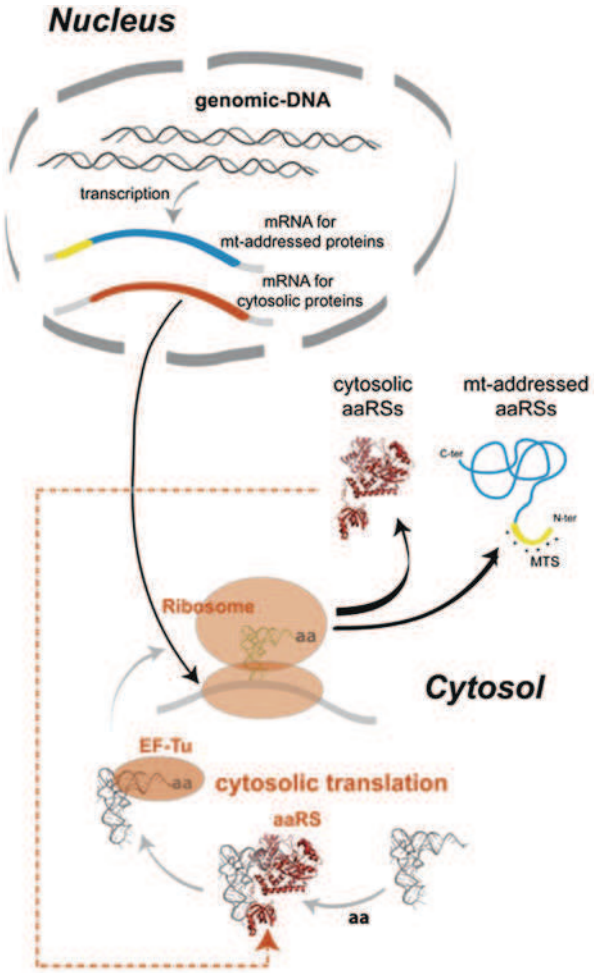


Fig. 4 Two sets of genes for human cytosolic and mitochondrial aaRSs. Achievement of human genome sequencing and gene annotation of the human aaRSs of mitochondrial location [8] reveal the presence of two distinct sets of nuclear genes for aaRSs (with two exceptions, see text). One set (in *red*) codes for the aaRSs of cytosolic location. The second set (in *blue*) codes for the aaRSs of mitochondrial location. The latter distinguishes by the presence of an encoded N-terminal mitochondrial targeting sequence (MTS, in *yellow*). The two sets of genes are translated via the cytosolic translation machinery. Sequences that are subsequently addressed to mitochondria are maintained unfolded in the cytosol, pending their entry into the organelle. Mutations are separated by half line spaces so that to illustrate those found in the two alleles of a single patient

2.2 *Mt-aaRSs are Imported Proteins*

Human mt-aaRSs are translated within the cytosol and subsequently imported into the mitochondria thanks to the presence of MTSs. As is the case for the vast majority of proteins of mitochondrial matrix and/or inner membrane location, these sequences

are predicted to be located at the N-termini of precursor mt-aaRSs [8]. However, to date, no consensus sequences have been deciphered. An MTS typically consists of ~15–50 amino acids including numerous positively charged residues (e.g., lysines and arginines), and forms amphipathic alpha-helices [52]. The MTS first directs the precursor proteins to mitochondria where they are further translocated by the translocase complexes of the outer (TOM) and inner (TIM) membranes (reviewed in, e.g., [53–59]). Upon arrival in the matrix, MTS are proteolytically cleaved by the mitochondrial processing peptidase (MPP). This process was recently shown possibly to affect the half-life of the proteins [60]. Removal of the pre-sequence exposes new amino-termini of the imported proteins, which may contain a stabilizing or a destabilizing amino acid (bulky hydrophobic residues are typically destabilizing). The N-end rule indeed states that regulation of the proteolytic degradation is closely related to the N-terminal residue of proteins [61]. It has recently been proposed that two additional peptidases (that function subsequently to MPP) are implicated in protein stabilization by removing the newly exposed N-terminal destabilizing residue(s). The first is the intermediate cleaving peptidase Icp55, which removes a single amino acid. The second is the mitochondrial intermediate peptidase Oct1, which removes an octapeptide. Accordingly, the processing of imported proteins is closely connected to protein turnover and quality control (reviewed in [54]).

Cleavage (maturation) sites of the MTS for human mt-aaRSs are so far mostly defined according to theoretical predictions based on computer programs (e.g., Predotar http://urgi.versailles.inra.fr/predotar/predotar.html, MitoProt http://ihg.gsf.de/ihg/mitoprot.html, TargetP http://www.cbs.dtu.dk/services/TargetP/, and iPSORT http://ipsort.hgc.jp/). However, the expression of recombinant human mt-aaRS proteins, deprived of theoretically predicted MTSs, appears to be difficult due to low solubility and the tendency of proteins to aggregate [9, 62]. This is likely to indicate that many predictions may be incorrect. Indeed, inaccurate prediction of the cleavage site was previously shown to be responsible for suboptimal expression of human mt-LeuRS. Only the LeuRS variant deprived of its 39 N-terminal amino acids was sufficiently overexpressed in *Escherichia coli*, efficiently purified, and fully active, while the variant deprived of the predicted 21 amino acids remained insoluble [63, 64]. Along the same line, the re-design of the N-terminus of human mt-AspRS enhances expression, solubility, and crystallizability of the mitochondrial protein [65, 66]. Discrepancies are also apparent between the predicted cleavages sites, the starting amino acid of the recombinant proteins, and the first residue visible in the established crystallographic structures to date (reviewed in [65]).

None of the above-mentioned examples have the experimental exact cleavage points of the mature proteins been established so far. These examples however emphasized that the preparation of stable recombinant molecules would gain from optimized criteria to predict MTS cleavage sites. A more systematic effort to determine experimentally the precise N-terminus of mature mitochondrial proteins (as done for, e.g., the yeast mt proteome [60]) would be of help to determine unambiguously the sequence of a functional mt protein.

However, this analysis remains difficult to perform experimentally (by sequencing or mass spectrometry) due to the minute amount of protein that can be isolated from human mitochondria and because of the risk of secondary proteolysis.

2.3 Structural Insights of mt-aaRSs

When considering their primary structures, all mt-aaRSs fall into the expected classes of the aaRSs as originally defined in [67, 68], with signature motifs being respectively HIGH and KMSK for class I enzymes, and motifs 1, 2, and 3 for class II enzymes. Striking divergences are, however, observed when considering their modular organizations. Modular design of the aaRSs is a result of a patchwork assembly of different functional modules during evolution. Minimal cores are the catalytic domain and the tRNA anticodon binding domain that are possibly surrounded by additional components for structural or functional purposes. Those are either remnants from early ancestors or structural inventions for functional expansion (e.g., [69]). The most striking observation concerns the divergence between the tetrameric cytosolic PheRS ($\alpha_2\beta_2$) and the monomeric (α) version found in human mitochondria [70], a situation also observed for PheRSs in yeast [71]. Human mt-PheRS possesses the minimum set of structural domains, making this enzyme the smallest exhibiting aminoacylation activity and the only class II monomeric synthetase [72]. Another striking observation is that two independent coding sequences have been found for mt-GluRS and mt-ProRS, as opposed to a single gene for both activities in the human cytosol leading to the bifunctional GluProRS [73].

Although the first crystallographic structure of a bacterial aaRS was published three decades ago (TyrRS from *Bacillus stearothermophilus* [74]), the first 3D structures of mt enzymes were only established recently: a bovine structure in 2005 (mt-SerRS [75]), the first human structure in 2007 (mt-TyrRS [76]), followed by two additional ones in 2012 (mt-PheRS [72] and mt-AspRS [66]) (Fig. 5). The time lag between resolving structures of bacterial aaRS and mammalian mt-aaRSs was mainly due to the difficulties involved in producing large amounts of stable mitochondrial proteins (defining the N-terminus leading to a soluble protein). Resolution of crystallographic structures and investigation of biophysical properties of mt-aaRSs reveal similarities, but also distinctive features, when compared to related prokaryotic homologs (the four mt-aaRSs for which crystallographic structures have been obtained are of prokaryotic origin [51]).

Structural idiosyncrasies (Fig. 5) concern, for instance, the very unique structural organization of mt-PheRS (monomer instead of heterotetramer [77]) or the presence of two distinctive insertions in mt-SerRS (an 8 aa amino-terminal "distal helix" and a carboxy-terminal "C-tail" composed of an over 40 Å long flexible loop stretching away from the body of the monomer [75]). More generally, it has been observed that, besides having similar architectures to prokaryotic homologs, mitochondrial enzymes are distinguished by more electropositive surface potentials (specifically along the

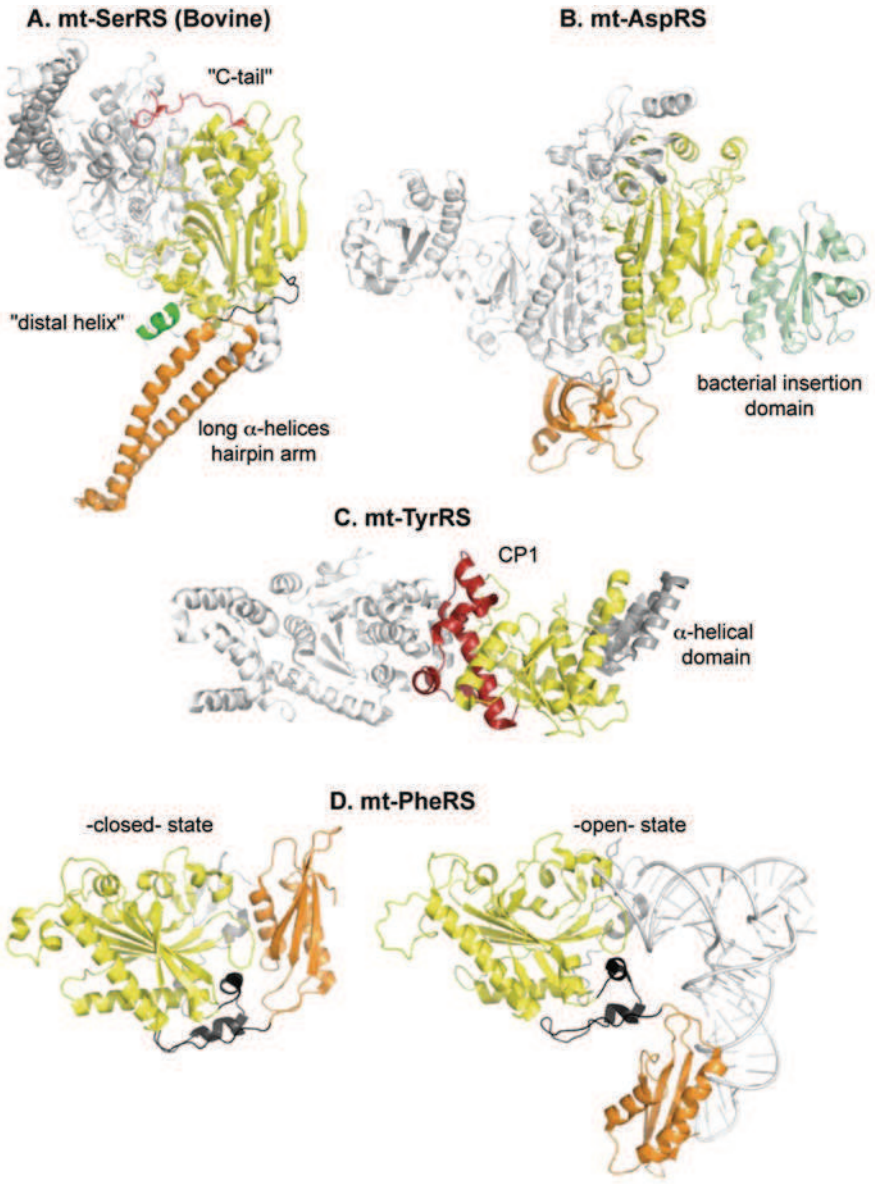


Fig. 5 Known crystallographic structures of mammalian mt-aaRSs. (**A**) Bovine mt-SerRS [75], where the specific "distal helix" and "C-tail" are emphasized in *green* and *red*, respectively. In addition, the bacterial-type N-terminal long α-helices hairpin arm is shown in *orange*. (**B**) Human mt-AspRS [66], where the bacterial insertion domain is highlighted in *light green*. (**C**) Human mt-TyrRS [76], where the CP1 and the α-helical domains are indicated in *red* and *gray*, respectively. Note that the S4-like domain is missing in this structure. (**D**) Human mt-PheRS in the –closed– state [77], and in the –open– state, complexed within *Thermus thermophilus* $tRNA^{Phe}$ (in *white*) [72]. Binding of tRNA engenders a drastic conformational change of mt-PheRS through ~160° hinge-type

tRNA binding surface) or by enlarged grooves for tRNA accommodation. The latest aspect is visible for instance in mt-AspRS, where the angle formed by the bacterial-specific insertion domain and the catalytic domain is more opened by 26°C [66]. This can also be seen in mt-PheRS where the PheRS/$tRNA^{Phe}$ complex formation was shown to be accompanied by a considerable rearrangement through an ~160° hinge-type rotation from a –closed– to an –open– state of the PheRS and the global repositioning of the anticodon binding domain upon tRNA binding [72, 77]. In addition, an alternative interaction network has been observed at the subunit interface of the dimeric mt-AspRS (weaker in terms of salt-bridges and hydrogen bonds). Biophysical investigations also demonstrated a thermal stability reduced by as much as 12°C for mt-AspRS, compared to *E. coli* AspRS [66]. Finally, a gain of plasticity is proposed for both the mt-TyrRS, where the KMSKS loop is rather remote from the active site, explaining the relative lacks of constraints in the structure [76], and the mt-AspRS, where unusual thermodynamic properties of tRNA binding are observed [66]. It has been suggested that the gain of plasticity may be a more general property of mt-aaRSs, as they have to deal with degenerated mt-tRNAs [66].

2.4 Some Functional Peculiarities

The bacterial origin (established for many of the mt-aaRSs) predicts that most of the mt-aaRSs should behave as prokaryotic aaRSs. However, this is not the case, raising interesting questions regarding the evolution of macromolecules of the mitochondrial translation machinery. The human mt-AspRS has been extensively studied along these lines. It shares 43% of identical residues (including residues specific for all AspRSs), the same modular organization (including the bacterial-type insertion and C-terminal extension domains), and the same architecture as *E. coli* AspRS, a representative bacterial homolog [66]. However, and despite the fact that the two enzymes are likely descendants from a common ancestor, numerous functional idiosyncrasies/discrepancies were reported. Indeed, the mt-AspRS exhibits a reduced catalytic efficiency [8, 9], requires a minimal set of recognition determinants within its cognate tRNA [78], displays a higher sensitivity to small substrate analogs [79], is able to cross aminoacylate bacterial tRNAs (while the bacterial enzyme unilaterally recognizes bacterial tRNAs [80, 81]), and shows an increased intrinsic plasticity when compared to its bacterial homolog [66].

It is proposed that all structural and functional peculiarities of the mt-aaRSs (exemplified here by the mt-AspRS) with respect to the bacterial homologs may represent an evolutionary process, allowing nuclear-encoded proteins to cooperate with degenerated organelle RNAs [66].

Fig. 5 (continued) rotation and the global repositioning of the anticodon binding domain. For all structures the catalytic core is in *yellow*, the anticodon binding domain in *orange*, and the hinge region in *black*. When appropriate, the second dimer is displayed in *light gray*

3 Mt-aaRSs in Human Disorders

As already evoked, a breakthrough took place in 2007 with the discovery of a first set of mutations present in the nuclear gene of an mt-aaRS, namely mt-AspRS [82]. This first discovery was followed very rapidly by the description of numerous additional mutations not only on the same gene but also on other mt-aaRS-coding genes, so that half of them are presently known to be affected [6, 17]. This discovery sheds light on a new family of nuclear genes involved in human disorders allowing the new naming of "mt-aaRS disorders."

3.1 Discovery of aaRS-Related Disorders

Mutations in *DARS2*, the nuclear gene coding for mt-AspRS, were first found in 2007 in patients with cerebral white matter abnormalities of unknown origin [82]. These abnormalities were part of a childhood-onset disorder called Leukoencephalopathy with Brain stem and Spinal cord involvement and Lactate elevation (LBSL). Since this first discovery, mutations in eight additional mt-aaRS-encoding genes have been reported (Table 3). They hit *RARS2* (patients with PontoCerebellar Hypoplasia type 6, PCH6), *YARS2* (Myopathy, Lactic Acidosis and Sideroblastic Anemia, MLASA syndrome), *SARS2* (HyperUricemia, Pulmonary hypertensions and Renal failure in infancy and Alkalosis, HUPRA syndrome), *HARS2* (Perrault Syndrome, PS), *AARS2* (Infantile Mitochondrial Cardiomyopathy), *MARS2* (Autosomal Recessive Spastic Ataxia with Leukoencephalopathy, ARSAL), *EARS2* (Early-onset Leukoencephalopathy with Thalamus and Brainstem involvement and High Lactate), and *FARS2* (Infantile mitochondrial Alpers Encephalopathy). An observation among the numerous reported cases is that despite a dominant effect on brain and neuronal system, sporadic manifestations also occur in skeletal muscle, kidney, lung, and/or heart.

3.2 The Present-Day List of Pathology-Related Mutations Within mt-aaRS Encoding Genes

Nine mt-aaRS genes are currently known to harbor a total of 65 mutations, found in patients in 64 different genetic combinations (Table 4). All mutations can also be visualized on schematic representations of modular organizations of the proteins (Fig. 6). The most prominent affected gene is *DARS2*, located on chromosome I. It comprises 32,475 bp, codes for 17 exons and is translated into a 645 aa long mt-AspRS. To date, 28 different mutations are known: 8 nonsense mutations (frameshift, premature stop), 16 missense mutations (amino acid exchange), and 4 insertions/deletions have been described in 13 different reports. These mutations

Table 3 Human mt-aaRSs involved in mitochondrial disorders

Gene	Pathogenic manifestation	Tissue	First report
DARS2 Aspartyl-tRNA Synthetase	Leukoencephalopathy with Brain stem and Spinal cord involvement and Lactate elevation (LBSL)	Brain, spinal cord	[82]
RARS2 Arginyl-tRNA Synthetase	PontoCerebellar Hypoplasia type 6 (PCH6)	Brain	[101]
YARS2 Tyrosyl-tRNA Synthetase	Myopathy, Lactic Acidosis and Sideroblastic Anemia (MLASA syndrome)	Blood, skeletal muscle	[97]
SARS2 Seryl-tRNA Synthetase	HyperUricemia, Pulmonary hypertension and Renal failure in infancy and Alkalosis (HUPRA syndrome)	Kidney, lung	[108]
HARS2 Histidyl-tRNA Synthetase	Perrault Syndrome (PS)	Ovarian, sensorineural system	[90]
AARS2 Alanyl-tRNA Synthetase	Infantil Mitochondrial Cardiomyopathy	Heart	[110]
MARS2 Methionyl-tRNA synthetase	Autosomal Recessive Spastic Ataxia with Leukoencephalopathy (ARSAL)	Brain	[98]
EARS2 Glutamyl-tRNA Synthetase	Early-onset Leukoencephalopathy with Thalamus and Brain stem involvment and high Lactate	Brain	[89]
FARS2 Phenylalanyl-tRNA Synthetase	Infantile mitochondrial Alpers Encephalopathy	Brain	[104]

Initial chronological reports connecting affected mt-aaRS genes with pathogenic manifestations, and affected tissues

Table 4 Pathology-related mutations on mitochondrial aaRS genes, transcripts, and proteins

AARS2 (VI)	Gene mutations (13,672 bp; gene ID: 57505)	Exons (22; NM_020745.3)	Proteins (985 aa; NP_065796.1)	Domains	Ref.
	1774C<T	13	R592W	CD	[110]
	464T>G	3	L155R	CD	
	1774C<T	13	R592W	CD	
	1774C<T	13	R592W	CD	
DARS2 (I)	Gene mutations (32,475 bp; gene ID: 55157)	Exons (17; NM_018122.4)	Proteins (645 aa; NP_060592.2)	Domains	Ref.
	228-20_21delTTinsC	3	R76SfsX5	ABD	[82]
	1876C>G	17	L626V	BED	
	228-20_-21delTTinsC	3	R76SfsX5	ABD	
	787C>T	9	R263X	CD	
	228-11C>G	3	R76SfsX5	ABD	
	536G>A	6	R179H	CD	
	228-20_-21delTTinsC	3	R76SfsX5	ABD	
	788G>A	9	R263Q	CD	
	228-20_-21delTTinsC	3	R76SfsX5	ABD	
		15	A522_K558del	CD	
	228-20_-21delTTinsC	3	R76SfsX5	ABD	
	492+2T>C	5	M134_K165del	ABD/HR	
	228-20_-21delTTinsC	3	R76SfsX5	ABD	
	455G>T	5	C152F	HR	
	228-20_-21delTTinsC	3	R76SfsX5	ABD	
	617-663del	7	F207CfsX24	CD	
	133A>G	2	S45G	ABD	
	228-15C>A	3	R76SfsX5	CD	

536G>A	6	R179H	CD
1273G>T	13	E425X	BID
1837C>T	17	L613F	BED
1876T>A	17	L626Q	BED
228-15C>G	3	R76SfsX5	ABD
1886A>G	17	Y629C	BED
228-16C>A	3	R76SfsX5	ABD
295-2A>G	4	A100_P132del	ABD
228-20_-21delTTinsC	3	R76SfsX5	CD
1679A>T	16	D560V	ABD
228-20_-21delTTinsC	3	R76SfsX5	ABD
550C>A	6	Q184K	CD
228-15C>A	3	R76SfsX5	ABD
396+2T>G	4	A100_P132del	ABD
228-20_-21delTTinsC	3	R76SfsX5	ABD
742C>A	8	Q248K	CD
228-20_-21delTTinsC	3	R76SfsX5	ABD
1272_1273GG>C	13	E424NfsX1	BID
228-20_-21delTTinsC	3	R76SfsX5	ABD
397-2A>G	5	M134_K165del	ABD/HR
228-20_-21delTTinsC	3	R76SfsX5	ABD
536G>A	6	R179H	CD
228-10C>A	3	R76SfsX5	ABD
492+2T>C	5	M134_K165del	ABD/HR

(continued)

Table 4 (continued)

DARS2 (I)	Gene mutations (32,475 bp; gene ID: 55157)	Exons (17; NM_018122.4)	Proteins (645 aa; NP_060592.2)	Domains	Ref.
	1345-17del13	14	C449_K521del	CD	[131]
	228-20_-21delTTinsC	3	R76SfsX5	ABD	
	228-16C>A	3	R76SfsX5	ABD	[85]
	716T>C	8	L239P	CD	
	228-20_21delTTinsC	3	R76SfsX5	ABD	[83]
	492+2T>C	5	M134_K165del	ABD/HR	
	228-20_21delTTinsC	3	R76SfsX5	ABD	
	455G>T	5	C152F	HR	
	228-16C>A	3	R76SfsX5	ABD	[84]
	745C>A	8	L249I	CD	
	228-16C>A	3	R76SfsX4	ABD	[99]
	?	?	?	?	
	228-16C>A	3	R76SfsX5	ABD	[94]
	228-16C>A	3	R76SfsX5	ABD	
	228-12C>G	3	R76SfsX5	CD	[112]
	228-24CinsT	3	R76SfsX6	CD	
	228-12C>A	3	R76SfsX5	ABD	[88]
	1069C>T	11	Q357X	CD	
	1825C>T	17	R609W	BED	[95]
	1825C>T	17	R609W	BED	
	228-20_21delTTinsC	3	R76SfsX5	ABD	[132]
	1395_1396delAA	14	T465TfsX7	CD	
	228-22T>A	3	R76SfsX5	ABD	[100]
	228-22T>A	3	R76SfsX5	ABD	

	172C>G	2	R58G	ABD	[102]
	406A>T	5	T136S	ABD	
EARS2 (XVI)	Gene mutations (32,980 bp; gene ID: 124454)	Exons (9; NM_001083614.1)	Proteins (523 aa; NP_001077083.1)	Domains	Ref.
	502A>G	4	R168G	CD	[89]
	1279_1280insTCC	7	T426_R427insL	ABD	
	322C>T	3	R108W	CD	
	322C>T	3	R108W	CD	
	1194C>G	6	Y398X	ABD	
	322C>T	3	R108W	CD	
	328G>A	3	G110S	CD	
	328G>A	3	G110S	CD	
	610G>A	4	G204S	CD	
	286G>A	2	E96K	CD	
	500G>A	4	C167Y	CD	
	322C>T	3	R108W	CD	
	949G>T	4	G317C	CD	
	164G>A	2	R55H	CD	
	670G>A	4	G224S	CD	
	1A>G	1	M1?	MTS	
	320G>A	3	R107H	CD	
	322C>T	3	R108W	CD	
	1547G>A	9	R516Q	CD	
	19A>T	1	R7X	MTS	
	322C>T	3	R108W	CD	

(continued)

Table 4 (continued)

EARS2 (XVI)	Gene mutations (32,980 bp; gene ID: 124454)	Exons (9; NM_001083614.1)	Proteins (523 aa; NP_001077083.1)	Domains	Ref.
	286G>A	2	E96K	CD	
	500G>A	4	C167Y	CD	
	193A>G	2	K65E	CD	[133]
	193A>G	2	K65E	CD	
FARS2 (VI)	Gene mutations (403,026 bp; gene ID: 10667)	Exons (6; NM_006567.3)	Proteins (451 aa; NP_006558.1)	Domains	Ref.
	986T>C	5	I329T	ABD	[104]
	1172A>T	6	D391V	CD	
	1275G>C	7	L425L	CD	
	1277C>T	7	S426F	CD	
	431A>G	3	Y144C	ABD	[134]
	431A>G	3	Y144C	ABD	
HARS2 (V)	Gene mutations (6,900 bp; gene ID: 234338)	Exons (13, NM_012208.2)	Proteins (506 aa; NP_036340.1)	Domains	Ref.
	1102G>T	10	V368L	ABD	[90]
	598C>G	6	L200V/L200_K211del	CD	
MARS2 (II)	Gene mutations (1,779 bp; gene ID: 92935)	Exons (1; NM_138395.3)	Proteins (593 aa, NP_612404)	Domains	Ref.
	681D268bpfs236X/wt				[98]
	Dup1/Dup1				
	Dup1/Dup2				
	Dup1/Dup-del				
	Yorube Insertion/wt				
RARS2 (VI)	Gene mutations (75,618 bp; gene ID: 57038)	Exons (20; NM_020320.3)	Proteins (562 aa, NP_612404)	Domains	Ref.
	1704A>G*	20	K568K	ABD	[101]
	872A>G*	10	K291R	CD	
	IVS2+5(A>G)*	2	L13RfsX15	MTS	
	35A>G	1	Q12R(Q12fsX25)	MTS	[87]
	1024A>G	12	M342V	CD	

	35A>G	1	Q12R(Q12fsX25)	MTS	[86]
	IVS2+5(A>G)	2	L13RfsX15	MTS	
	721T>A	9	W241R	CD	[91]
	35A>G	1	Q12R(Q12fsX25)	MTS	
	25A>G	1	I9V	MTS	
	1586+3A>T	18	R504_L528del	ABD	
	734G>A	9	R245Q	CD	
	1406G>A	16	R469H	ABD	
	IVS2+5(A>G)	2	L13RfsX15	MTS	[92]
	IVS2+5(A>G)	2	L13RfsX15	MTS	
	1211T>A	14	M404K	CD	
	471-473del	7	K158del	CD	
SARS2 (XIX)	Gene mutations (15,127 bp; gene ID: 54938)	Exons (16; NM_017827.3)	Proteins (518 aa; NP_060297.1)	Domains	Ref.
	1169A>G	13	D390G	CD	[108]
	1169A>G	13	D390G	CD	
YARS2 (XII)	Gene mutations (8,630 bp; gene ID: 51067)	Exons (5; NM_001040436.2)	Proteins (477 aa; NP_001035526.1)	Domains	Ref.
	156C>G	1	F52L	CD	[97]
	156C>G	1	F52L	CD	
	137G>A	1	G46D	CD	[109]
	137G>A	1	G46D	CD	

Gene names of affected mt-aaRSs with chromosomal location under brackets are given on the left. Reported nucleotide mutations such as, e.g., replacement, insertion, deletion or depletion (Gene mutations), affected exons by the mutations (Exons), related amino acid modification (Proteins), affected domains of the protein (Domains), and referred literature (Ref.) are listed. Domains attributions were done on the basis of Pfam definition of the proteins. ABD stands for Anticodon Binding Domain, CD for Catalytic Domain, BED for Bacterial Extension Domain, BID for Bacterial Insertion Domain, HR for Hinge Region and MTS for predicted Mitochondrial Targeting Signal. Properties of each wild-type gene (length, gene accession number, total number of exons, mRNA accession number) and corresponding proteins (total length, protein accession number) are given on the first lane of each reported case

*Those three mutations are found on a same patient in a homozygous status

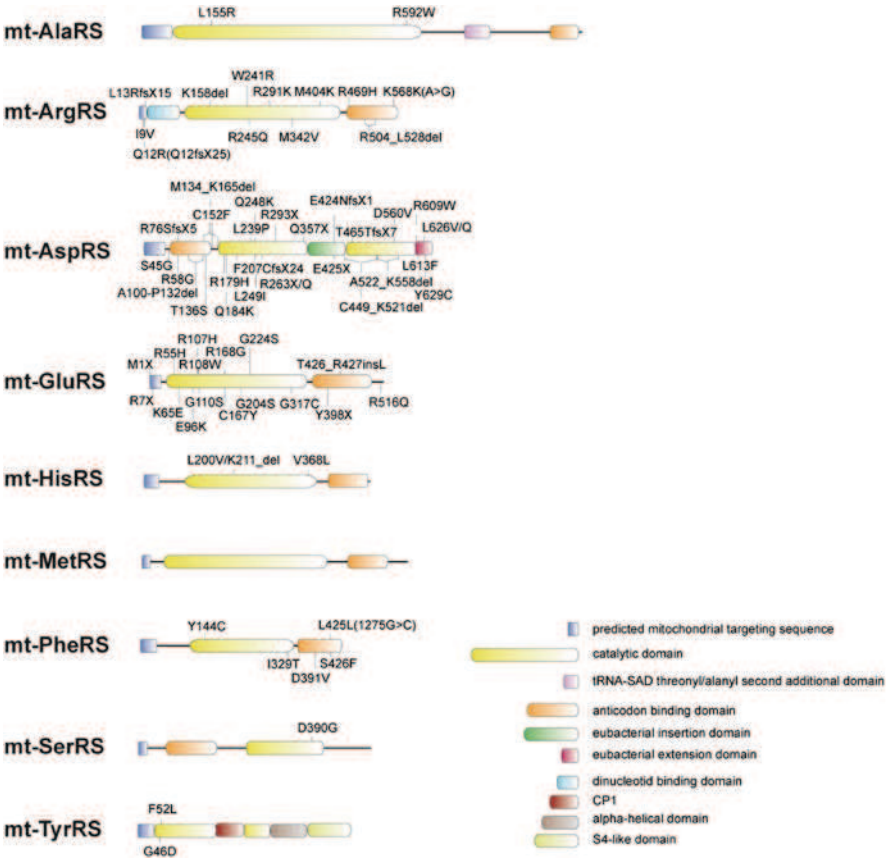


Fig. 6 Display of pathology-related mutations on modular organizations of mt-aaRSs. Color code of all domains is given

are distributed over nearly all exons and, thus, are found in all protein domains, including the predicted MTS (S45G) and the hinge region (C152F and M134_K165del). In addition to these mutations found on one allele of the gene, almost all LBSL patients have a mutation in a polypyrimidine tract at the 3′-end of intron 2, which is found on the second allele. This mutation affects correct splicing of the third exon, which leads to a frameshift and a premature stop (R76SfsX5). This frameshift mutation is "leaky," leading to a decrease but not zero expression of full-length mt-AspRS. Beside these typically compound heterozygous states of LBSL patients, homozygous *DARS2* mutations have been described more recently. Homozygous patients harbor either the R76SfsX5 mutation or the R609W mutation (in the bacterial insertion domain).

The *EARS2* gene codes for mt-GluRS and is located on chromosome XVI. It comprises 32,980 bp, nine exons and encodes a 523 aa long protein. Today, 16 different mutations are known: 3 nonsenses, 12 missenses, and 1 insertion.

The catalytic domain is dominantly affected. The two mutations (R7X and M1X) localized in the predicted MTS are nonsense mutation and certainly lead to truncated translation products. Among those, only one mutation (K65E) is in a homozygous state. The gene *RARS2* codes for mt-ArgRS, is located on chromosome VI, consists of 75,618 bp, 20 exons, and codes for a 562 aa long protein. Mutations in this gene affect all parts of the protein except the dinucleotide-binding domain. Ten different mutations are reported: a single nonsense (Q12fsX25), seven missenses, and two deletions (K158del, R504_L528del). In addition, three homozygous patients were found harboring a combination of two silent mutations (K568K and K291R) with one additional mutation (IVS2 + 5) causing exon 2 skipping. The patient's major transcript lacked exon 2, but a faint, normal-sized fragment was also seen. The gene *FARS2* is 403,026 bp long, codes for mt-PheRS, and is located on chromosome VI. It has six exons and codes for a 451 aa long protein. Today four different missense mutations have been reported. Just the Y144C one (in the anticodon binding domain) is in a homozygote state. The gene *HARS2* codes for mt-HisRS, is located on chromosome V, and consists of 6,900 bp and 13 exons. The corresponding protein is 506 aa long. There is presently only one reported case in which patients are compound heterozygotes. The mutation on one allele is a missense (V368L). The mutation on the second allele produces either a missense replacement (L200V) or creates an additional splice site, inducing the deletion of 11 aa (L200_K211del). The *YARS2* gene, coding for mt-TyrRS, is localized on chromosome XII, is 8,630 bp long, harbors five exons, and codes for a 477 aa long protein. For this aaRSs, only two homozygote mutations are known, both being localized in the catalytic domain. The *AARS2* gene codes for mt-AlaRS, is 13,672 bp long, localized on chromosome VI, has 22 exons, and codes for a 985 aa long protein. There are presently two-reported cases in which patients are harboring missense mutations (either homozygote or compound heterozygote), all localized in the catalytic domain. The *SARS2* gene codes for mt-SerRS, is located on chromosome XIX, is 15,127 bp long, and contains 16 exons. Only one homozygote mutation is reported, localized in the catalytic domain of the 518 aa long protein. It should be noted that the *MARS2* gene, located on chromosome II, is composed of just one exon of 1,779 bp length. It codes for a 593 aa long protein, named mt-MetRS. Interestingly, no "classical" mutation has been reported for this gene but complex rearrangements were shown to be the cause of ARSAL. The single-exon composition of the gene permits duplication events of either the full exon or part of it, leading to the homozygous state or compound heterozygous state of the patients. An additional situation with a large insertion in one of the alleles has also been reported.

In summary, none of the chromosomes is a "hot spot" for pathology-related mutations affecting mt-aaRSs, and neither exons nor protein domains have obvious favored mutation sites. Presently, *DARS2* is the most frequently hit gene. However this may not indicate a peculiar mutational exposure of this gene, but is more likely to be due to intensive investigations of this firstly reported example of an mt-aaRS gene correlated with a disease.

3.3 Compound Heterozygous vs. Homozygous States

Among the 64 reported combinations of mutations, 53 are compound heterozygous and 11 are homozygous (excluding the puzzling combinations found for *MARS2*; see below). In all cases, the parents are unaffected heterozygous carriers of one mutation, leading to autosomal recessive mutations that affect the two alleles in the children. Figure 7 schematically summarizes all observed situations that are combining splicing, missense, nonsense, deletion, and rearrangements defects.

Compound heterozygous status is dominantly observed. In most of the cases, mutations in the first allele produce a splicing defect, resulting in reduced expression of the protein. The second allele carries a missense mutation. Such combinations were reported for *DARS2* [82–85] and *RARS2* [86, 87]. In both situations the splicing defect is "leaky," so that a small amount of wild-type protein remains expressed, which is likely to be sufficient to support basal aminoacylation activity. The residual expression of wild-type protein is mandatory in the situation where the splicing defect is combined with nonsense mutations (which completely abolishes protein expression) as was reported for, e.g., *DARS2* (R76SfsX5/Q357X [88]) and *RARS2* (Q12fsX25/L13RfsX15 [86]). Other examples of nonsense mutations are found, but combined with missense mutations of likely moderate consequences. This is, for instance, reported for *EARS2* (R7X or Y398X combined with R108W [89]). Another possible impact of mutation is the deletion of one or more amino acids within the protein. Mt-HisRS harboring the L200_K211 deletion has been suggested to be unstable when mutants are transiently expressed in human cells [90]. Deletions have also been reported in *RARS2* and *DARS2* (combined in the latest with the abundant R76SfsX5 mutation). However, their possible impacts on protein expression level or stability remains unclear [82, 83, 91, 92]. As a last example, combinations of two different missense mutations have been reported for *EARS2*, *FARS2*, *RARS2*, *DARS2*, and *AARS2*. In these cases, it is assumed that proteins are expressed but folding, structure, stability, and/or activity could be affected (see below).

Figure 7 recalls the natural oligomeric status of mt-aaRSs: some are monomers (mt-PheRS, mt-GluRS, mt-ArgRS), some are dimers (mt-AspRS, mt-SerRS, mt-TyRS, mt-HisRS, mt-MetRS), and one (mt-AlaRS) is homotetramer. A consequence of heterozygosity is the production of two distinct polypeptide chains that can randomly associate to build a dimer (or a tetramer), which theoretically leads to an equal proportion between the two possible homodimers (each harboring the same mutation) and the heterodimer (where each constitutive monomer is harboring a different mutation). However, it has been reported that some of the mutations have an impact on the oligomerization rate, leaving out the random association of mutated polypeptide chains. For instance, mt-HisRS, having the V368L mutation, oligomerizes more efficiently than any combinations, including the wild-type polypeptide [90].

The discovery of *homozygous mutations* was quite unexpected. In fact, mutations in *DARS2* were found initially only in a compound heterozygous state, suggesting that the activity of mutant mt-AspRS homodimers may be incompatible

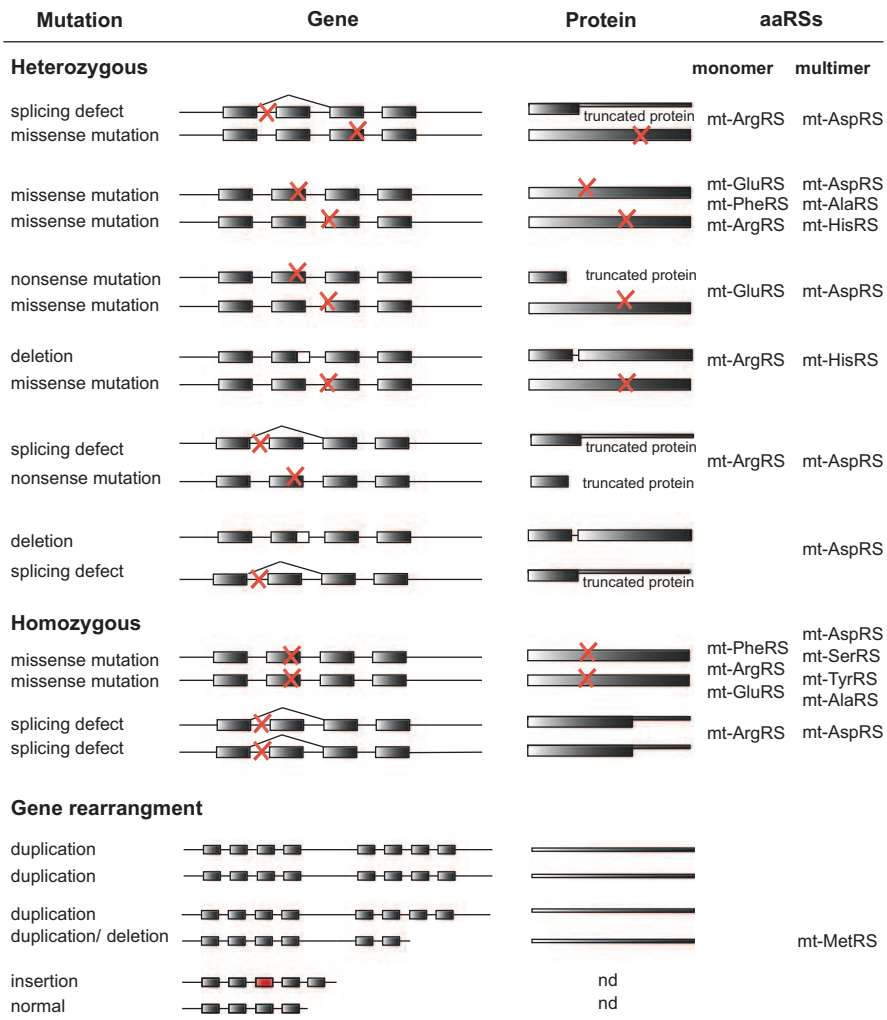


Fig. 7 From the gene to the protein: impact of pathology related mutations. The schematic views illustrate reported combinations of compound heterozygous and homozygous mutations (*red crosses*), and gene rearrangements on a genetic level, as well as possible impacts on corresponding proteins. Exonic organization is represented here in a simplified way (*gray boxes* schematize exons). Truncated proteins result from either nonsense mutations or splicing defects (reported to be "leaky" so that a reduced amount of full-length protein remains expressed, as shown by the *thinner bars*). The natural oligomeric status of mt-aaRSs is recalled on the right. *nd* stands for not determined

with life. It was then proposed that dimers carrying two different mutations would have more residual functional activity than those carrying the same mutation, and thus would not yield to a lethal phenotype [93]. The recent discovery of homozygous mutations of *DARS2* correlated with LBSL in a German patient and in a Japanese family [94, 95] ruled out this conviction. Therefore, the sole

possibility for homozygous mutations to be compatible with the survival of the patient is that the mutation does not exert a too severe effect. As an example, the homozygous R76fsX5 mutation in *DARS2* from an LBSL patient induces a splicing defect that was demonstrated to be "leaky." This allows the expression of a small amount of wild-type protein likely to be sufficient to support some basal activity [96]. Today, homozygous missense mutations are found in *AARS2*, *DARS2*, *EARS2*, *FARS2*, *RARS2*, *YARS2*, and *SARS2*. Their "moderate" effect is not obvious since none of the observed-mutations (e.g., K65E in *EARS2*, Y144C in *FARS2*, D390G in *SARS2*, or F52L and G46D in *YARS2*) conserve any of the physico-chemical properties of the amino acids (isostericity, net charge, hydrophobicity, . . .). It has however been demonstrated that mt-TyrRS carrying the F52L mutation remains catalytically active, with only a twofold reduced aminoacylation rate [97], emphasizing that volume/polarity of the amino acid cannot be the only parameter to take into account. Possible neighborhood effect and structural impacts are discussed below.

In this discussion, *MARS2* is an exception. No "classical" missense or nonsense mutations are observed, but instead complex gene rearrangements have been reported. As already mentioned, this gene is composed of a single exon. A local genomic instability and/or recombination errors have been hypothesized to be the cause of template switching during DNA replication [98]. As a consequence, duplication events of either the full exon or part of it are observed in patients, leading to either homozygous or compound heterozygous states. An additional situation with a large insertion of 300 bp in one of the alleles has also been reported.

4 Diverse Molecular Impacts

A mutation in an mt-aaRS gene can have numerous molecular consequences, affecting either the biogenesis of the enzyme itself, and/or its import and maturation within the organelle, and/or its functional properties. Figure 1 (Part 1) schematized the link of these different steps with ATP synthesis. Biogenesis of mt-aaRSs involves expression and processing of the corresponding mRNA, protein expression, and stability in the cytosol and addressing to mitochondria. Importing into mitochondria requires several steps. It is followed by maturation steps of imported proteins upon entry into mitochondria. Housekeeping function of mature mt-aaRS corresponds to amino acid activation, tRNA recognition, and tRNA aminoacylation. Pathology-related mutations may thus have a direct effect on the mitochondrial translation machinery by impacting mt-aaRS biogenesis, localization (see Sect. 4.1) and/or function (see Sect. 4.2). As a consequence, the translation efficiency/rate of the full set of (or specific) mt-DNA-encoded subunits of respiratory chain complexes may be affected as well, impacting respiratory chain complexes activities and, finally, ATP synthesis (see Sect. 4.3).

4.1 Impact of Mutations on mt-aaRSs Biogenesis

4.1.1 Defects in mt-aaRS mRNAs Processing and Expression

The previously mentioned complex gene rearrangement of *MARS2* leads to an increased amount of transcript (due to duplication of full size gene and/or of regulatory elements). However, no increase in the amount of proteins is observed. It was suggested that corresponding mRNAs undergo transcriptional regulatory event(s) drastically lowering mRNA stability and thus leading to a reduction of 40–80% of the normal protein level [98]. In *DARS2* and *RARS2*, several intronic mutations were reported to affect pre-mRNAs processing. Most of the LBSL patients have mutations (of different types) within intron 2, which affects the correct splicing of exon 3. Similarly, some PCH6 patients have an IVS2 + 5(A > G) mutation, which leads to exon 2 skipping. Exclusion of exon 3 (in *DARS2*) or of exon 2 (in *RARS2*) causes a frameshift (since the two exons are asymmetric) and generates a premature stop codon. In the two situations, it is speculated that the leaky nature of the splicing defect allows for the synthesis of a small (sufficient) amount of wild-type mRNA in most tissues [94, 96, 99, 100]. However, the selective vulnerability within the nervous system is explained by tissue-specific differences in the concentration of the splicing factors (reduced in neural cell) and the presence of rather weak splice sites. In agreement, 5′-splice site of exon 3 in *DARS2* and 3′-splice site of exon 2 in *RARS2* were shown to deviate markedly from the consensus and to have low splicing scores (even in the absence of disease-causing mutations). As a consequence, the exclusion of related exons induced by the mutations is augmented in neural cells [96, 101]. Another example of mis-splicing concerns mutation Q12R(Q12fsX25) in *RARS2*. This mutation interferes with a splicing-enhancer element and causes the retention of 221 bp from intron 1, a consequent frameshift, and the truncation of protein after residue 25 [87].

4.1.2 Defects in mt-aaRSs Expression and Stability

A defect in protein expression and stability can straightforwardly be associated with surveillance pathways, such as, e.g., nonsense-mediated mRNA decay. The main function of these pathways is to reduce errors in gene expression by eliminating mRNA transcripts that contain deletions or premature stop codons (resulting from, for instance, mis-splicing events). However, the easy correlation between decreased mRNA stability and decreased protein stability is not obvious. Pierce and co-workers have identified in *HARS2* a mutation in a compound heterozygous state (L200V), which creates an alternative splice-site and leads to an in frame deletion of 12 codons in exon 6 (L200_K211). The level of the spliced-mRNA is significantly increased in the affected child compared to the unaffected (but carrier) father. Transient expression of the spliced-mRNA into HEK293T cells results in a poorly detectable mutant protein, suggesting its instability.

However, the mechanism involved in stability deficiency for the mutant protein remains unclear [90]. In three of the PCH6 patients, compound heterozygous mutations I9V/R504_L529, R245Q/R469H, and W241R/Q12R were found in *RARS2*. It was shown by western blot experiments on cultured fibroblasts that the expression level of the total proteins was reduced down to approximately 28% of the wild-type content. However, the level of mt-ArgRS-encoding mRNAs, as measured by quantitative PCR, remains low but normal in the patients, excluding transcripts instability induced by the mutations [91]. Three mutations (C152F, Q184K, and D560V) within *DARS2* were also shown to have an impact on the expression of mt-AspRS. Western blotting on transiently transfected HEK293T cells with mutated sequences indicates strongly reduced steady state levels of the mutant proteins. Further analyses on cycloheximide (to inhibit production of newly synthesized proteins) treated cells indicate a decreased stability of ~50% of the three mutant proteins as compared to the wild-type protein [102].

4.1.3 Defects in mt-aaRSs Import

Several pathology-related mutations are found within (or close to) the predicted MTS. Mutations M1X and R7X, in *EARS2*, are nonsense and likely lead to untranslated and truncated products, respectively. None will have the opportunity to be imported into mitochondria [89]. Two mutations, Q12R(Q12fsX25) and I9V, are located within the MTS-encoding sequence of *RARS2*. The mutation Q12R was predicted to enhance import efficiency of mt-ArgRS. As already discussed, this mutation also and mostly interferes with a splicing-enhancer element and causes the retention of 221 bp from intron 1, leading to a frameshift that truncates the protein after residue 25 (Q12fsX25) and likely prevents its import into the mitochondria [87]. The role of the I9V mutation on import of mt-ArgRS into mitochondria is difficult to anticipate since both amino acids are hydrophobic/aliphatic residues, with very similar physico-chemical properties. Predictions on the probability of the I9V mutant to be imported into mitochondria indeed suggest a modest effect of the mutation [91]. Confocal microscopy imaging revealed that mutation S45G located in the predicted MTS of mt-AspRS affects neither the targeting nor the binding of the protein to the mitochondria. However, by combining in vitro import and processing assays, the translocation step was found to be impaired by the mutation [103]. A more recent study investigating the impact of nine mutations found in *DARS2* on sub-cellular localization of mt-AspRS by immuno-cytochemistry (using antibodies against transiently expressed tagged mutant proteins in HEK293T cells) did not confirm a localization defect [102].

4.2 Impact of Mutations on mt-aaRSs Function

4.2.1 Impact on mt-aaRSs Enzyme Activities

The housekeeping activity of the aaRSs is to provide aminoacylated-tRNAs (aa-tRNAs) for translation. The effectiveness of aminoacylation can be detected either by in vitro or by in vivo methods (Table 2). ATP-PPi exchange assays and in vitro aminoacylation reactions can be used to establish kinetic parameters k_{cat}, K_M for ATP, amino acid, and/or tRNA. As examples, mutations L200V and V368L (*HARS2*) and I329T (*FARS2*) were shown to affect ATP-PPi exchange ability of mt-HisRS [90] and mt-PheRS [104], respectively. As an alternative procedure of the ATP-PPi exchange experiment, Cassandrini and co-workers applied a colorimetric-based measurement of Pi production [105]. The authors demonstrated that crude mitochondrial extracts, extracted from cultured skin fibroblasts of patients harboring either R245Q/R469H or W241R/Q12R mutations on mt-ArgRS, have residual activities in Pi formation of only 33% and 19%, respectively. However, they also demonstrate that the level of the protein itself is affected by the mutations [91]. Impacts of mutations on in vitro aminoacylation efficiency of recombinant mt-aaRSs have been investigated at several instances. It has been reported, for example, that mutations within mt-AspRS impact the enzymatic activity (measured as nmol of incorporated amino acid per milligram of enzyme per minute) in a maximum range of ~135-fold (for R263Q, [102]). Other studies have revealed impacts of the Y144C and the F52L mutations on respectively mt-PheRS and mt-TyrRS catalytic efficiencies (relative ratios of k_{cat}/K_M, expressed in s^{-1} μM^{-1}) of 2.3-fold [104] and 9-fold [97]. It should be noted that these effects remain in a moderate range as compared to what has been observed for pathology-related mutations on mt-tRNAs (e.g., >5,000-fold for variants of human mt-$tRNA^{Lys}$, [106]).

Measurement of in vivo steady state levels of aa-tRNAs is performed by separation of aminoacylated- and non-aminoacylated-tRNAs on acidic gels. In vivo steady state levels of aa-tRNAs were investigated using total RNA extracted from patients biopsies or cultured cells and northern blotting on acidic gels (reviewed in, e.g., [107]). The total amount of $tRNA^{Arg}$ is reduced in patient cells by comparison to the amount found in cells from healthy individuals. They observed, however, that remaining $tRNA^{Arg}$ were fully aminoacylated, probably by the few wild-type mt-ArgRS that escaped from the splicing defect engendered by the L13RfsX15 mutation. It is thus suggested that uncharged tRNA might undergo degradation [91, 101]. Belostostky and coworkers showed that the amount of $tRNA^{Ser}$(AGY) in cells harboring the D390G mutation in mt-SerRS was reduced to 10–20% as compared to unaffected control, and that the residual pool of $tRNA^{Ser}$(AGY) was not aminoacylation. Interestingly, the same mutation affected neither stability nor aminoacylation properties of $tRNA^{Ser}$(UCN) [108]. In contrast to these observations, no effect on the steady state level of Met-$tRNA^{Met}$ was observed in patient cells, despite a clear reduced amount of mt-MetRS [98].

Additional methods were developed to investigate the possible impact of mutations on enzyme activity in vivo. For instance, human mutations were modeled on yeast strains, deleted from either the *MSR1* or the *HTS1* gene (homologues of human *RARS2* and *HARS2*, respectively). Homologue mutants of R469H and R245Q were able to complement *MSR1*-deleted strains under fermentable conditions (a situation where the respiratory chain is dispensable), but unable (R469H) or barely able (R245Q) to complement *MSR1*-deleted strain under respiratory conditions. In contrast, a homologue mutant of W241R fully complements the same yeast strain [91]. The homologue mutant L200_K211del was unable to complement a yeast *HTS1*-deleted strain, and the corresponding human sequence couldn't be expressed in bacteria or in human cells, suggesting that this mutant is likely to be unable to provide any activity in vivo and confirming the instability of the protein [90]. As an alternate experimental procedure, the retroviral expression of *YARS2* rescues the translation defect observed in patient muscle cells [109]. Finally, the correlation between mutations in *MARS2* and the ARSAL pathology was confirmed by using the fly as a model organism [98].

As an outcome, defects in the aminoacylation properties of the mt-aaRSs are not necessarily sufficient to explain the pathogenicity of a mutation. Therefore, the cellular environment and additional physiological conditions have to be considered to understand clearly their pathogenic impacts. As examples, tissue-related concentration of the different substrates (e.g., [104, 110]) and/or alternated yet unidentified functions of the mt-aaRSs (e.g., [82, 86, 109]) have been conjectured.

4.2.2 Structure–Function Connections

AaRSs are modular enzymes, composed of well-defined and organized domains, with conserved amino acid residues having either a structural or architectural role, or a function in the chemistry of substrate recognition or in the aminoacylation reaction. Evidence suggests that replacement of key conserved residues may alter essential physico-chemical properties (side chain length, net charge, polarity, hydrophobicity, hydrophilicity, . . .) and thus may have a key impact on the properties of the protein. Present-day knowledge of the 3D structures of mt-aaRSs, or of aaRSs from evolutionary related species, is of great help for connecting the structural impact of pathology-related mutations with its possible functional consequences. For instance, investigation of the crystallographic structures of prokaryotic HisRSs revealed that the mutation L200 and V368 are both implicated in packing interactions with highly conserved hydrophobic amino acids, involved respectively in ATP binding and in histidine recognition [90]. The two mutations (L200V and V368L) may destabilize the packing interactions, engendering movements that are perturbing some secondary structure elements and possibly reducing binding affinities for either ATP or histidine. In agreement, activities of both mutant proteins, measured by ATP-PPi exchange assay, are significantly reduced relative to the wild-type protein. In a second example, the structural and functional impacts of three

pathology-related mutations affecting mt-PheRS could be connected [104]. The crystallographic structure of the enzyme [72, 77] reveals that I329 is located within the ATP-binding site. Replacement of this residue by the small and uncharged threonine should result in a widened ATP-binding site likely decreasing the affinity for the small substrate. In agreement, ATP-PPi exchange kinetic assay confirms a 2.5-fold decreased affinity for ATP for the mutant enzyme, while the binding for phenylalanine remains unaffected. In addition, D391 and Y144 are situated on both sides of the contact surface between the catalytic core and the anticodon-binding domain of mt-PheRS and are stabilizing (by forming hydrogen bounds with key conserved residues) the –closed– state of mt-PheRS. The enzyme was shown to undergo drastic conformational changes upon tRNA binding towards the functional –open– state [72] (Fig. 5). Mutations D391V and Y144C are likely to alter the rotation mechanisms upon tRNA binding and thus to affect the conformational stability of the protein. Along this line, a clear decrease for $tRNA^{Phe}$ binding (increased K_M) was observed, but only for the Y144C mutant enzyme. Instead, a decreased affinity for phenylalanine was measured for the D391V mutant, despite D391 not being situated in the catalytic core. As an explanation, the authors are emphasizing that D391 is involved into a close network of interactions encompassing conserved residues of motif 2 (Y188) and near motif 3 (R330). The D391V replacement may cause R330 and other neighboring residues to adopt different conformations, leading to perturbations in a loop, which is critical for binding and coordination of phenylalanine [104]. The ninefold loss of catalytic efficiency (k_{cat}/K_M) measured for the F52L mutant of mt-TyrRS by in vitro aminoacylation assay [97] might also be explained by the localization of this residue near the catalytic center, as observed within the crystallographic structure of the enzyme [76].

In contrast, 3D representations of aaRSs are not always sufficient to explain or predict the molecular effect of pathology-related mutations. As previously underlined, the V368L mutation was shown to reduce the enzymatic activity of mt-HisRS, in agreement with its location inside the conserved HisB motif (which is specific to HisRSs and contributes to histidine binding pocket). However, this same mutation was also shown to influence the rate of protein dimerization, although it is not localized at the dimerization interface [90]. Similarly, R592 and L155 are respectively situated in 21.15 Å of a site that could have an editing activity, and in the surrounding of conserved catalytic residues of mt-AlaRS. However, their mutations neither affect editing activity nor aminoacylation properties of the enzyme, leaving the connection between structural predictions and functional mechanisms unclear [110]. Mt-SerRS has the functional peculiarity of being able to recognize two isoacceptor tRNAs ($tRNA^{Ser}_{AGY}$ and $tRNA^{Ser}_{UCN}$). Investigation of the crystallographic structure of bovine mt-SerRS revealed key residues responsible for recognition but also discrimination of the two isoacceptors. Those residues are situated within the helical arm of the synthetase and within or flanking the “distal helix” (Fig. 5), shown to be a structural peculiarity of the mitochondrial enzyme [75]. Analysis of aminoacylation properties revealed that the D390G mutation significantly impacts the acylation of $tRNA^{Ser}_{AGY}$ but does not alter that of $tRNA^{Ser}_{UCN}$. Unexpectedly, the D390 residue is not situated in

the isoacceptors discriminating area but in a beta-strand from the catalytic core, far away from the "distal helix" and the helical arm [108].

Despite these few examples where the connection between the structure and the function is not obvious, any future knowledge on crystallographic structures and/or on biophysical properties of mt-aaRSs will be of help to understand the mechanistic aspect and functional impacts of some of the pathology-related mutations. It will also be of help to predict and direct functional investigations. As an example, the crystallographic structure of the yeast cytosolic ArgRS was investigated to assess the possible consequences of the M404K and K158del mutations and to predict a tRNA binding deficiency for the first mutation, and an altered aminoacylation property for the second [92]. Similar predictions could then be drawn for other pathology-related mutations.

4.3 Impact of Mutations on Mitochondrial Translation and Activity of the Respiratory Chain Complexes

As stated above, ATP synthesis is dependent on the coordinated expression of nuclear and mitochondrial genes and requires precise recognition between tRNAs and mt-aaRSs to allow for accurate synthesis of aa-tRNAs. Accordingly, key links between the aminoacylation activity of mt-aaRSs in charge of the synthesis of the 13 mt-DNA-encoded proteins and the activity of respiratory chain (RC) complexes are foreseen. It can be anticipated that any dysfunction of a single macromolecule of the translation machinery may have severe impacts either on the translation and/or on the activity of the mt-DNA-encoded RC subunits (for all complexes except Complex II, of complete nuclear origin). However, this view appears too simplistic. Figure 8 summarizes observed molecular defects (at the levels of mRNAs expression and mt-aaRSs biosynthesis and functioning) and possible defects in the activity of the RC complexes for all reported mutations. It also schematizes that the routes linking the molecular impact of a mutation with its possible phenotypic effect are not yet always fully deciphered.

4.3.1 Translation and/or Activity of the Respiratory Chain Complexes Sub-units are Differentially Affected

Defects in the translation of mt-DNA-encoded RC subunits were reported to be correlated with mutations within *YARS2* [97, 111], *MARS2* [98], and *FARS2* [104]. In addition, defects in the activity of those complexes were reported to be correlated with mutations within *YARS2* [97, 111], *SARS2* [108], *RARS2* [92, 101], *MARS2* [98], *FARS2* [104], *EARS2* [89], *DARS2* [112], and *AARS2* [110]. However, affected subunits may vary from one case to another. Combinations of translation and/or activity defects are numerous. For example, the translation of all complexes is affected by mutation in *YARS2*, while solely the expression of Complex IV

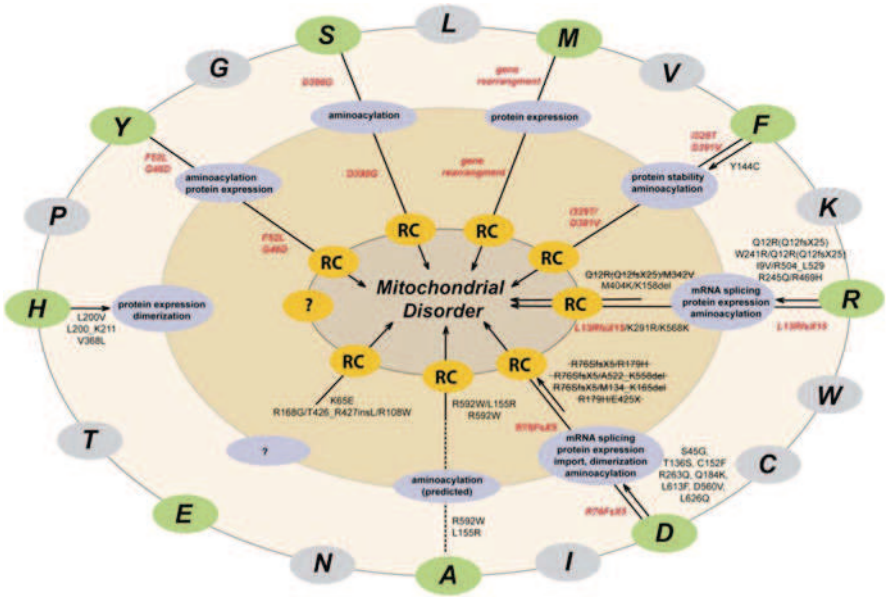


Fig. 8 Summary of the impacts of pathology-related mutations on the different steps of mt-aaRS life cycle and subsequent products activities. Mt-aaRSs (on the *outer circle*) are represented using the one-letter code. Mt-aaRSs affected by pathology-related mutations are in *green*; those yet unrelated to human pathologies are in *gray*. Observed molecular defects at the levels of mRNAs expression and mt-aaRSs biosynthesis and functioning are represented in *blue* in the *middle circle*. Possible defects on activity of the RC complexes are schematized in *orange* in the *inner circle*. The routes that link the molecular impact of a mutation with its possible phenotypic effect and lead to mitochondrial disorders are schematized (*plain* and *dashed arrows*). Mutations displayed in *red* were shown to engender obvious connections between mt-aaRS expression and/or aminoacylation defects with RC dysfunctioning. However, these routes are diverse, and interruptions indicate either the absence of full investigation, or the absence of clear molecular link (*discontinuous* or *dashed arrows*)

is impaired by mutations in *FARS2*. Similarly, the activity of all the complexes was shown to be impaired by a mutation in *SARS2*, while only Complex IV is moderately affected by the M404K/K158del mutations in *RARS2,* leaving unaffected the activities of Complexes I and III. To be noticed as well, gene rearrangement of *MARS2* engenders defects in the translation of all subunits, but noticeably causes activity defects solely in Complex I.

4.3.2 The Same Affected Gene, but Different Impacts on Translation and/or Activity of the Sub-units

A striking observation is that mutations affecting the same gene have different consequences. For instance, mutations M342V/Q12R and IVS2 + 5(A>G)/K291R/K568K, both located in *RARS2*, respectively do not affect and drastically impair respiratory chain complexes [92, 101]. More confusingly, the same combination of

mutations (IVS2 + 5(A>G)/K291R/K568K) does not engender the same effects on three affected individuals. This combination of mutations reduced activities of Complexes I, III, and IV in muscles from patient II-2, but reduced activity of Complexes I in muscles from patient II-4, or activity of Complex IV from patient II-5 [87].

4.3.3 Tissues Specificity

Phenotypic manifestations of mutations have been investigated several times on mitochondria extracted from cell cultures established from biopsies specimens obtained from patients. Here again, discrepancies are observed depending on the investigated tissues. Fibroblasts are frequently used, although investigations performed on those cells showed effects on RC subunits solely for mutations within *EARS2* [89]. Conversely, measurements made on muscle, brain, or heart cells show defects on RC subunits at several instances (muscle cells affected by *DARS2*, *RARS2*, *YARS2*, *SARS2*, *AARS2*, *EARS2*, and *FARS2* mutations, brain cells affected by *AARS2* and *FARS2* mutations, and heart cells affected by *AARS2* mutations). Consequently, discrepancies can be observed when testing the effect of a given mutation on different types of cells. For instance, the F52L mutation found in *YARS2* does not show any impact on mitochondrial translation in fibroblasts, but engenders drastic impairment in skeletal muscle and myotubes [97]. Similarly, mutations in *AARS2* affect the RC activity in only muscle and brain cells, but not in liver cells [110]. Of note, it is not yet possible to make a direct link with the tissue prevalence of the disease (see Table 3) since cellular models haven't been systematically investigated.

In summary, the diverse effects of mutations on the respiratory chain complexes highlight the numerous routes linking the molecular impact of a mutation with its possible phenotypic effect. These routes are summarized and schematized in Fig. 8. Some examples show obvious connections between mt-aaRS expression and/or aminoacylation defects with RC dysfunctioning (e.g., mutations in *MARS2*, *YARS2*, and *SARS2*). Conversely, it can be seen that, for some of the reported cases, investigations are not yet complete so that it is not possible to draw the pathogenic pathway (e.g., mutations in *HARS2*, *EARS2*, *AARS2*, *DARS2*, *RARS2*, and *FARS2*). Another situation is visible in the scheme, where the molecular routes are diverse for the set of mutations affecting the same gene (mutations in, e.g., *DARS2*, *RARS2*, and *FARS2*). It can thus be hypothesized that a diversity of mechanisms, with tissue specificity, contributes to the heterogeneous manifestations of mitochondrial disorders.

5 Outlook

Human mitochondrial aminoacylation systems deserve specific attention as a consequence of their recently recognized connection with human pathologies. Understanding the fundamental properties/peculiarities of aminoacylation systems

is critical to resolving the link between mutation and pathology. In the past 20 years a large number of human neuromuscular and neurodegenerative disorders have been reported to be correlated with point mutations in the mt-mRNAs, mt-rRNAs, and especially in mt-tRNAs (reviewed in, e.g., [5, 6, 10, 113–115]). Among the mutations leading to "mitochondrial disorders," more than 230 are distributed all over the 22 tRNAs. Numerous studies have attempted to unravel the molecular impacts of the mutations on the various properties of the tRNA that lead to a mosaïcity of phenotypic effects. Whilst there is no general rule, a trend towards a structural perturbation as initial molecular impact of mutations can be retained [116, 117]. However, the housekeeping function of tRNAs, namely their capacity to become esterified by an amino acid, is not systematically affected in mutated variants, so that alternative functions of tRNAs or at least alternative partnerships have to be considered [118]. The discovery of a new family of impacted genes, the mt-aaRSs, further extends the complexity of mitochondrial disorders.

The distance between the birth of an mt-aaRS and its final role in ATP synthesis is very long, so that the gap between the comprehension of the molecular impact of mutations in macromolecules and a dysfunction of the respiratory chain is very large. Also, considering the sole/common outcome from any defect in mt-aaRSs (or in mt-tRNAs) results from an effect on the translation rate of each of the 13 mt-DNA-encoded proteins is a view that rapidly appeared too simplistic. Indeed, the major outcome from the present review is that there is no common combination of affected steps that correlates the 64 reported cases (combining 65 mutations within nuclear-encoded mt-aaRS genes) to the various observed phenotypic expressions. There is obviously no "favored" mt-aaRS gene. Nine are reported today, but their recent, rapid and exponential correlations with human pathologies suggest as evidence that all mt-aaRS genes are likely to be affected by pathology-related mutations, but remain to be revealed. There are also no "favored" affected parts of the protein, which is in agreement with the fact that all steps of the mt-aaRS life cycle can be impacted. Finally, despite the fact that primary observations would suggest an exclusive connection of mt-aaRS disorders to the nervous system and to inherited neurological diseases, sporadic manifestations were lately observed in, e.g., skeletal muscle, kidney, lung, and/or heart. Thus, links between the activity of a given aaRS to mitochondrial translation on one hand and ATP production on the other, involve a number of issues that need to be further explored. Those issues may consider a possible combined effect of mutations affecting other gene (s), such as, for instance, affecting the tRNA modifying enzyme as observed in MLASA patients [109]. Those issues should also take into account the possibility that aminoacylation may turn out not to be the sole function of mt-aaRSs in a living cell, and that these enzymes may also participate in other processes and/or be implicated in various fine-tuning mechanisms, as already shown for various bacterial and eukaryal aaRSs; see below. Thus it becomes obvious that we have to integrate mt-aaRSs into a functional network at the cellular level. In other words, it is of outstanding interest to nail down all the potential interacting components of mt-aaRSs and study their dynamic location within the cell.

In support of this assumption, developments in genomics and post-genomics, associated with conventional biochemical studies, led to the finding of unexpected non-conventional auxiliary functions for human cytosolic aaRSs and connections to other cellular activities (reviewed in, e.g., [119]). Examples include enzymes secreted as procytokines that, after activation, operate in pathways linked to the immune system or angiogenesis (e.g., cyt-TyrRS, cyt-TrpRS [120, 121]), or involved in the vascular development (cyt-SerRS [122]). In addition, accumulating evidence indicates that disruption of non-canonical functions of cytosolic aaRSs connects to various types of diseases, including neural pathologic conditions and cancer [123]. For example, point mutations in human cytosolic TyrRS and cytosolic GlyRS are associated with Charcot-Marie-Tooth (CMT) diseases. Examination of the aminoacylation activities demonstrates that CMT disease can occur without loss of aminoacylation activity [124, 125]. Finally, nowadays evidence indicates that macromolecular assemblies might be sources for proteins with auxiliary functions. Multiprotein complexes containing aaRSs are widely found in all three domains of life, playing roles in apoptosis, viral assembly, and regulation of transcription and translation (reviewed in [126]. As an example, the cytosolic translational apparatus in human cells is highly organized. Nine of the cytosolic aaRS are assembled into the MARS complex (with three auxiliary proteins), which has been emphasized as an anchoring platform for multitasking proteins [127–129]. Those were shown to be recognizable not only by displaying atypical functional activities (possibly linked to structural inventions [69, 119, 130]), but also by their atypical cellular organization, atypical selection pressure, or by atypical "omics" behaviors. The organization of the aaRS within human mitochondria remains mostly unknown at present but will merit full attention in the near future.

Acknowledgements We thank Redmond Smyth for many stylistic improvements of the manuscript. Our work is supported by Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg (UdS), and the French National Program "Investissement d'Avenir" (Labex MitCross), administered by the "Agence National de la Recherche," and referenced ANR-10-IDEX-002-02. The ADIRAL association is acknowledged. HS was supported by Région Alsace, Université de Strasbourg, Association Française contre les Mytopathies (AFM) and Fondation des Treilles.

References

1. Scheffler IE (2001) A century of mitochondrial research: achievements and perspectives. Mitochondrion 1:3–31
2. Anderson S, Bankier AT, Barrell BG et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465
3. Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. Nature 290:470–474
4. Christian BE, Spremulli LL (2012) Mechanism of protein biosynthesis in mammalian mitochondria. Biochim Biophys Acta 1819:1035–1054

5. Florentz C, Sohm B, Tryoen-Tóth P et al (2003) Human mitochondrial tRNAs in health and disease. Cell Mol Life Sci 60:1356–1375
6. Suzuki T, Nagao A, Suzuki T (2011) Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases. Annu Rev Genet 45:299–329
7. Watanabe K (2010) Unique features of animal mitochondrial translation systems. The non-universal genetic code, unusual features of the translational apparatus and their relevance to human mitochondrial diseases. Proc Jpn Acad Ser B Phys Biol Sci 86:11–39
8. Bonnefond L, Fender A, Rudinger-Thirion J et al (2005) Toward the full set of human mitochondrial aminoacyl-tRNA synthetases: characterization of AspRS and TyrRS. Biochemistry 44:4805–4816
9. Sissler M, Pütz J, Fasiolo F, Florentz C (2005) Mitochondrial aminoacyl-tRNA synthetases. In: Ibba M, Francklyn C, Cusack S (eds) Aminoacyl-tRNA synthetases. Landes Biosciences, Georgetown, pp 271–284
10. Ylikallio E, Suomalainen A (2012) Mechanisms of mitochondrial diseases. Ann Med 44: 41–59
11. Shoffner JM, Lott MT, Lezza AM et al (1990) Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation. Cell 61: 931–937
12. Goto Y, Nonaka I, Horai S (1990) A mutation in the tRNALeu(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. Nature 348:651–653
13. Dimauro S, Davidzon G (2005) Mitochondrial DNA and disease. Ann Med 37:222–232
14. Diaz F (2010) Cytochrome c oxidase deficiency: patients and animal models. Biochim Biophys Acta 1802:100–110
15. Stumpf JD, Copeland WC (2011) Mitochondrial DNA replication and disease: insights from DNA polymerase γ mutations. Cell Mol Life Sci 68:219–233
16. Scheper GD, Van der Knaap MS, Proud CG (2007) Translation matters: protein synthesis defects in inherited disease. Nat Rev Genet 8:711–723
17. Konovalova S, Tyynismaa H (2013) Mitochondrial aminoacyl-tRNA synthetases in human disease. Mol Genet Metab 108:206–211
18. Reeve AK, Krishnan KJ, Turnbull D (2009) Mitochondrial DNA mutations in disease, aging, and neurodegeneration. Ann N Y Acad Sci 1147:21–29
19. Wallace DC (2010) Mitochondrial DNA mutations in disease and aging. Environ Mol Mutagen 51:440–450
20. McFarland R, Elson JL, Taylor RW et al (2004) Assigning pathogenicity to mitochondrial tRNA mutations: when "definitely maybe" is not good enough. Trends Genet 20:591–596
21. Zeviani M, Di Donato S (2004) Mitochondrial disorders. Brain 127:2153–2172
22. Smits P, Smeitink J, Van den Heuvel L (2010) Mitochondrial translation and beyond: processes implicated in combined oxidative phosphorylation deficiencies. J Biomed Biotechnol 2010:737385
23. Fernández-Silva P, Acín-Pérez R, Fernández-Vizarra E et al (2007) In vivo and in organello analyses of mitochondrial translation. Methods Cell Biol 80:571–588
24. DiMauro S, Hirano M (2003) Mitochondrial DNA deletion syndromes. In: Pagon RA, Bird TD, Dolan CR, Stephens K, Adm MP (eds) GeneReviews, Seattle
25. Kunz WS, Kudin A, Vielhaber S et al (2000) Flux control of cytochrome c oxidase in human skeletal muscle. J Biol Chem 275:27741–27745
26. Munnich A, Rustin P (2001) Clinical spectrum and diagnosis of mitochondrial disorders. Am J Med Genet 106:4–17
27. Barrientos A (2002) In vivo and in organello assessment of OXPHOS activities. Methods 26:307–316
28. Brand MD, Nicholls DG (2011) Assessing mitochondrial dysfunction in cells. Biochem J 435:297–312
29. Chance B, Williams GR (1955) A simple and rapid assay of oxidative phosphorylation. Nature 175:1120–1121

30. N'Guessan B, Zoll J, Ribera F et al (2004) Evaluation of quantitative and qualitative aspects of mitochondrial function in human skeletal and cardiac muscles. Mol Cell Biochem 256–257:267–280
31. Veksler VI, Kuznetsov AV, Sharov VG et al (1987) Mitochondrial respiratory parameters in cardiac tissue: a novel method of assessment by using saponin-skinned fibers. Biochim Biophys Acta 892:191–196
32. Letellier T, Malgat M, Coquet M et al (1992) Mitochondrial myopathy studies on permeabilized muscle fibers. Pediatr Res 32:17–22
33. Bouitbir J, Charles A-L, Echaniz-Laguna A et al (2012) Opposite effects of statins on mitochondria of cardiac and skeletal muscles: a "mitohormesis" mechanism involving reactive oxygen species and PGC-1. Eur Heart J 33:1397–1407
34. Nijtmans LG, Henderson NS, Holt IJ (2002) Blue native electrophoresis to study mitochondrial and other protein complexes. Methods 26:327–334
35. Brown WM, George M Jr, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA 76:1967–1971
36. Castellana S, Vicario S, Saccone C (2011) Evolutionary patterns of the mitochondrial genome in Metazoa: exploring the role of mutation and selection in mitochondrial protein coding genes. Genome Biol Evol 3:1067–1079
37. Giegé R, Jühling F, Pütz J et al (2012) Structure of transfer RNAs: similarity and variability. Wiley Interdiscip Rev RNA 3:37–61
38. Helm M, Brulé H, Friede D et al (2000) Search for characteristic structural features of mammalian mitochondrial tRNAs. RNA 6:1356–1379
39. Mudge SJ, Williams JH, Eyre HJ et al (1998) Complex organisation of the 5′-end of the human glycine tRNA synthetase gene. Gene 209:45–50
40. Shiba K, Schimmel P, Motegi H, Noda T (1994) Human glycyl-tRNA synthetase. Wide divergence of primary structure from bacterial counterpart and species-specific aminoacylation. J Biol Chem 269:30049–30055
41. Tolkunova E, Park H, Xia J et al (2000) The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual alternative splicing of the primary transcript. J Biol Chem 275:35063–35069
42. Rinehart J, Krett B, Rubio MA et al (2005) Saccharomyces cerevisiae imports the cytosolic pathway for Gln-tRNA synthesis into the mitochondrion. Genes Dev 19:583–592
43. Ibba M, Soll D (2000) Aminoacyl-tRNA synthesis. Annu Rev Biochem 69:617–650
44. Pujol C, Bailly M, Kern D et al (2008) Dual-targeted tRNA-dependent amidotransferase ensures both mitochondrial and chloroplastic Gln-tRNAGln synthesis in plants. Proc Natl Acad Sci USA 105:6481–6485
45. Schön A, Kannangara CG, Gough S, Söll D (1988) Protein biosynthesis in organelles requires misaminoacylation of tRNA. Nature 331:187–190
46. Frechin M, Duchêne A-M, Becker HD (2009) Translating organellar glutamine codons: a case by case scenario? RNA Biol 6:31–34
47. Nagao A, Suzuki T, Katoh T et al (2009) Biogenesis of glutaminyl-mt tRNAGln in human mitochondria. Proc Natl Acad Sci USA 106:16209–16214
48. Frechin M, Senger B, Brayé M et al (2009) Yeast mitochondrial Gln-tRNA(Gln) is generated by a GatFAB-mediated transamidation pathway involving Arc1p-controlled subcellular sorting of cytosolic GluRS. Genes Dev 23:1119–1130
49. Alfonzo JD, Söll D (2009) Mitochondrial tRNA import–the challenge to understand has just begun. Biol Chem 390:717–722
50. Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. Science 283:1476–1481
51. Brindefalk B, Viklund J, Larsson D et al (2007) Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases. Mol Biol Evol 24:743–756
52. Pfanner N (2000) Protein sorting: recognizing mitochondrial presequences. Curr Biol 10: R412–R415
53. Baker MJ, Frazier AE, Gulbis JM, Ryan MT (2007) Mitochondrial protein-import machinery: correlating structure with function. Trends Cell Biol 17:456–464

54. Becker T, Böttinger L, Pfanner N (2012) Mitochondrial protein import: from transport pathways to an integrated network. Trends Biochem Sci 37:85–91
55. Bolender N, Sickmann A, Wagner R et al (2008) Multiple pathways for sorting mitochondrial precursor proteins. EMBO Rep 9:42–49
56. Gakh O, Cavadini P, Isaya G (2002) Mitochondrial processing peptidases. Biochim Biophys Acta 1592:63–77
57. Van der Laan M, Hutu DP, Rehling P (2010) On the mechanism of preprotein import by the mitochondrial presequence translocase. Biochim Biophys Acta 1803:732–739
58. Neupert W, Herrmann JM (2007) Translocation of proteins into mitochondria. Annu Rev Biochem 76:723–749
59. Schmidt O, Pfanner N, Meisinger C (2010) Mitochondrial protein import: from proteomics to functional mechanisms. Nat Rev Mol Cell Biol 11:655–667
60. Vögtle F-N, Wortelkamp S, Zahedi RP et al (2009) Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. Cell 139:428–439
61. Varshavsky A (2011) The N-end rule pathway and regulation by proteolysis. Protein Sci 20(8):1298–1345
62. Sissler M, Lorber B, Messmer M et al (2008) Handling mammalian mitochondrial tRNAs and aminoacyl-tRNA synthetases for functional and structural characterization. Methods 44: 176–189
63. Bullard JM, Cai YC, Spremulli LL (2000) Expression and characterization of the human mitochondrial leucyl-tRNA synthetase. Biochim Biophys Acta 1490:245–258
64. Yao Y-N, Wang L, Wu X-F, Wang E-D (2003) Human mitochondrial leucyl-tRNA synthetase with high activity produced from Escherichia coli. Protein Expr Purif 30:112–116
65. Gaudry A, Lorber B, Neuenfeldt A et al (2012) Re-designed N-terminus enhances expression, solubility and crystallizability of mitochondrial protein. Protein Eng Des Sel 25:473–481
66. Neuenfeldt A, Lorber B, Ennifar E et al (2012) Thermodynamic properties distinguish human mitochondrial aspartyl-tRNA synthetase from bacterial homolog with same 3D architecture. Nucleic Acids Res 41:2698–2708
67. Cusack S, Berthet-Colominas C, Härtlein M et al (1990) A second class of synthetase structure revealed by X-ray analysis of Escherichia coli seryl-tRNA synthetase at 2.5Å. Nature 347:249–255
68. Eriani G, Delarue M, Poch O et al (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. Nature 347:203–206
69. Guo M, Schimmel P, Yang X-L (2010) Functional expansion of human tRNA synthetases achieved by structural inventions. FEBS Lett 584:434–442
70. Bullard JM, Cai YC, Demeler B, Spremulli LL (1999) Expression and characterization of a human mitochondrial phenylalanyl-tRNA synthetase. J Mol Biol 288:567–577
71. Sanni A, Walter P, Boulanger Y et al (1991) Evolution of aminoacyl-tRNA synthetase quaternary structure and activity: Saccharomyces cerevisiae mitochondrial phenylalanyl-tRNA synthetase. Proc Natl Acad Sci USA 88:8387–8391
72. Klipcan L, Moor N, Finarov I et al (2012) Crystal structure of human mitochondrial PheRS complexed with tRNAPhe in the active "open" state. J Mol Biol 415:527–537
73. Kaiser E, Hu B, Becher S et al (1994) The human EPRS locus (formerly the QARS locus): a gene encoding a class I and a class II aminoacyl-tRNA synthetase. Genomics 19:280–290
74. Bhat TN, Blow DM, Brick P, Nyborg J (1982) Tyrosyl-tRNA synthetase forms a mononucleotide-binding fold. J Mol Biol 158:699–709
75. Chimnaronk S, Gravers Jeppesen M, Suzuki T et al (2005) Dual-mode recognition of non-canonical tRNAsSER by seryl-tRNA synthetase in mammalian mitochondria. EMBO J 24: 3369–3379
76. Bonnefond L, Frugier M, Touzé E et al (2007) Crystal structure of human mitochondrial tyrosyl-tRNA synthetase reveals common and idiosyncratic features. Structure 15: 1505–1516
77. Klipcan L, Levin I, Kessler N et al (2008) The tRNA-induced conformational activation of human mitochondrial phenylalanyl-tRNA synthetase. Structure 16:1095–1104

78. Fender A, Sauter C, Messmer M et al (2006) Loss of a primordial identity element for a mammalian mitochondrial aminoacylation system. J Biol Chem 281:15980–15986
79. Messmer M, Blais SP, Balg C et al (2009) Peculiar inhibition of human mitochondrial aspartyl-tRNA synthetase by adenylate analogs. Biochimie 91:596–603
80. Fender A, Gaudry A, Jühling F et al (2012) Adaptation of aminoacylation identity rules to mammalian mitochondria. Biochimie 94:1090–1097
81. Kumazawa Y, Himeno H, Miura K, Watanabe K (1991) Unilateral aminoacylation specificity between bovine mitochondria and eubacteria. J Biochem 109:421–427
82. Scheper GC, Van der Klok T, Van Andel RJ et al (2007) Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. Nat Genet 39:534–539
83. Isohanni P, Linnankivi T, Buzkova J et al (2010) DARS2 mutations in mitochondrial leucoencephalopathy and multiple sclerosis. J Med Genet 47:66–70
84. Labauge P, Dorboz I, Eymard-Pierre E et al (2011) Clinically asymptomatic adult patient with extensive LBSL MRI pattern and DARS2 mutations. J Neurol 258:335–337
85. Lin J, Chiconelli Faria E, Da Rocha AJ et al (2010) Leukoencephalopathy with brainstem and spinal cord involvement and normal lactate: a new mutation in the DARS2 gene. J Child Neurol 25:1425–1428
86. Namavar Y, Barth PG, Kasher PR et al (2011) Clinical, neuroradiological and genetic findings in pontocerebellar hypoplasia. Brain 134:143–156
87. Rankin J, Brown R, Dobyns WB et al (2010) Pontocerebellar hypoplasia type 6: a British case with PEHO-like features. Am J Med Genet A 152A:2079–2084
88. Sharma S, Sankhyan N, Kumar A et al (2011) Leukoencephalopathy with brain stem and spinal cord involvement and high lactate: a genetically proven case without elevated white matter lactate. J Child Neurol 26:773–776
89. Steenweg ME, Ghezzi D, Haack T et al (2012) Leukoencephalopathy with thalamus and brainstem involvement and high lactate 'LTBL' caused by EARS2 mutations. Brain 135:1387–1394
90. Pierce SB, Chisholm KM, Lynch ED et al (2011) Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. Proc Natl Acad Sci USA 108:6543–6548
91. Cassandrini D, Cilio MR, Bianchi M et al (2012) Pontocerebellar hypoplasia type 6 caused by mutations in RARS2: definition of the clinical spectrum and molecular findings in five patients. J Inherit Metab Dis 36:43–53
92. Glamuzina E, Brown R, Hogarth K et al (2012) Further delineation of pontocerebellar hypoplasia type 6 due to mutations in the gene encoding mitochondrial arginyl-tRNA synthetase, RARS2. J Inherit Metab Dis 35:459–467
93. Antonellis A, Green ED (2008) The role of aminoacyl-tRNA synthetases in genetic diseases. Annu Rev Genomics Hum Genet 9:87–107
94. Miyake N, Yamashita S, Kurosawa K et al (2011) A novel homozygous mutation of DARS2 may cause a severe LBSL variant. Clin Genet 80:293–296
95. Synofzik M, Schicks J, Lindig T et al (2011) Acetazolamide-responsive exercise-induced episodic ataxia associated with a novel homozygous DARS2 mutation. J Med Genet 48: 713–715
96. Van Berge L, Dooves S, Van Berkel CG et al (2012) Leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation is associated with cell-type-dependent splicing of mtAspRS mRNA. Biochem J 441:955–962
97. Riley LG, Cooper S, Hickey P et al (2010) Mutation of the mitochondrial tyrosyl-tRNA synthetase gene, YARS2, causes myopathy, lactic acidosis, and sideroblastic anemia–MLASA syndrome. Am J Hum Genet 87:52–59
98. Bayat V, Thiffault I, Jaiswal M et al (2012) Mutations in the mitochondrial methionyl-tRNA synthetase cause a neurodegenerative phenotype in flies and a recessive ataxia (ARSAL) in humans. PLoS Biol 10:e1001288

99. Mierzewska H, Van der Knaap MS, Scheper GC et al (2011) Leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation in the first Polish patient. Brain Dev 33:713–717
100. Yamashita S, Miyake N, Matsumoto N et al (2012) Neuropathology of leukoencephalopathy with brainstem and spinal cord involvement and high lactate caused by a homozygous mutation of DARS2. Brain Dev 35:312–316
101. Edvardson S, Shaag A, Kolesnikova O et al (2007) Deleterious mutation in the mitochondrial arginyl-transfer RNA synthetase gene is associated with pontocerebellar hypoplasia. Am J Hum Genet 81:857–862
102. Van Berge L, Kevenaar J, Polder E et al (2012) Pathogenic mutations causing LBSL affect mitochondrial aspartyl-tRNA synthetase in diverse ways. Biochem J 450:345–350
103. Messmer M, Florentz C, Schwenzer H et al (2011) A human pathology-related mutation prevents import of an aminoacyl-tRNA synthetase into mitochondria. Biochem J 433: 441–446
104. Elo JM, Yadavalli SS, Euro L et al (2012) Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy. Hum Mol Genet 21:4521–4529
105. Chang GG, Pan F, Yeh C, Huang TM (1983) Colorimetric assay for aminoacyl-tRNA synthetases. Anal Biochem 130:171–176
106. Sissler M, Helm M, Frugier M et al (2004) Aminoacylation properties of pathology-related human mitochondrial tRNA(Lys) variants. RNA 10:841–853
107. Köhrer C, Rajbhandary UL (2008) The many applications of acid urea polyacrylamide gel electrophoresis to studies of tRNAs and aminoacyl-tRNA synthetases. Methods 44:129–138
108. Belostotsky R, Ben-Shalom E, Rinat C et al (2011) Mutations in the mitochondrial seryl-tRNA synthetase cause hyperuricemia, pulmonary hypertension, renal failure in infancy and alkalosis, HUPRA syndrome. Am J Hum Genet 88:193–200
109. Sasarman F, Nishimura T, Thiffault I, Shoubridge EA (2012) A novel mutation in YARS2 causes myopathy with lactic acidosis and sideroblastic anemia. Hum Mutat 33:1201–1206
110. Götz A, Tyynismaa H, Euro L et al (2011) Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. Am J Hum Genet 88:635–642
111. Sasarman F, Karpati G, Shoubridge EA (2002) Nuclear genetic control of mitochondrial translation in skeletal muscle revealed in patients with mitochondrial myopathy. Hum Mol Genet 11:1669–1681
112. Orcesi S, La Piana R, Uggetti C et al (2011) Spinal cord calcification in an early-onset progressive leukoencephalopathy. J Child Neurol 26:876–880
113. Rötig A (2011) Human diseases with impaired mitochondrial protein synthesis. Biochim Biophys Acta 1807:1198–1205
114. Wallace DC (1999) Mitochondrial diseases in man and mouse. Science 283:1482–1488
115. Yarham JW, Elson JL, Blakely EL et al (2010) Mitochondrial tRNA mutations and disease. Wiley Interdiscip Rev RNA 1:304–324
116. Levinger L, Mörl M, Florentz C (2004) Mitochondrial tRNA 3′ end metabolism and human disease. Nucleic Acids Res 32:5430–5441
117. Wittenhagen LM, Kelley SO (2003) Impact of disease-related mitochondrial mutations on tRNA structure and function. Trends Biochem Sci 28:605–611
118. Jacobs HT, Holt IJ (2000) The np 3243 MELAS mutation: damned if you aminoacylate, damned if you don't. Hum Mol Genet 9:463–465
119. Guo M, Schimmel P (2013) Essential nontranslational functions of tRNA synthetases. Nat Chem Biol 9:145–153
120. Wakasugi K, Slike BM, Hood J et al (2002) Induction of angiogenesis by a fragment of human tyrosyl-tRNA synthetase. J Biol Chem 277:20124–20126
121. Wakasugi K, Slike BM, Hood J et al (2002) A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. Proc Natl Acad Sci USA 99:173–177

122. Kawahara A, Stainier DYR (2009) Noncanonical activity of seryl-transfer RNA synthetase and vascular development. Trends Cardiovasc Med 19:179–182
123. Park SG, Schimmel P, Kim S (2008) Aminoacyl tRNA synthetases and their connections to disease. Proc Natl Acad Sci 105:11043–11049
124. Seburn KL, Nangle LA, Cox GA et al (2006) An active dominant mutation of glycyl-tRNA synthetase causes neuropathy in a Charcot-Marie-Tooth 2D mouse model. Neuron 51:715–726
125. Storkebaum E, Leitão-Gonçalves R, Godenschwege T et al (2009) Dominant mutations in the tyrosyl-tRNA synthetase gene recapitulate in Drosophila features of human Charcot-Marie-Tooth neuropathy. Proc Natl Acad Sci USA 106:11782–11787
126. Hausmann CD, Ibba M (2008) Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. FEMS Microbiol Rev 32:705–721
127. Han JM, Lee MJ, Park SG et al (2006) Hierarchical network between the components of the multi-tRNA synthetase complex: implications for complex formation. J Biol Chem 281: 38663–38667
128. Kaminska M, Havrylenko S, Decottignies P et al (2009) Dissection of the structural organization of the aminoacyl-tRNA synthetase complex. J Biol Chem 284:6053–6060
129. Ray PS, Arif A, Fox PL (2007) Macromolecular complexes as depots for releasable regulatory proteins. Trends Biochem Sci 32:158–164
130. Guo M, Yang X-L, Schimmel P (2010) New functions of aminoacyl-tRNA synthetases beyond translation. Nat Rev Mol Cell Biol 11:668–674
131. Uluc K, Baskan O, Yildirim KA et al (2008) Leukoencephalopathy with brain stem and spinal cord involvement and high lactate: a genetically proven case with distinct MRI findings. J Neurol Sci 273:118–122
132. Tzoulis C, Tran GT, Gjerde IO et al (2012) Leukoencephalopathy with brainstem and spinal cord involvement caused by a novel mutation in the DARS2 gene. J Neurol 259:292–296
133. Talim B, Pyle A, Griffin H et al (2013) Multisystem fatal infantile disease caused by a novel homozygous EARS2 mutation. Brain 136:e228
134. Shamseldin HE, Alshammari M, Al-Sheddi T et al (2012) Genomic analysis of mitochondrial diseases in a consanguineous population reveals novel candidate disease genes. J Med Genet 49:234–241

5 Non-translational functions of aminoacyl-tRNA synthetases

One of the major outcomes of the systematic analysis of the pathological implication of mt-aaRS in mitochondrial disorders is the assumption that aminoacylation may not be the sole function of mt-aaRSs in a living cell. It has been suggested that these enzymes may also participate in other processes and/or be implicated in various fine-tuning mechanisms, as already shown for various bacterial and eukaryal aaRSs. The following paragraphs give a brief overview on the different non-translational (non-canonical) function(s) of cytosolic aaRSs, the triggers leading to the non-translational function and connection of them with human diseases.

5.1 Organization of cytosolic aaRSs

For more than 20 years it is known that cytosolic synthetases are partly organized in multi aminoacyl-tRNA synthetase complex (MARS) of approximately 1.5 MDa. This ubiquitous complex contains in higher eukaryotes nine synthetases (ArgRS, AspRS, GlnRS, IleRS, LeuRS, LysRS, MetRS and fusion protein GluRS-ProRS (EPRS)) and three scaffold proteins (p18, p38 and p43) (Mirande et al., 1982). Several approaches, such as cross-linking and pull down were used to dissect the organization of the macromolecular complex. It has been suggested that the non-synthetase proteins form a platform that stabilizes the complex, with p38 serving as an interaction partner with the majority of the aaRSs. The complex arrangement is grouped in two sub-domains depending on the affiliation of the proteins to p38 (N-or C-terminal interaction) (**Figure 6**). It has been proposed that this stable complex serves as a “depot” in which aaRSs fulfill their canonical function, but also from which released proteins have additional non-translational (“moonlight”) function(s) (Ray et al., 2007). Of note, the involvement of aaRSs in multi-protein complexes exist also in other domains of life and lower eukaryotes, such as the

transamidosome in bacteria (AspRS in complex with GatCAB amidotransferase) or the MetRS:Arc1p:GluRS complex in yeast (reviewed in e.g. Hausmann and Ibba, 2008).

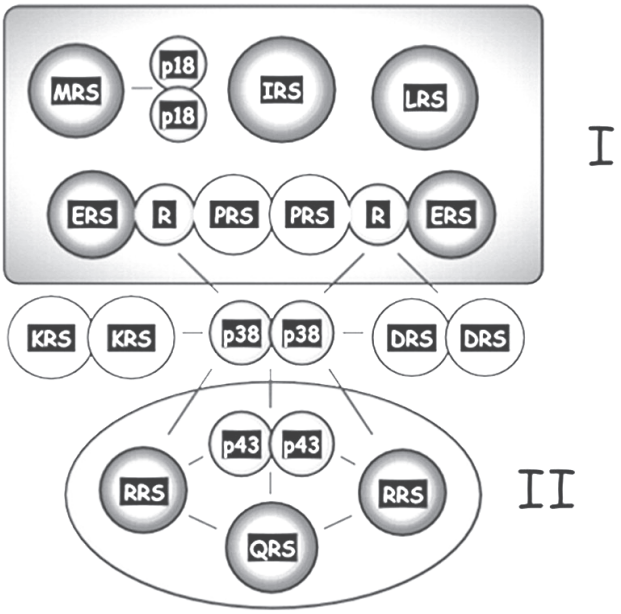


Figure 6: Organization of the human cytosolic multi-aaRS complex (MARS) (Kaminska et al., 2009)**.** AaRSs are illustrated in the one-letter amino acid code. Roman numbers indicate the two sub-domains of the MARS complex.

5.2 AaRSs beyond translation

In addition to their essential role in translation, unexpected alternative functions have been reported for many of the cytosolic aaRSs over the last decade, connecting them to additional pathways. These pathways concern glucose and amino acid metabolism (e.g. LeuRS), tissue development (e.g. SerRS), angiogenesis (e.g. TyrRS or TrpRS), immune (e.g. LysRS or AsnRS) and inflammatory response (e.g. GluProRS or LysRS), tumorigenesis or neuronal system (e.g. GlyRS or AlaRS) (review e.g. Park et al., 2008; Guo and Schimmel, 2013). Beside the aaRSs, the three scaffold proteins of the MARS complex are also involved in several alternative pathways connected to turmorgenesis (p18, p38 and p43) or the inflammatory response (e.g. p43). Thus, the non-translational functions of aaRSs are not anymore seen as exceptions, but instead it is assumed that multi-

functionality is a frequent process that links the translation machinery with cell signaling or other biological pathways (Park et al., 2008; Kim et al., 2011; Guo and Schimmel, 2013). Nowadays, aaRSs are considered as proteins from one of the most ancient family and it is assumed that they have been available for adaptation and recruitment to emerging cell signal pathways. Thus, it is considered that aaRSs are one of the earliest cytokines and cell signaling molecules (Park et al.; 2008) and since they have essential canonical function, the evolutionary retention enhances the ability to achieve and keep new functions (Guo and Schimmel, 2013).

5.3 General concepts of aaRSs to gain new functions (Figure 7)

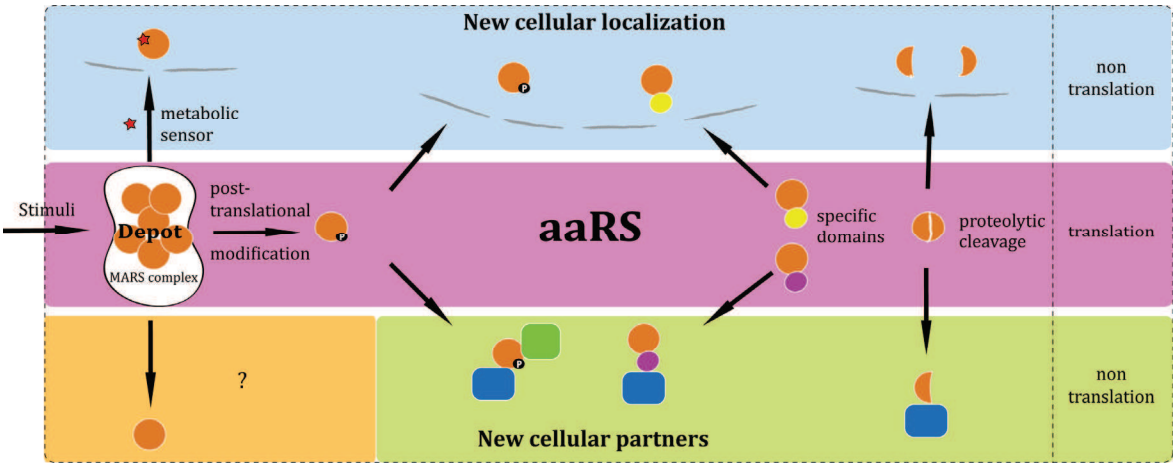


Figure 7: General concepts to of cytosolic aaRSs to gain new functions. Cytosolic aaRSs can gain non-translational function by a new cellular localization or new cellular partners. Triggers leading to non-translational functions are indicated. External stimuli can lead to release of mt-aaRS, organized in the MARS complex, by e.g. post translation modification (phosphorylation) or metabolites. Evolutionary acquired domains can harbor translocation signals or support the interaction with new partners. Proteolytic cleavage can lead to new active peptides.

5.3.1 Depot hypothesis

Cytosolic aaRSs fulfill non-translational function(s) in diverse ways. For instance, the MARS complex plays a central role in the regulation of aaRS alternative function. This

complex serves as a molecular reservoir ("depot hypothesis"), where anchored proteins are perform aminoacylation, while released proteins are free to perform alternate functions (Ray et al.; 2007). An example of such mechanism concerns the IFN-γ induced release of the Glu-ProRS from the MARS complex and its assembly to the GAIT complex, which mediates the translation inhibition of mRNA coding for VGEF-A (Sampath et al., 2004). Another example is the release of the LysRS from the MARS complex. Depending on some stimuli, the LysRS can be phosphorylated on either T52 or S207 residues (through MAPK pathway), whereby it is released from the MARS complex. The T52 phosphorylation leads to the translocation of LysRS to the plasma membrane where it inhibits the degradation of 67 kDa laminin receptor and promotes laminin-induced cell migration (Kim et al., 2012). In contrast, the S207 phosphorylation leads to nuclear import and association with the oncogenic transcription activator MITF (Yannay-Cohen et al., 2009).

5.3.2 Change of location and interaction with different partners

Cytosolic aaRSs, which do not belong to the MARS complex, may also have non-translational function(s). The change of the cellular compartment is a general mechanism for aaRSs to fulfill a new cellular function. For instance, the SerRS is translocated to the nucleus where it modulates the expression of the VEGF-A, a central factor in the vascular development (Fukui et al., 2009). Further studies discovered a robust nuclear localization signal (NLS) directing SerRS into the nucleus (Xu et al., 2012). As another example, the TrpRS is up-regulated by the IFN-γ whereby it accumulates in the nucleus to bridge the DNA-dependent proteins kinase (DNA-Pkcs) to the poly(ADP-ribose) polymerase 1 (PARP1) leading to the activation of p53 (Sajish et al., 2012). Other aaRS, such as TyrRS or TrpRS, can also be secreted from the cell and serve as cytokines (Wakasugi and Schimmel, 1999; Wakasugi et al., 2002).

5.3.3 Old domains, new function

AaRSs are modular proteins with typical catalytic and tRNA binding domains. Higher eukaryotic aaRSs acquired additional domains. Beside the classical function of the domains, they also play a role in the non-translational function of the aaRSs. The amino acid binding pocket from LeuRS also serves as an instrinsic leucine sensor being independent from $tRNA^{Leu}$. In the presence of leucine, LeuRS translocates into the lysosome and serves as an activator of the mTOR pathway (Han et al., 2012). Other metabolic sensors are LysRS, TrpRS or GlnRS and playing roles in MITF activation, anti-angiogenesis or anti-apoptosis, respectively (reviewed in Guo and Schimmel 2013).

5.3.4 New domains in aaRSs

Recent analyses indicate that the acquisition of new, non-translational functions of aaRSs through additional non-catalytic domains correlate with the increasing complexity of the organism. The appearance of new domains in aaRSs during the evolution marks the manifestation of new biological functions. Today, about 20 domains are identified in eukaryotic cytosolic aaRSs among which GST, WHEP or leucine zipper domains are responsible for new protein-protein interactions. For example, the WHEP domain in TrpRS of higher eukaryotes is responsible for the nuclear interaction of TrpRS with DNA-Pkcs and PARP1. In addition to these well-known domains, unfamiliar domains (named UNEs) were also found. These domains are unique for a family of aaRSs, have almost no sequences homologies with other proteins, and their functions are only partially understood (Guo et al., 2010). For example, the nucleus localization of the SerRS is triggered by the UNE-S domain (the letter S represent the aa of the corresponding aaRS in the single letter code), which includes a nuclear localization signal (Xu et al., 2012). In general, it turns out that the addition of these new domains has no obvious influence on the aminoacylation activity of the aaRSs.

5.3.5 Natural fragmentation as source for novel function

Another strategy to gain new functional proteins is the natural fragmentation of proteins by proteolytic cleavage, truncation or alternative splicing. In humans, the TyrRS is, under certain conditions, secreted and proteolytically cleaved. The cleavage generates two newly active cytokines, the mini-TyrRS and the endothelial monocyte-activating polypeptide II (EMAPII), with distinct functions in angiogenesis and immune activation. The removal of the N-terminal WHEP domain of TrpRS by proteolysis or alternative splicing generates an active angiostatic cytokine that interact with VE-cadherin and blocks blood vessel formation (Kise et al., 2004). In general, it seems that the newly generated peptides gain functions initially absent in the full-length proteins (e.g. Guo et al., 2010).

5.3.6 Post-translational modifications as on and off switches

Phosphorylation plays an import role in the activity of some aaRSs. LysRS has two phosphorylation sites regulating the switch from the translational to the non-translational activity. For example, in mast cells, the phosphorylation state of S207 regulates the involvement of LysRS in either translational or transcriptional activities. LysRS is translationally active if it is bound to the MARS complex. Upon antigen activation of the MAP kinase pathway, LysRS can be phosphorylated and change its conformation from a “closed form” to an “open form”. The “open form” is unable to bind the MARS complex and is translocated from the cytoplasm to the nucleus, where the “open form” regulates the transcription activator MITF (Ofir-Birin et al., 2013). In contrast, phosphorylation in Glu-ProRS are not only required for the release from the MARS complex, but also for the interaction of the Glu-ProRS with proteins from the GAIT complex (Arif et al., 2009). It is proposed that phosphorylation is a general principle that regulates the release of aaRSs from the MARS complex, which is critical to achieve non-translational functions (Ray et al., 2007, Arif et al., 2009).

Of note and for sake of simplicity, alternate functions of the three scaffold proteins contained within the MARS complex were not discussed but it is worth stating that they are involved in the regulation of development, metabolism, tumorigenesis, neuronal and inflammatory processes.

5.4 Non-translational functions and their link to diseases

Aberrant function or expressions of cytosolic aaRSs are correlated with different types of diseases such as autoimmune disorders, neurological disorders, and cancers (Park et al., 2008). Nowadays, pathology-related mutations of cytosolic aaRSs (e.g. GlyRS, Antonellis et al., 2003; or TyrRS, Jordanova et al., 2006) were mainly found in Charcot-Marie-Tooth diseases, a heterogeneous group of diseases affecting the peripheral nervous system (Patzkó and Shy, 2011). AaRSs are involved in these diseases not only through an impact on their canonical function but also on their non-translational function (Park et al., 2008; Guo et al., 2010; Guo and Schimmel 2013). Since aaRSs play multiple roles in pathways other than translation, it is expected that a dysfunction in aaRSs can lead to a wide range of aberrant cellular functions. A recent study systematically connects cytosolic aaRSs in a network model with more than 120 cancer-associated genes (Kim et al., 2011). It was speculated that aberrant expression and post-translation modifications play a significant role. For example, in some colon cancers an increased expression of the MetRS and the GlyRS was observed. Since, the GlyRS has a pro-apoptotic and tumor-growth suppressing activity, its overexpression might reflect an innate anti-tumor response (Park et al., 2012). Other examples, where non-translational functions are associated with diseases comprise the expression of autoantibodies against a subset of aaRSs (e.g. AlaRS, HisRS or ThrRS) in some autoimmune disorders such as diabetes type I, polymyositis or dermatomyositis. Reasons for the production of antibodies against the aaRS remain unclear, but it is speculated that some activities as chemoattractant might play a role (Guo and Schimmel 2013).

Objectives of the thesis

Since the complete annotation of mt-aaRSs in 2005 and the first discovery of pathology-related mutations in the nuclear genes coding for the human mt-aaRSs in 2007, our lab (and others) have started several biochemical, structural and biophysical characterization of some of the mt-aaRSs and their involvement in diseases. More and more peculiarities of the mitochondrial translation system are reported. It turns out that the etiology of mitochondrial diseases (correlated with mutations in mt-aaRSs) is quite diverse and sometimes cannot only be explained with the current knowledge on the mt-aaRS ('Missing-link' hypothesis).

Alternate functions have already been described for several aaRSs of cytosolic location, and were shown to be due to alternate cellular organization, alternate partnerships or localization. While the organization of human cytosolic aaRSs is well explored and their involvement into alternate functions clearly established for some of them, the properties of the human mt-aaRSs remain scarcely known. The purpose of this thesis was double tracked. On one site, it should reveal new peculiarities of mt-aspartyl-tRNA synthetase and integrate it into new functional networks (sub-mitochondrial localization and partnership). On the other site, this thesis should expand the view of the sub-mitochondrial organization (location and repartition) to the full set of mt-aaRSs and should ultimately shed light into the molecular mechanisms underlying some of the pathologies. The thesis follows three major objectives, which are presented in five chapters in the result section.

I. The mitochondrial translation machinery is of dual origin. The mt-tRNAs are encoded in the mitochondrial genome while the mt-aaRS are encoded in the nucleus one. The mitochondrial mt-tRNAs are structurally 'degenerated' while the mt-aaRSs are mostly conserved in their canonical structural properties if compared to the bacterial ancestors. Strong molecular adaptations of the mt-aaRS to the tRNAs have been hypothesized so that the partnership between both

macromolecules is maintained for accurate mitochondrial translation. Along this line, we observed decreased sequence conservation in one domain of one mt-aaRSs (mt-AspRS), which paired together with the occurrence of alternative splicing in same domain.

The goal of ***chapter I*** was to understand and explain the occurrence of an alternative spliced transcript for the mt-AspRS and to connect this observation with decreased sequence conservation.

II. Recent evidence indicates that macromolecular assemblies and dynamic cellular organization might be sources for proteins with auxiliary functions. In contrast, there is no solid knowledge about the sub-mitochondrial organization in mitochondria. It becomes thus necessary to study dynamic location (e.g. import from cytosol into mitochondria or sub-localization inside mitochondria) and to determine all potential interacting components of mt-aaRSs within the organelle. These objectives are analyzed and discussed in three separated chapters.

The goal of ***Chapter II*** was to establish the sub-mitochondrial organization of the mt-aaRSs (sub-mitochondrial location and repartition of the full set of mt-aaRSs).

The goal of ***Chapter III*** was to establish and validate a unique and robust tool to enable exploration of the impact of precursor sequences on the import process and ultimately to have access to naturally processed mitochondrial proteins.

The goal of ***Chapter IV*** was to identify possible interacting components of the mt-AspRS and to integrate it in functional interaction network.

III. The distance between the birth of mt-aaRS and its final role in ATP synthesis is very long, so that the gap between the comprehension of the molecular impact of mutations in macromolecules and a dysfunction of the respiratory chain is very large. Thus, links between the activity of a given aaRS to mitochondrial translation on one hand and ATP production on the other, involve a number of

issues that need to be further explored. New routes towards understanding of the mosaïcity of effects outside the frame of classical defects mitochondrial translation (e.g. reduced aminoacylation activity) should become opened along these lines.

The goal of ***chapter V*** was to initiate the *de novo* analysis of some discovered pathology-related mutations in tRNA and mt-aaRS genes. In addition, we want to re-analyze known and characterized mutants of mt-AspRS in regard to possible physic-chemical defects and their impact on their sub-mitochondrial organization.

Altogether, this thesis aims to contribute to a better understanding of mosaïcity of effects connected with an emerging list of mitochondrial diseases correlated with mt-aaRSs. The establishment of the sub-mitochondrial organization (sub-mitochondrial location and repartition of the full set of mt-aaRSs) and partnership is expected to reveal new peculiarities, alternate/novel function(s) of mt-aaRSs, their possible involvement into different cellular pathways, and ultimately should shed light into molecular mechanisms underlying some of some of the pathologies.

RESULTS

Chapter I: Bacterial insertion domain of mitochondrial AspRS

1 Introduction

Mitochondria are of endosymbiotic origin and thought to result from the engulfment of a bacterial like organism. The source of nuclear genes coding for mt-adressed aaRSs is however diverse, reflecting numerous post-endosymbiotic and/or lateral gene transfer events (Brindefalk et al., 2007). Among others, the human mt-AspRS originates from the bacterial domain of life. The following chapter integrates the mitochondrial AspRSs in an evolutionary scenario that suggests the reshaping of an ancestral structural domain (named "Bacterial Insertion Domain", present in aspRSs from bacteria to metazoan) in connection with the probable loss of function of this domain in the mitochondrial environment. Our findings combine biochemical and bioinformatics evidence, and connect extensive sequence divergence, observed only within opisthokont and protist mitochondrial sequences, with the occurrence of alternative splicing events concentrated on this domain. The main results are summarized and discussed in the article #1. They include: i) the unusual conservation pattern of the BID, leading to a partition between the bacterial and the viridiplantae mitochondrial sequences on the one hand, and the opisthokont and protist mitochondrial sequences on the other hand; ii) the discovery of alternatively spliced forms of mitochondrial AspRSs, impacting differently the bacterial insertion domain in human and mouse; iii) the ubiquitous presence of the alternative transcript in all tested human tissues and iv) the active translation of the alternative transcript, but the decreased protein stability. These results are in line with other characteristics (Bonnefond et al., 2005; Fender et al., 2006, 2012; Neuenfeldt et al., 2013), established for the human mt-AspRS that

indicate a functional relaxation of mt-AspRSs compared to bacterial and plant mt-AspRSs, despite their common ancestry.

In addition to the results presented in the article, two paragraphs describe additional experiments. Paragraph 3 describes the numerous attempts to detect a possible protein corresponding to the alternatively spliced human mitochondrial AspRS *in vivo* by a combination of 2D-Gel electrophoresis and mass spectrometric analyses. Paragraph 4 describes knock down experiments that were performed to decipher a possible functional implication of the spliced mt-AspRS.

2 Article #1

Released selective pressure on a structural domain gives new insights on the functional relaxation of mitochondrial aspartyl-tRNA synthetase

Hagen Schwenzer, Gert C. Scheper, Nathalie Zorn, Luc Moulinier, Agnès Gaudry, Emmanuelle Leize, Franck Martin, Catherine Florentz, Olivier Poch and Marie Sissler

(2013*) submitted*

Released selective pressure on a structural domain gives new insights on the functional relaxation of mitochondrial aspartyl-tRNA synthetase

Hagen Schwenzer[1], Gert C. Scheper[2#], Nathalie Zorn[3], Luc Moulinier[4], Agnès Gaudry[1], Emmanuelle Leize[3], Franck Martin[1], Catherine Florentz[1], Olivier Poch[4] and Marie Sissler[1*]

[1] Architecture et Réactivité de l'ARN, CNRS, Université de Strasbourg, IBMC - 15 rue René Descartes, F-67084 Strasbourg Cedex, France
[2] Department of Pediatrics and Child Neurology, VU University Medical Center, 1081 HV Amsterdam, The Netherlands.
[3] Laboratoire de Spectrométrie de Masse des Interactions et des Systèmes, Chimie de la Matière Complexe, 1 rue Blaise Pascal, F-67008 Strasbourg Cedex France
[4] Laboratoire de Bioinformatique et de Génomique Intégratives, IGBMC, 1 rue Laurent Fries BP-10142, F-67404 Illkirch Cedex, France
[#]Present address: Crucel Holland BV, Archimedesweg 4-6, 2333 CN Leiden, The Netherlands

* To whom correspondence should be addressed: Marie Sissler, IBMC - 15 rue René Descartes, 67084 Strasbourg Cedex, France, Phone: 33 (0)3 88 41 70 62; Fax: 33 (0)3 88 60 22 18; E-mail: M.Sissler@ibmc-cnrs.unistra.fr

Keywords: mitochondria, aminoacyl-tRNA synthetase, bioinformatics, translation, moleculare biology

Abbreviations: aaRS: aminoacyl-tRNA synthetase (specificity is indicated by the name of the amino acid abbreviated in a three-letter code, transferred to the cognate tRNA, e.g. AspRS stands for aspartyl-tRNA synthetase), MTS: mitochondrial targeting sequence, mt: mitochondrial

HIGHLIGHTS

- Mitochondrial AspRSs possess a bacterial insertion domain (BID) of unclear function
- Probable relaxation of the selective pressure on the BID in the eukaryotic branch
- Sequence divergences and alternative splicing are concentrated on BID of mt-AspRSs
- In line with discrepancies between mt and bacterial AspRSs despite common ancestry

ABSTRACT

Mammalian mitochondrial aminoacyl-tRNA synthetases are nuclear-encoded enzymes that are essential for mitochondrial protein synthesis. Due to an endosymbiotic origin of the mitochondria, many of them share structural domains with homologous bacterial enzymes of same specificity. This is also the case for human mitochondrial aspartyl-tRNA synthetase (AspRS) that shares the so-called bacterial insertion domain with bacterial homologs. The function of this domain in the mitochondrial proteins is unclear. Here, we show by bioinformatic analyses that the sequences coding for the bacterial insertion domain are less conserved in opisthokont and protist than in bacteria and viridiplantae. The divergence suggests a loss of evolutionary pressure on this domain for non-plant mitochondrial AspRSs. This discovery is further connected with the herein described occurrence of alternatively spliced transcripts of the mRNAs coding for some mammalian mitochondrial AspRSs. Interestingly, the spliced transcripts alternately lack one of the four exons that code for the bacterial insertion domain. Although we showed that the human alternative transcript is present in all tested tissues; co-exists with the full-length form, possesses 5'-and 3'-UTRs, a poly-A tail and is bound to polysomes, we were unable to detect the corresponding protein. The relaxed selective pressure combined with the occurrence of alternative splicing, involving a single structural sub-domain, favors the hypothesis of the loss of function of this domain for AspRSs of mitochondrial location. This evolutionary divergence is in line with other characteristics, established for the human mt-AspRS, that indicate a functional relaxation of non-viridiplantae mt-AspRSs when compared to bacterial and plant ones, despite their common ancestry.

INTRODUCTION

Aminoacyl-tRNA synthetases (aaRSs) are housekeeping enzymes involved in the essential process of protein biosynthesis, *i.e.* the translation of the genetic information from mRNA into proteins. In every cell and organelle, each of the 20 aaRSs esterifies specifically its corresponding tRNA(s) with the correct amino acid, which is then transferred to the growing peptide chain on the ribosome. AaRSs have been extensively explored during the past decades to unravel their structural, functional and evolutionary properties (reviewed in *e.g.* [1-4]). In human mitochondria, protein biosynthesis is dedicated to the production of 13 proteins, all sub-units of the respiratory chain complexes, which are the seats for energy production (reviewed in [5]). Human mitochondria possess a specific set of aaRSs, all encoded by the nuclear genome, synthesized within the cytosol and imported into the mitochondria thanks to a mitochondrial (mt) targeting sequence (MTS) [6]. Their genes differ from the ones encoding cytosolic-addressed aaRSs, with solely two exceptions where one single gene encoded both forms (GlyRSs and LysRSs) [7]. Despite the conventional view of the endosymbiotic origin of mitochondria [8], the source of nuclear gene for mt-addressed aaRSs is diverse, reflecting numerous post-endosymbiotic and/or lateral gene transfer events [9]. Nevertheless, many of the mt-aaRSs originate from the bacterial domain. This is also the case for human mt-aspartyl-tRNA synthetase (mt-AspRS) [7].

The human mt-AspRS shares 43% of identical residues, the same secondary

structure organization (including the bacterial-type insertion and C-terminal extension domains), and a same architecture with *Escherichia coli* AspRS, a representative bacterial homolog [10]. Despite the fact that the two enzymes are likely descendants from a common ancestor, numerous functional idiosyncrasies and discrepancies were reported for human mt-AspRS when compared to *E. coli* AspRS. These concern a reduced catalytic efficiency [6, 7], the requirement of a minimal set of determinants within cognate tRNA [11], a higher sensitivity to small substrate analogs [12], the capacity of cross aminoacylation of bacterial $tRNA^{Asp}$ [13], and an increased structural plasticity of the mitochondrial enzyme when compared to its bacterial homolog [10].

To further investigate evolutionary discrepancies, sequence analyses at the genomic and proteic levels were performed and revealed an extensive divergence in the bacterial insertion domain (BID), a domain shared by all bacterial-type AspRSs. In contrast to its strong conservation in bacteria, we uncovered a high sequence variability of the BID in almost all eukaryotes mitochondrial sequences (with the exception of green plants). In vertebrates, the BID is encoded by exons 11, 12, 13, and 14. Interestingly, we found an alternatively spliced transcript of the mRNA coding for the human mt-AspRS, which misses exon 13. Further characterization of this spliced transcript revealed that both full-length and exon13-deleted mRNAs co-exist in all tested tissues, are processed and handled by the polysomes, suggesting their active translations. Although we were unable to detect a protein corresponding to the spliced form, the large sequence diversity of the BID in almost all eukaryotes, together with the fact that we found alternatively spliced forms of mt-AspRS mRNA that lack other BID exons in other mammals, led us to speculate that there has been a relaxation of the selective pressure on the BID in the eukaryotic branch. It is possible that the appearance of mitochondrial AspRSs splice variants, defective in the BID, reflects the functional relaxation of the nuclear-encoded mt-aaRSs.

MATERIAL AND METHODS

Generation of the Multiple Sequence Alignment (MSA)

Sequences of the mt-AspRS proteins were examined in 81 eukaryotic organisms with: 27 Metazoa, 40 Fungi, 9 Archaeplastida (*Viridiplantae*), and 5 Protists. To obtain an objective evaluation of the sequence divergence, the bacterial homologues were retrieved from 104 organisms encompassing all bacterial subgroups and removing redundancy by counting only non-identical sequences. Finally, the phylogenetic analysis (see below) was performed using 9 sequences from archaeal and eukaryotic cytosolic AspRSs (*Cyanidioschyzon merolae*, *Pyrobaculum aerophilum*, *Aeropyrum pernix*, *Thermococcus kodakarensis*, *Methanothermobacter thermautotrophicus*, *Arabidopsis thaliana*, *Drosophila melanogaster*, *Saccharomyces cerevisiae*, *Homo sapiens*) as outgroup. The complete list of species is provided in the Supplementary Table I.

Initial BlastP searches [14] were conducted at the National Center for Biotechnology Information site (http://www.ncbi.nlm.nih.gov/BLAST) in the non-redundant protein database ($E \leq 0.001$) using *Homo sapiens*, *Saccharomyces cerevisiae* and *Escherichia coli* proteins as queries. When necessary, additional similarity searches were performed at the genomic level using TblastN to uncover potentially missing or mispredicted proteins. The likely homologous sequences detected by BLAST searches were aligned using PipeAlign [15] providing a final Multiple Sequence Alignment (MSA). The quality of this MSA was evaluated using the NorMD

(normalized Mean Distance) [16]. Finally, the manual verification and correction, paying attention to secondary structures based on the bacterial protein AspRS structures found in the PDB database (4ah6 for the human mitochondrial enzyme and 1l0w and 1c0a for *Thermus thermophilus* and *E. coli* enzymes, respectively), was performed using ORDALIE alignment viewer and editor (manuscript in preparation). Subfamilies were automatically defined using DPC (Density of Point Clustering) [17] and validated by human expertise on the basis of global conservation, domain organization and inter-subfamily residue conservation analysis. The complete alignment of the bacterial and mt-AspRS protein sequences is given in Supplementary Figure 1.

Sequence Analysis

To incorporate detailed analyses of the protein families at the structural, functional, and evolutionary levels, the protein present in the MSA were scanned using MACSIMS [18]. Sequence conservation within each family was estimated using the ORDALIE software to evaluate various conservation scores corresponding to the global and pairwise sequence identities and similarities between various sets of complete sequences including i) the 194 organisms; ii) the distinct subfamilies (viridiplantae, bacteria, fungi and metazoan plus protists). Finally, for the 194 organisms and for each subfamily, mean percent identity is measured in a sliding-window analysis (window length=50, step=1) along the length of the alignment.

The FamID defines the conservation within each subfamily and is calculated as the mean of all pairwise percent identities and the positions in the alignment with gaps within the subfamily are excluded from the calculation.

$$FamID = 2\frac{\sum_{1 \le i < j \le n} ID_{S_i,S_j}}{n(n-1)}$$

where n = total number of sequence tested, S_i and S_j the *i*th and *j*th sequence, and $ID_{Si,Sj}$ = pairwise percent identity between the *i*th and *j*th sequence, excluding gap regions.

Cell culture

HEK-293 (Human Embryonic Kidney) and HEPA1-6 (mouse hepatoma cells, gift from M. Frugier, Strasbourg) cells were maintained in DMEM (Dulbecco's Modified Eagle's Medium, Invitrogen) supplemented with 10% FCS (Fetal Calf Serum), 100 units/ml penicillin and 100 µg/ml streptomycin at 37°C, in a 10% CO_2 humidified incubator. BHK-21 (Baby Hamster Kidney) cells (ATCC#CRL-12072) were cultivated in GMEM (Glasgow Modification of Minimum Essential Medium, Invitrogen) supplemented with 10% FCS, 1.5 g/ml bacto-tryptose phosphate, 100 units/ml penicillin and 100 µg/ml streptomycin at 37°C.

PCR amplification of human and mouse mt-AspRS coding sequences

The human mt-AspRS coding sequence was amplified from a human cDNA library (obtained from HEK-293 cells) using Dynazyme EXT polymerase (Finnzymes). Primers for PCR were designed to amplify the mature enzymes, *i.e.* without of its theoretical N-terminal MTS [7] (see Supplementary Fig. 2). PCR products were directly used for cloning into pCR2.1® following the TA Cloning Kit protocol (Invitrogen). After transformation of One Shot TOP10 *E. coli* cells (Invitrogen), plasmid DNA was purified and sequenced [7].

Two micrograms of total RNA obtained from HEPA1-6 cells were reverse transcribed following the First Strand cDNA synthesis kit protocol (Fermentas). A fragment of the mouse mt-AspRS sequence (spanning from exon 6 to exon 15) was amplified using a proofreading DNA Polymerase. PCR products were purified and cloned into pCR blunt II vector following the instructions of the Zero Blunt® PCR Cloning Kit (Invitrogen),

plasmid DNA was purified and sequenced.

RT-PCR of mt-AspRS in different tissues

cDNA synthesis, DNA amplification and quantitative PCR were carried out as described before [19]. Human total RNAs from 20 tissues were purchased from Ambion (FirstChoice® Human Total RNA Survey Panel) or purified from HEK cells as described before [20]. Primer sequences and target sites were as described in Supplementary Fig. 2. qPCR values obtained for each tissue were corrected for GAPDH mRNA levels. Relative expression level of Δexon13 variant was calculated using the ΔΔCt method.

Recombinant protein expression in mammalian cell

Recombinant proteins (mt-AspRS or mt-AspRS-Δexon13) were expressed in BHK cells using the transient MVA-T7 (Modified Vaccinia Virus Ankara) transfection-infection approach as described in [21]. Briefly, BHK-21 cells were grown for 24 h on 10 cm culture dishes, infected with a mixture containing MVA-T7 and 20 µg of plasmid DNA. 30 min after infection, cells were induced with 1 mM IPTG. After 24-48 h, cells were harvested by scraping and centrifuging to obtain cell pellets. Crude protein extracts were recovered from the pellets using Pierce RIPA buffer (Rockford, IL, USA) containing 1x protease arrest (GBiosience, USA) following the manufactures instructions.

Sucrose-Gradient Analysis

HEK-293 cells were grown on 30x150 mm dishes. At 75% confluence, cells were harvested, washed three times in phosphate-buffered saline (PBS) and resuspended in lysis buffer (10 mM Hepes pH 7.5, 10 mM KAc, 0.5 mM $MgAc_2$, 5 mM DTT) and lysed by 20 strokes with a syringe. After centrifugation at 13000g, the cleared supernatant was concentrated in a centrifugal filter device (cutoff 10 kDa). The poly-ribosomes were separated from 80S ribosomes and free RNA on a 7%–47% linear sucrose gradient (25 mM Tris-HCl pH 7.4, 50 mM KCl, 5 mM $MgCl_2$, 1 mM DTE, 1 mM AEBSF) in a SW41 rotor (Beckman coulter) for 2.5 hr at 37 000 rpm at 4°C. Gradients were fractionated, phenol/chloroform extraction of RNA was performed [20] and two micrograms of total RNA were used for reverse transcription following the First Strand cDNA synthesis kit protocol (Fermentas). Mt-aaRS sequences were amplified as described above. PCR products were purified, cloned into pCR blunt II vector following the instructions of the Zero Blunt® PCR Cloning Kit (Invitrogen), and plasmids were sequenced.

RESULTS

1. Different selective pressures within bacterial insertion domains of AspRSs

AspRSs evolution has been extensively studied [22], emphasizing that the nuclear-encoded protein of mitochondrial location, present in all eukaryotes, has a bacterial origin [7, 9]. The trypanosomatids are an exception, where the two AspRSs are of eukaryotic type [23]. To determine the conservation patterns of mt-AspRS subfamily proteins in more detail, we searched for the presence of mt-AspRS and bacterial AspRS genes across an extensive range of organisms (see Supplementary Table I for a complete list) and performed sequence comparison on the basis of a structurally and manually refined sequence alignment of the 194 sequences (Supplementary Figure 1). As previously shown [7, 9, 22], this analysis confirmed that, from bacteria to eukaryotes, all bacterial-type AspRSs share a common domain architecture (Figure 1A) with an N-terminal anticodon-binding domain, a hinge region, a catalytic domain embedding the three class II signature motifs and the BID, and finally, a C-terminal bacterial extension. Mitochondrial proteins have an additional N-terminal MTS.

As exemplified by the phylogenetic analysis (data not shown), the bacterial

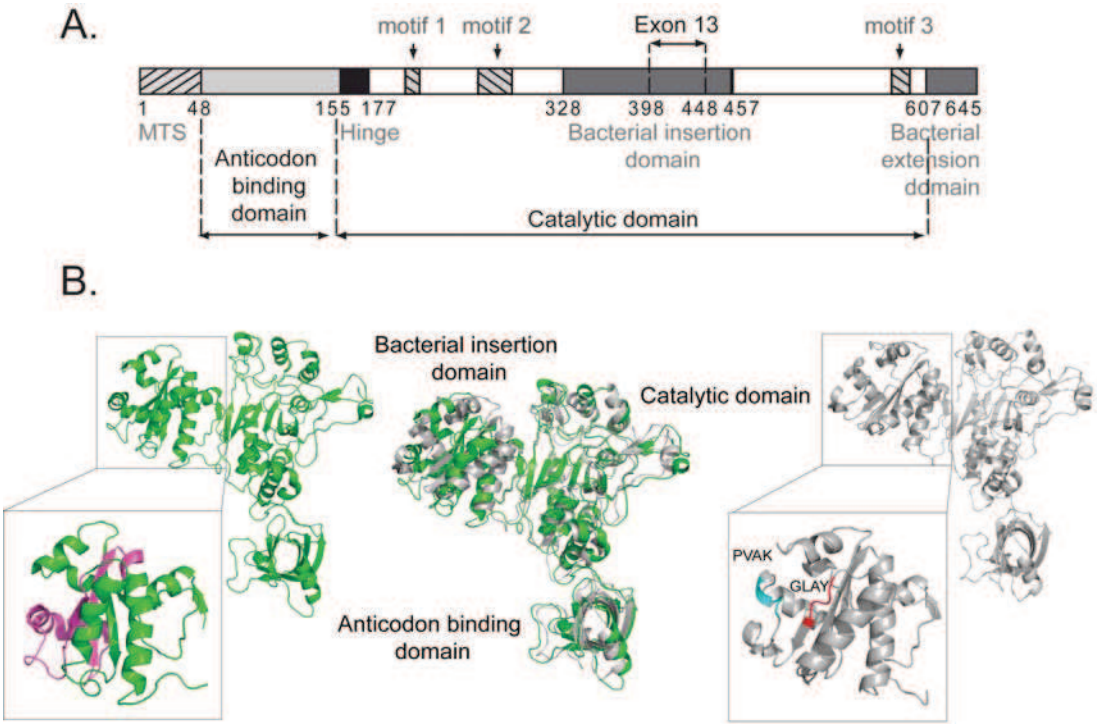


Figure 1 : Exon13-encoded peptide in mt-AspRS. A. Modular organization of full-length human mt-AspRS [7]. MTS stands for mitochondrial targeting sequence. Motifs 1, 2, and 3 constitute the catalytic core of the synthetase. **B.** Crystallographic structure of the human mt-AspRS (green, on the left) (PDB: 4AH6) [10]. The bowed part shows the BID, where the location of the exon13-encoded peptide is highlighted in magenta. Grey structure (right): the crystal structure of the *E. coli* AspRS (1EQR) [39], including a more detailed view of the BID where the highly conserved "G(L,I)Φ(Y,W,Φ)" and "PVAK" motifs in bacteria and viridiplantae, are colored in red and cyan, respectively. In the middle: superposition of both structures, emphasizing the overall fold conservation. Major structural domains are indicated. For sake of simplicity, only one monomer of AspRS (naturally found in a dimeric state) is shown in this figure. Structures have been drawn using PyMol (http://www.pymol.org)

AspRS and the viridiplantae mt-AspRS are closely related while the remaining mt-AspRS sequences are more divergent and dispersed on separate branches corresponding to opisthokonts (fungal and metazoan) and protists. The stronger conservation of the bacterial and plant enzymes is further confirmed by the amino acid percentages calculated for each pair of sequences, which range between 39.5 and 93.9 % (mean 51.5%) while the percentages observed for the other eukaryotic enzymes range from 25.1% up to 85.4% (mean 36.3%). To further investigate the conservation patterns present in the bacterial and mitochondrial AspRS sequences, mean percent identities were calculated using a sliding-window of 50 residues on the sequences from all organisms or from distinct subsets of organisms corresponding to the viridiplantae, the bacteria, the fungi and the metazoan plus protist (Figure 2). This analysis emphasizes that sequence conservation is not uniformly distributed but rather differentially scattered between the distinct functional domains. It also shows that the differential conservation pattern that distinguishes bacterial and viridiplantae from the fungi, metazoan and protists is mainly linked to the highly divergent BID in the latter subgroups. Indeed, the conservation scores obtained for the bacterial and viridiplantae subgroups are globally similar along the distinct domains. Conversely, when calculated in either the fungi or the protists and metazoan subgroups, the scores fall dramatically for the BID, emphasizing a weaker selective pressure at work in the opisthokont and

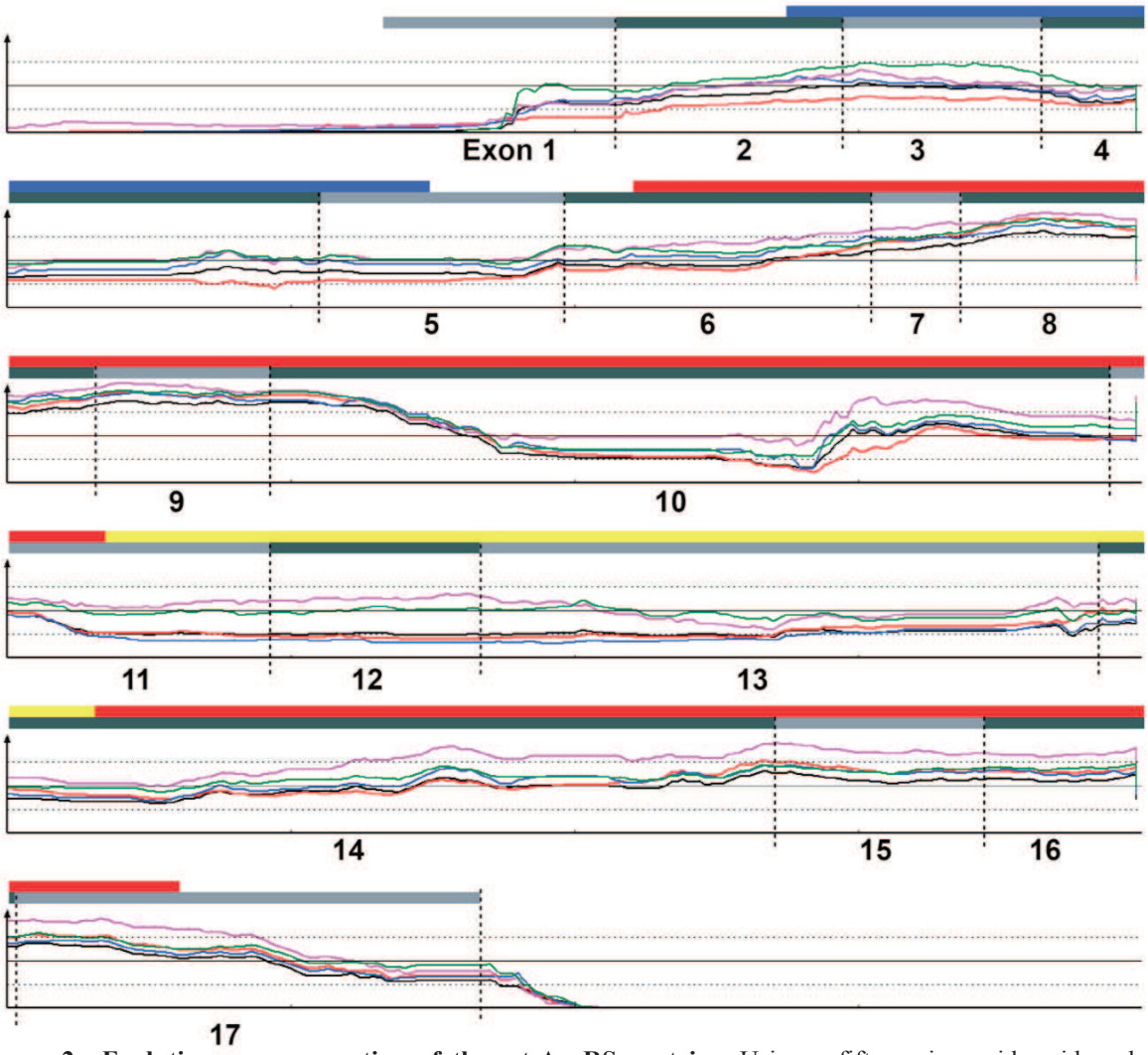


Figure 2 : Evolutionary conservation of the mt-AspRS proteins. Using a fifty amino acid residues long sliding window the average sequence identity in the mt-AspRS proteins is visualized. Functional domains are indicated by colored bars above the conservation diagram: blue bar: N-terminal anticodon-binding domain; red bar: catalytic domain; yellow bar: bacterial insertion domain. Mean conservation is indicated by dotted black lines corresponding from top to bottom to 75%, 50% and 25% mean identity. The colored lines represent the following: black: all organisms; red: fungi; blue: metazoan plus protists; magenta: viridiplantae and green: bacteria. Drops in the four conservation diagrams in the N-terminal and C-terminal regions and within the conserved regions are due to large extensions/insertions in some sequences. Numbers below the conservation diagram schematically illustrate the location of the exons in the mRNA corresponding to the displayed amino acid sequence (based on human mt-AspRS sequence); dashed lines and light/dark grey boxes indicate exon limits.

protist sequences. Next, we focused on the BID and calculated the mean amino acid identity percentages of this domain based on each pair of sequences belonging to one of the organisms' subgroups described above. The outcome emphasizes high conservation scores obtained for the bacterial (40% identity ± 9.8 standard deviation) and viridiplantae (43.2% identity ± 10.5 standard deviation) sequences in comparison to the values for the fungal (24.3% identity ± 13.6 standard deviation) and metazoan and protist (19.3% identity ± 8.4 standard deviation) sequences.

Finally, we searched for specific residues or motifs that may underlie the observed differential conservation behavior. This allowed us to uncover two quasi-invariant motifs "G(L,I)Φ(Y,W,Φ)" (where Φ stands for hydrophobic residues) located between residues 348 – 351 (using the *E. coli* numbering) and "PVAK" (aa 367 – 370

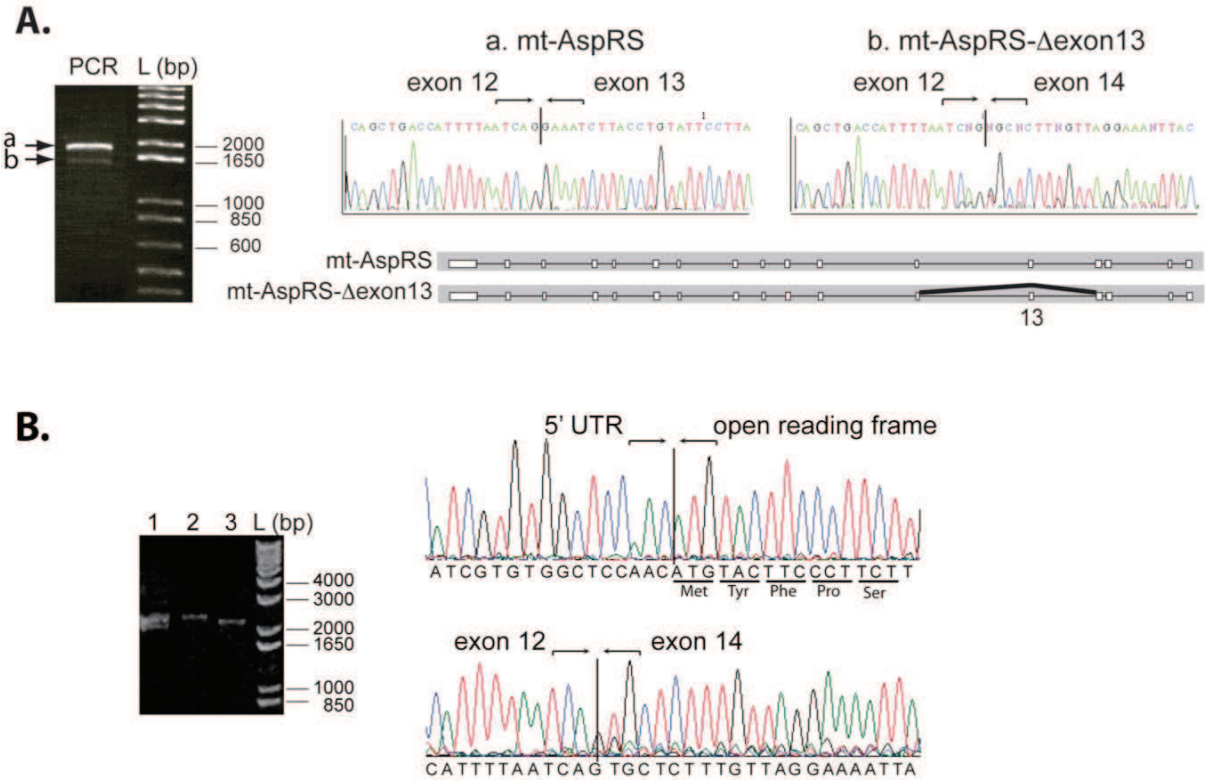


Figure 3 Two mRNAs for human mt-AspRS. A. Two PCR products of mt-AspRS coding sequences were amplified from a human cDNA library (arrows). Sequencing of PCR amplified products identified the full-length mt-AspRS sequence (a) and a sequence lacking 153 nucleotides corresponding to the deletion of the symmetric exon 13 (b). A schematic representation of the corresponding pre-mRNAs organization is given below the chromatograms (exons: white boxes; introns: lines). **B.** Detection of a full-length mt-AspRS-Δexon13 mRNA. PCR fragments containing the complete open reading frame were generated with primers hybridizing within the 5'- and 3'-UTRs, resulting in two bands (lane 1). These two bands were cut separately and purified (lanes 2 and 3). Sequence analysis of the shorter of the two fragments focused on the region containing the start codon (upper part) and the region spanning exon 13 (lower part), showing that the alternatively spliced mRNA contains an open reading frame that contains the normal initiation codon and encode the N-terminal MTS. L stands for molecular weight ladder. All primer sequences are detailed in Supplementary Figure 2.

following *E. coli* numbering) which are common to all bacterial and viridiplantae sequences and strictly absent in the other sequences (Figure 1B and Supplementary Figure 1). The extensive sequence divergence between the fungal, metazoan and protists sequences imply that no specific residue or motif common to the three groups could be identified.

2. Co-existence of two mRNAs for human and mouse mitochondrial aspartyl-tRNA synthetases

In addition to the observed strong sequence divergence, biochemical investigations reveal further peculiarities of the BID. This domain is encoded in vertebrates by exon 12 and exon 13, and parts of exons 11 and 14. When amplifying by PCR the coding sequence of the human mt-AspRS out of a human cDNA library, one shorter DNA fragment was amplified in addition to the PCR product of expected size (Figure 3A). Sequencing of the two PCR products identified a gap of 153 nucleotides (in the smaller amplicon), corresponding to the deletion of exon 13 in mt-AspRS mRNA sequence. This exon is qualified symmetric since its length is an exact multiple of three. Thus, its deletion does not affect the downstream reading frame, but shortens the translated sequence by 51 amino acids. Subsequent amplification of the full-length mRNA with primers specific to 5'- and 3'-UTRs (Supplementary Figure 2) and sequencing, showed the presence of a MTS also for the alternative splice transcript mt-AspRS-

Δexon13 (Figure 3B), suggesting a mitochondrial location of the corresponding protein, if expressed and stable.

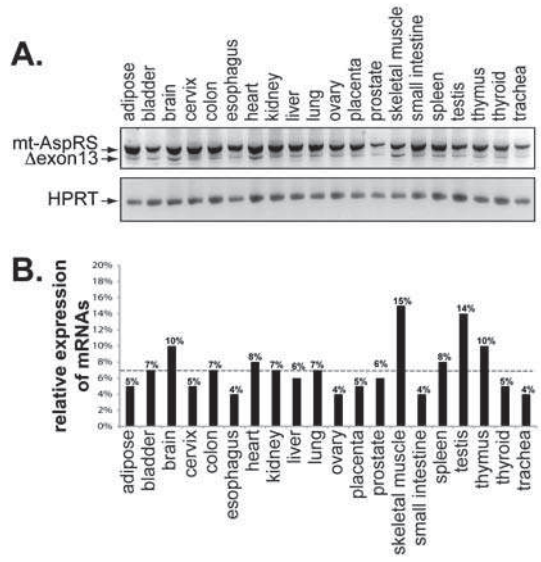


Figure 4 The Δexon13 splice-variant is expressed in different human tissues. **A.** PCR amplification (upper panel) of full-length and mt-AspRS-Δexon13 mRNAs using total cDNA from 20 different human tissues as templates, with oligonucleotides hybridizing in neighboring exon 11 (forward) and 3'-UTR (reverse) (see Supplementary Figure 2). Control PCR reactions with primers for hypoxanthine-guanine phosphoribosyl-transferase (HPRT) were included (lower panel). **B.** Relative expression level of Δexon13 variant, calculated using the ΔΔCt method. qPCR experiments were performed with the set of primers given in Supplementary Figure 2. One of these primers spans the exon 12-exon 14 boundary and recognizes only the mRNA missing exon 13. qPCR values obtained for each tissue were corrected for GAPDH mRNA levels. The dotted line emphasizes the mean calculated value.

The region of human mt-AspRS mRNA spanning exon 13 was amplified from cDNAs derived from total mRNAs extracted from 20 human tissues. As seen in figure 4A, two PCR products were retrieved from all investigated tissues. Similar results were obtained in several immortalized lymphoblast cell lines (not shown). Sequence analysis confirmed the presence or the absence of exon 13 respectively in the DNA fragments of high and low molecular weight. The relative abundance of full-length and the Δexon13-variant mRNAs were measured by quantitative PCR in each cDNA sample (Figure 4B). Two primer sets, one that detects both the full-length and the Δexon13-variant mRNAs and one specific to the splice-variant spanning the exon 12/exon 14 boundary, were used (Supplementary Figure 2). The shorter fragment was present in all tested tissues in different ratios (as compared to the longer fragment) ranking from 4% to 15%, with the highest level in skeletal muscle (15%) and the lowest level in esophagus, ovary, small intestine and trachea (4%). This demonstrates the wide distribution of mt-AspRS-Δexon13 mRNA in human tissues.

Expression data searches using BLAST identified two organisms (*Pongo abelii* NM_0011332290.1 and *Mus Musculus* BC057943.1) with alternative splice variants of mt-AspRS mRNA that lack exon 12. The presence of both the full-length and Δexon12 mt-AspRS mRNAs was experimentally validated using mouse cells. The mt-AspRS sequence was amplified from a cDNA bank obtained from HEPA1-6 (*Mus musculus*). Cloning of amplicons and sequencing confirmed the existence of an additional transcript that lacked 63 nt corresponding exactly to exon 12.

3. Human mt-AspRS-Δexon13 mRNA is polyadenylated and handled by polysomes

Polyadenylated total mRNAs, extracted from HEK-293 cells, were reverse transcribed using oligo-dT primers. Subsequent amplification with mt-AspRS gene-specific primers provided the two expected DNA fragments. Sequence analysis demonstrated that the high and low molecular weight DNAs correspond respectively to the polyadenylated full-length and Δexon13-variant transcripts (data not shown).

Next, to investigate whether full-length and Δexon13 mRNAs are both handled by the translation machinery and

expressed, all polysome-bound mRNAs were retrieved from polyribosomal fractions purified by sucrose sedimentation of crude cellular extract from HEK293 cells (Supplementary Figure 3). The two mRNAs were detected by RT-PCR amplification, and their sequences were confirmed after retrieval and cloning of the amplicons.

4. Absence of detection of the protein corresponding to human mt-AspRS-Δexon13

To characterize structural and functional properties of human mt-AspRS-Δexon13 *in vitro*, we attempted to clone and express the protein according to the procedure previously used for recombinant full-length mature mt-AspRS [7]. This approach was however unsuccessful when applied to mt-AspRS-Δexon13 despite a large number of alternate cloning or expression strategies. These strategies included fusions to the maltose binding protein (MBP) [24], and/or to "peri" [24] or "ompA" [25] secretion signals allowing targeting of the protein to the periplasm. Multiple bacterial strains were tested, including strains that expressed chaperone proteins [26]. In addition, a variety of different cell growth conditions were investigated (see Supplementary Method 1). Only one condition led to low yield of soluble protein that was however contaminated with non-reversely bound chaperones. This recombinant expressed protein had very low aminoacylation activity that was estimated to be 700-fold lower than the one of full-length mt-AspRS (data not shown).

Subsequently, *in cellulo* expression of mt-AspRS-Δexon13 was attempted. Plasmids containing either the full-length sequence of mt-AspRS or the Δexon13 variant, with an N-terminal flag-tag (DYKDDDDK), were introduced into mammalian cells (BHK21, baby hamster kidney cells) following the procedure described in [21]. Detection of flag-tagged proteins in crude cell extracts was performed by western blot using a monoclonal antibody specific for the flag peptide. While a flag-tagged protein was detected in cells expressing the full-length mt-AspRS, expression of mt-AspRS-Δexon13 was undetectable (not shown), suggesting either non-expression or non-stability of the splice variant.

Another attempt consisted on the *in vivo* detection of mt-AspRS-Δexon13 by mass spectrometry (Supplementary Method 2). The peptide N_{389}FAADHFNQ/CSLLGK$_{403}$ corresponds to the boundary exon 12/exon 14 -encoded regions, is experimentally generated by trypsin digestion (as verified on recombinant protein, not shown), and is specific to mt-AspRS-Δexon13 (not found in any other human protein as verified when screening human sequences database). This specific peptide was searched for by nanoLC-MS/MS on crude protein extracts, extracted from mitochondria (~400 mg of wet mitochondrial pellet) purified out of ~2.10^8 human cells (either HEK293 or myoblasts), and separated by either 1-D or 2-D gel electrophoresis. The detection limit of nanoLC-MS/MS experiment was estimated to be 32 ng of injected protein (mt-AspRS) by a sensitivity test. In this case, mt-AspRS was identified by four specific peptides (sequence coverage 9%). In-gel tryptic digestion of recombinant mt-AspRS-Δexon13 showed 68,90% sequence coverage including the specific 389-403 peptide. Western blot detections on a standard range of known concentrations of mt-AspRS were performed (not shown) and allowed to estimate an approximate amount of 8 µg of mt-AspRSs per 2.10^8 cells. The estimated expression level of mt-AspRS-Δexon13 (based on mRNAs ratios) is 5-10% as compared to the full-length sequence leading to a theoretical amount of ~400 to ~800 ng of protein produced from the spliced variant, a value 4 to 8 times above the theoretical detection level. However, despite good theoretical experimental conditions, the exon12/exon14 boundary peptide could not be detected (while peptides corresponding

to full-length mt-AspRS were retrieved), suggesting that additional factors (such as *e.g.* low stability of the protein) may limit the detection of mt-AspRS-Δexon13 *in vivo*. To exclude an unexpected sub cellular localization, similar experiments were also performed with crude cellular extracts, but without success.

DISCUSSION

The present study demonstrates divergent evolutionary characteristics of the bacterial insertion domains of AspRSs, leading to an unusual partition between the bacterial and the viridiplantae mitochondrial sequences on the one hand, and the opisthokont and protists mitochondrial sequences on the other hand. This unusual evolutionary picture is reflected by a combination of extensive sequence divergences and the appearance of mitochondrial AspRS splice variants alternatively defective in the BID in some mammals. Concerning sequence divergences, our calculations of mean amino acid identity percentages demonstrate that solely the BID is divergent and distinguishes the opisthokont and protists sequences from the bacterial and the viridiplantae ones. In addition, the two quasi-invariant motifs found in all bacterial and viridiplantae sequences are strictly absent in the other sequences. Altogether, these data firstly demonstrate that mitochondrial sequences display two distinct scenarios, emphasizing that this evolutionary discrepancy did not arise from the ancestor of mt-AspRSs. The data also demonstrate a specific and strong selective pressure that is exerted on the insertion domain in bacterial and viridiplantae AspRSs, suggesting potential important catalytic/functional mechanisms and/or interactions with partners. Conversely, the release of selective pressure within the BID in mt-AspRSs suggests an absence of function for this domain for non-viridaeplantae AspRSs of mitochondrial location, implying that divergent enzymes may emerge in these organisms.

The appearance of mitochondrial AspRS splice variants alternatively defective in the BID possibly further reflects the functional relaxation of the nuclear-encoded mt-AspRSs. We show, in human, that the amplicon corresponds to an mRNA coding for a variant missing the symmetric exon 13 that is present in all tested tissues together with the wild-type mRNA. This spliced mRNA possesses 5'- and 3'-UTRs, a polyA-tail and is bound to polysomes, suggesting that this mRNA is actively translated. The corresponding protein, however, turned out to be nearly impossible to express *in vitro* and was not detectable *in cellulo* or *in vivo*, strongly suggesting a decreased stability. Exon 13-encoded peptide is actually situated in the BID (Figure 1). The recently published crystallographic structure of the human mt-AspRS [10] indicates that excision of the exon 13-encoded peptide occurs at extremities of well-defined structural domains (Figure 1) and the start and the end of the peptide are adjacent in the 3D structure. This excision removes, however, half of a ß-sheet and leaves one of the major ß-strand alone, disrupting inter-strands hydrogen bounding, and thus disrupting inter-strands stability, which could result in a highly unstable protein.

The BID is alternatively named GAD in the Pfam annotation, according to its presence in both some **G**lutamyl-tRNA **A**midotransferase and bacterial-type AspRSs (**D**RS). Exploration of topology conservation leads to some hints towards a possible role (reviewed in [22]). Composed by more than 130 residues, the BID in *T. thermophilus* or in *E. coli* AspRSs is made up of a four-stranded ß-sheet surrounded by two anti-parallel α-helices on the side facing the active-site core, and smaller helices on the side facing the solvent [27, 28]. A similar fold has been observed in other proteins, e.g. glutamine synthase [29], nucleoside diphosphate kinase [30], and histidine-containing phosphocarrier protein [27], but no functional link with bacterial AspRSs could

be made so far. Of note, it has been demonstrated that the BID from *E. coli* AspRS interacts indirectly with the tRNA acceptor stem *via* a network of hydrogen bounds built up by a shell of more than 20 water molecules [28]. Supporting a functional role for this domain in *E. coli* AspRS, the mutation Arg383 to Cys383 was shown to modulate tRNA binding [31]. In the human mt-AspRS, the overall fold of the insertion domain remains unchanged, it however underwent a rigid-body counter-clockwise rotation of 26°, rendering unlikely a contribution for tRNA binding [10].

In vertebrates, the BID is encoded by part of exon 11, exon 12, exon 13, and part of exon 14. All these exons are qualified symmetric since their length is an exact multiple of three, and their removal will preserve the reading frame [32]. Of note, this is not the case for all mt-AspRS encoding exons. Data mining within expressed sequence tag databases revealed that alternative forms of mt-AspRS mRNAs exist in other mammals. Forms missing either exon 12 (in *Pongo abelii* and in *Mus musculus*) or exon 13 (in human) co-exist with full length mRNA. This observation has been validated experimentally in human and mouse cells, which indicates the absence of a specific and defined targeted-exon for splicing, but instead reveals the random possibility for the removal of either exon 12 or exon 13 in mammalian species. The absence of conservation of the splice variant in the human and mouse lineages likely excludes the possibility that the new isoform would have been retained so that to generate a new protein, with a new function. Also and to our knowledge, solely the BID is subjected to exon skipping under normal conditions. One additional event of exon skipping (deletion of exon 3) has been reported, but as a consequence from a genetically inherited mutation involved in a human leukoencephalopathy [33].

The combination of loss of selective pressure with the occurrence of exon skipping is in line with the hypothesis of Modrek and Lee [34], who suggested that new molecular isoforms are acquired through several steps, combining over time e.g. alternative splicing with mutational events. Such a hypothesis follows the accepted proposal that alternative splicing serves as a testing ground for molecular evolution [35, 36]. This proposal has been further enriched by the introduction of the stochastic noise notion in alternative splicing mechanisms, i.e. the occurrence of alternative isoforms with no apparent functional roles [37]. This "noise" may play a critical role by creating a landscape of opportunities allowing trial/error approach for the evolution of the gene [34, 35, 37]. Various processes might be at work with first the occurrence of a "low fitness" form, which keeps the reading frame and thus escapes the nonsense-mediated mRNA decay. The "low fitness" form is likely inert and well tolerated by the cell since it represents only a few of the transcripts and since the original transcript form would still accomplish its function. This "low fitness" form is thus free to evolve, without altering the activity of the original isoform, until reaching "high fitness" [35, 38]. The final fate of the isoform depends on the benefit it will bring to the cell [34]. Accordingly, the Δexon13 mRNA version of mt-AspRS-encoding mRNA in human (or Δexon12 mRNA in mouse) would be such "low fitness" form(s) and might be considered as a "snapshot" of the ongoing evolution. The overall tissue distribution of mt-AspRS-Δexon13 excludes the random occurrence of an isoform generated from splicing errors but instead suggests a mechanism fixed by evolution.

Outlook

The revealing of extensive sequence divergence combined with the occurrence of alternative splicing concentrated on a single sub-domain of a protein raises numerous questions. What is the function of the BID of AspRS and why is this function specifically conserved in viridiplantae mitochondria? More importantly, do the

relaxed evolutionary constraints illustrate some specific adaptation to the mitochondrial environment of non-viridiplantae organisms? This would be in line with numerous functional and structural discrepancies between human mt-AspRS and *E. coli* AspRS, despite the fact that they are both descendants from a common ancestor. Such discrepancies are demonstrated by i) the reduced catalytic efficiency of the mitochondrial enzyme [7]; ii) the loss of a primordial determinant within its cognate tRNA, which is compensated by the replacement of an highly conserved aspartic acid residue by a glycine in the mitochondrial enzyme [11]; and iii) an increased intrinsic structural plasticity of the mitochondrial enzyme when compared to its bacterial homolog [10]. All those examples illustrate the relaxation of the mitochondrial protein, which likely represent evolutionary mechanisms for adaptation of nuclear-encoded proteins to degenerated mitochondrial-encoded RNAs [10]. The present study may enlarge this list, with the hypothesis of an ongoing decline of the bacterial insertion domain, a domain that is already missing in AspRS of cytosolic location in eukaryotes, or from archaea.

ACKNOWLEDGMENTS

E. Westhof is acknowledged for critical reading of the manuscript. We thank L. Echevarria for her contribution; A. Fender for help in the initial cloning step; K. de Groot for technical assistance; G. Eriani, M. Frugier, and R. Smyth for constructive discussions; A. Lombès, M. Frugier, A. Lescure, L. Maréchal-Drouard and M. Mörl for gift of material; and R. Ripp for generating html link for supplementary Figure 1. L. Kuhn and P. Hammann are acknowledged for 2D gel electrophoresis and contribution to MS analyses (dilution tests). This work was supported by Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg (UdS), Association Française contre les Myopathies (AFM), ANR MITOMOT [ANR-09-BLAN-0091-01/03] for financial support and MYOSIX for gift of myoblast cells. This work was partially supported by the French National Program “Investissement d’Avenir” (Labex MitCross), administered by the "Agence National de la Recherche", and referenced [ANR-10-IDEX-002-02]. H.S was supported by Région Alsace, Université de Strasbourg, AFM, and Fondation des Treilles.

COMPETING INTERESTS

Dr. Scheper is an employee of Crucell Holland BV. There are no patents, products in development or marketed products to declare.

REFERENCES

[1] Cusack S., Aminoacyl-tRNA synthetases, Curr. Opin. Struct. Biol. 7 (1997) 881-889.

[2] Giegé R., Sissler M., Florentz C., Universal rules and idiosyncratic features in tRNA identity, Nucleic Acids Res. 26 (1998) 5017-5035.

[3] Ibba M., Francklyn C., Cusack S., The aminoacyl-tRNA synthetases., Landes Biosciences, Georgetown, TX, (2005).

[4] Guo M., Yang X.L., Schimmel P., New functions of aminoacyl-tRNA synthetases beyond translation., Nat Rev Mol Cell Biol 11 (2010) 668-674.

[5] Florentz C., Sohm B., Tryoen-Tóth P., Pütz J., Sissler M., Human mitochondrial tRNAs in health and disease., Cell. Mol. Life Sci. 60 (2003) 1356-1375.

[6] Sissler M., Pütz J., Fasiolo F., Florentz C., Mitochondrial aminoacyl-tRNA synthetases., in: Ibba M., Francklyn C. and Cusack S. (Eds.), Aminoacyl-tRNA synthetases, Landes Biosciences, Georgetown, TX, (2005), pp. 271-284.

[7] Bonnefond L., Fender A., Rudinger-Thirion J., Giegé R., Florentz C., Sissler M., Toward the full set of

human mitochondrial aminoacyl-tRNA synthetases: characterization of AspRS and TyrRS., Biochemistry 44 (2005) 4805-4816.

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

[9] Brindefalk B., Viklund J., Larsson D., Thollesson M., Andersson S.G., Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases., Mol. Biol. Evol. 24 (2007) 743-756.

[10] Neuenfeldt A., Lorber B., Ennifar E., Gaudry A., Sauter C., Sissler M., Florentz C., Thermodynamic properties distinguish human mitochondrial aspartyl-tRNA synthetase from bacterial homolog with same 3D architecture, Nucleic Acids Res. 41 (2013) 2698-2708.

[11] Fender A., Sauter C., Messmer M., Pütz J., Giegé R., Florentz C., Sissler M., Loss of a primordial identity element for a mammalian mitochondrial aminoacylation system, J. Biol. Chem. 281 (2006) 15980-15986.

[12] Messmer M., Blais S.P., Balg C., Chênevert R., Grenier L., Lagüe P., Sauter C., Sissler M., Giegé R., Lapointe J., Florentz C., Peculiar inhibition of human mitochondrial aspartyl-tRNA synthetase by adenylate analogs., Biochimie 91 (2009) 596-603.

[13] Fender A., Gaudry A., Jühling F., Sissler M., Florentz C., Adaptation of aminoacylation rules to mammalian mitochondria, Biochimie 94 (2012) 1090-1097.

[14] Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs., Nucleic Acids Res. 25 (1997) 3389-3402.

[15] Plewniak F., Bianchetti L., Brelivet Y., Carles A., Chalmel F., Lecompte O., Mochel T., Moulinier L., Muller A., Muller J., Prigent V., Ripp R., Thierry J.C., Thompson J.D., Wicker N., Poch O., PipeAlign: A new toolkit for protein family analysis., Nucleic Acids Res. 31 (2003) 3829-3832.

[16] Thompson J.D., Plewniak F., Ripp R., Thierry J.C., Poch O., Towards a reliable objective function for multiple sequence alignments., J Mol Biol. 314 (2001) 937-951.

[17] Wicker N., Dembele D., Raffelsberger W., Poch O., Density of points clustering, application to transcriptomic data analysis., Nucleic Acids Res. 30 (2002) 3992-4000.

[18] Thompson J.D., Muller A., Waterhouse A., Procter J., Barton G.J., Plewniak F., Poch O., MACSIMS: multiple alignment of complete sequences information management system., BMC Bioinformatics 7 (2006) 318.

[19] van Kollenburg B., van Dijk J., Garbern J., Thomas A.A., Scheper G.C., Powers J.M., van der Knaap M.S., Glia-specific activation of all pathways of the unfolded protein response in vanishing white matter disease, J. Neuropathol Exp Neurol. 65 (2006) 707-715.

[20] Chomczynski P., Sacchi N., Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction., Analytical Biochem. 162 (1987) 156-159.

[21] Jester B.C., Drillien R., Ruff M., Florentz C., Using Vaccinia's innate ability to introduce DNA into mammalian cells for production of recombinant proteins., J. Biotechnol. 156 (2011) 211-213.

[22] Giegé R., Rees B., Aspartyl-tRNA synthetases, in: Ibba M., Francklyn C. and Cusack S. (Eds.), Aminoacyl-tRNA Synthetases, Landes Bioscience, Georgetown, TX, (2005), pp. 210-226.

[23] Charrière F., O'Donoghue P., Helgadóttir S., Maréchal-Drouard L.,

Cristodero M., Horn E.K., Söll D., Schneider A., Dual targeting of a tRNAAsp requires two different aspartyl-tRNA synthetases in Trypanosoma brucei., J Biol Chem. 284 (2009) 16210-16217.

[24] Nallamsetty S., Austin B.P., Penrose K.J., Waugh D.S., Gateway vectors for the production of combinatorially-tagged His6-MBP fusion proteins in the cytoplasm and periplasm of Escherichia coli., Protein Sci. 14 (2005) 2964-2971.

[25] Hytönen V.P., Laitinen O.H., Airenne T.T., Kidron H., Meltola N.J., Porkka E.J., Hörhä J., Paldanius T., Määttä J.A., Nordlund H.R., Johnson M.S., Salminen T.A., Airenne K.J., Ylä-Herttuala S., Kulomaa M.S., Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli., Biochem. J. 384 (2004) 385-390.

[26] Ferrer M., Chernikova T.N., Timmis K.N., Golyshin P.N., Expression of a temperature-sensitive esterase in a novel chaperone-based Escherichia coli strain., Appl. Environ. Microbiol. 70 (2004) 4499-4504.

[27] Delarue M., Poterszman A., Nikonov S., Garber M., Moras D., Thierry J.-C., Crystal structure of a prokaryotic aspartyl-tRNA synthetase., EMBO J. 13 (1994) 3219-3229.

[28] Eiler S., Dock-Bregeon A.C., Moulinier L., Thierry J.-C., Moras D., Synthesis of aspartyl-tRNA^{Asp} in *Escherichia coli*-a snapshot of the second step., EMBO J. 18 (1999) 6532-6541.

[29] Almassy R.J., Janson C.A., Hamlin R., Xuong N.H., Eisenberg D., Novel subunit-subunit interactions in the structure of glutamine synthetase., Nature 323 (1986) 304-309.

[30] Janin J., Dumas C., Morera S., Xu Y., Meyer P., Chiadmi M., Cherfils J., Three-dimensional structure of nucleoside diphosphate kinase., J. Bioenerg. Biomembr. 32 (2000) 215-225.

[31] Martin F., Barends S., Eriani G., Single amino acid changes in AspRS reveal alternative routes for expanding its tRNA repertoire in vivo, Nucleic Acids Res. 32 (2004) 4084-4089.

[32] Magen A., Ast G., The importance of being divisible by three in alternative splicing., Nucleic Acids Res. 33 (2005) 5574-5582.

[33] Scheper G.C., van der Klok T., van Andel R.J., van Berkel C.G., Sissler M., Smet J., Muravina T.I., Serkov S.V., Uziel G., Bugiani M., Schiffmann R., Krageloh-Mann I., Smeitink J.A., Florentz C., Coster R.V., Pronk J.C., van der Knaap M.S., Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation, Nat Genet 39 (2007) 534-539.

[34] Modrek B., Lee C., A genomic view of alternative splicing, Nature Gen. 30 (2002) 13 - 19.

[35] Boue S., Letunic I., Bork P., Alternative splicing and evolution., Bioessays 25 (2003) 1031-1041.

[36] Ermakova E.O., Nurtdinov R.N., Gelfand M.S., Fast rate of evolution in alternatively spliced coding regions of mammalian genes., BMC Genomics 7 (2006) 84.

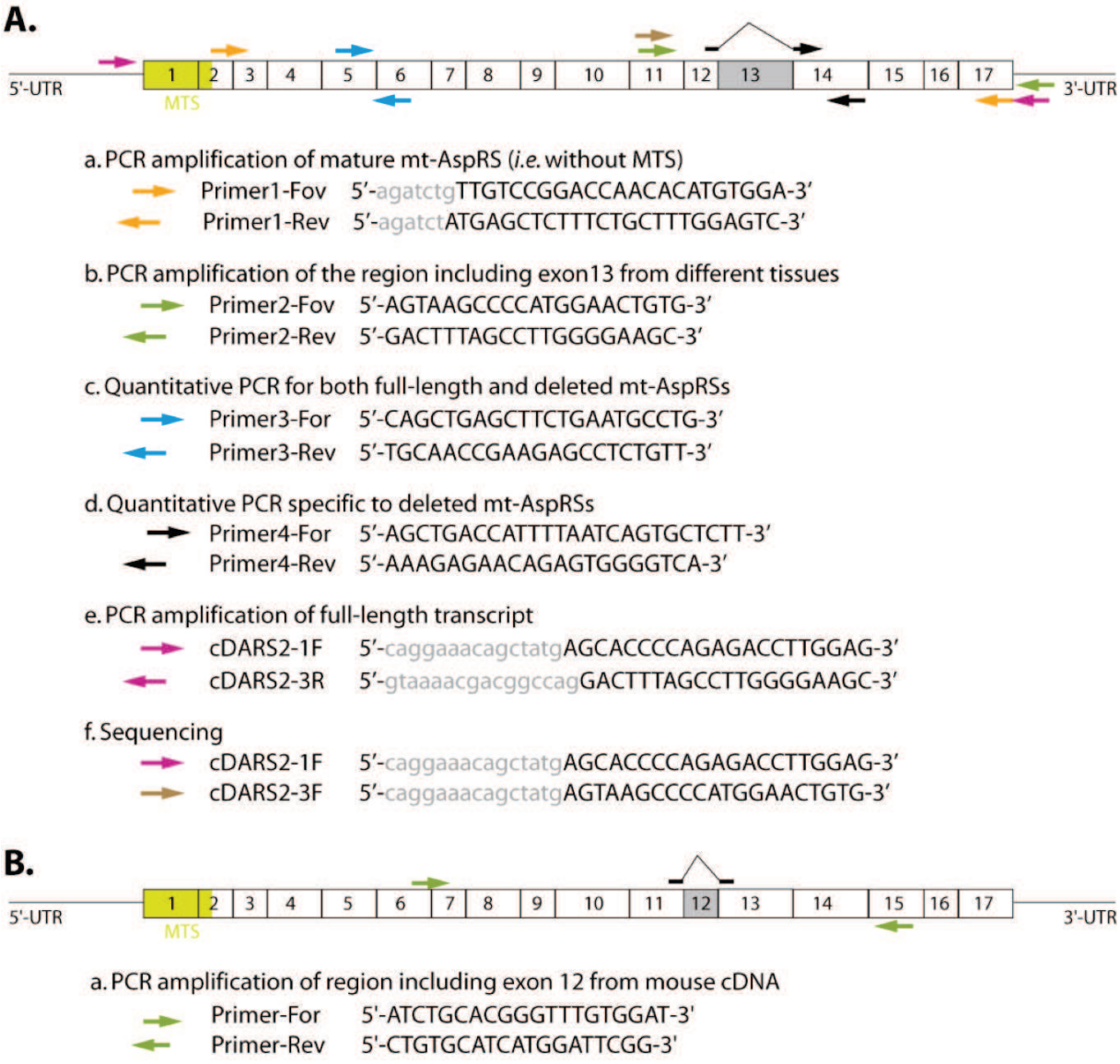
[37] Melamud E., Moult J., Stochastic noise in splicing machinery., Nucleic Acids Res. 37 (2009) 4873-4886.

[38] Keren H., Lev-Maor G., Ast G., Alternative splicing and evolution: diversification, exon definition and function., Nat Rev Genet. 11 (2010) 345-355.

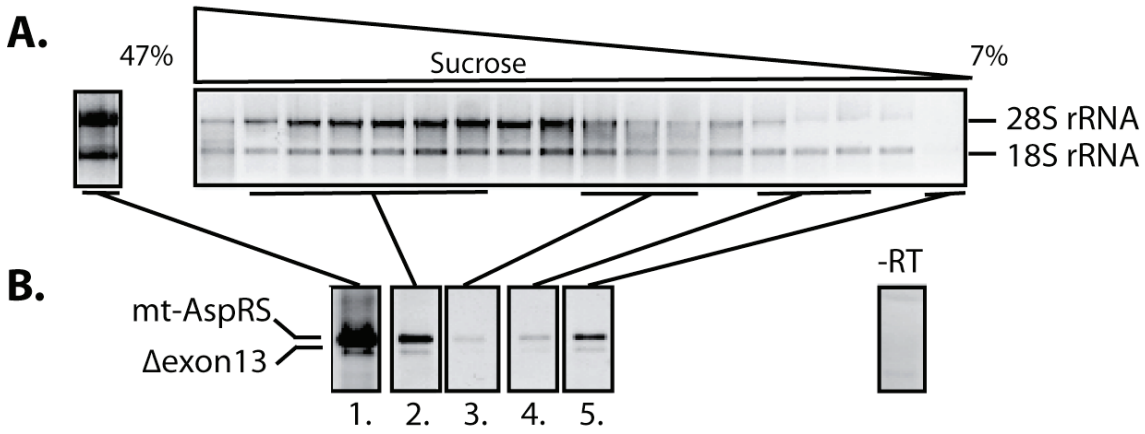
[39] Rees B., Webster G., Delarue M., Boeglin M., Moras D., Aspartyl-tRNA synthetase from *Escherichia coli:* flexibility and adaptability to the substrates, J. Mol. Biol. 299 (2000) 1157-1164.

SUPPLEMENTARY DATA

Legend to Supplementary Fig. 1: Multiple Sequences Alignment of 194 AspRS sequences (the alignment is provided as a separate file named Schwenzer_Supp_Fig3.jpg or can be retrieved at http://lbgi.fr/sissler/AspRSs). Sequences of the mt-AspRS proteins were examined in 81 eukaryotic organisms (27 Metazoa, 40 Fungi, 9 Archaeplastida (*Viridiplantae*), and 5 Protists) and aligned with AspRSs from 104 bacterial species. Nine sequences from archaeal and eukaryotic cytosolic AspRSs (*Cyanidioschyzon merolae*, *Pyrobaculum aerophilum*, *Aeropyrum pernix*, *Thermococcus kodakarensis*, *Methanothermobacter thermautotrophicus*, *Arabidopsis thaliana*, *Drosophila melanogaster*, *Saccharomyces cerevisiae*, *Homo sapiens*) were used as outgroup. The complete list of species is provided in the Supplementary Table 1.



Supplementary Fig. 2. Primers for mt-AspRSs analyses. A. Location, experimental purposes and sequences of oligonucleotides specific to **human** mt-AspRS. The scheme represents the full length mt-AspRS mRNA with boxes corresponding to the different exons. Exon 13 is highlighted in grey and predicted MTS in green. In some oligonucleotides, lower case grey letters correspond to nucleotides, which do not hybridize with mt-AspRS mRNA sequence, but have been added for cloning purposes. Purified oligonucleotides for mt-AspRS were either from Proligo (Boulder, CO), Invitrogen, and Sigma. **B.** Location, experimental purpose and sequences of oligonucleotides specific to **mouse** mt-AspRS. Exon 12 is highlighted in grey and predicted MTS in green.



Supplementary Fig. 3: Retrieval of polysome-associated mRNAs coding for mt-AspRS and mt-AspRS-Δexon13 from crude cytosolic extracts separated by sucrose gradient sedimentation. A. Total cellular extract, containing polysome bounded mRNA, were fractionated on a 7%-47% sucrose gradient. Aliquotes of RNAs, extracted from 18 fractions, were analyzed on a 1% agarose gel. Visualization of 28S and 18S rRNAs shows them as prominent RNA species and confirms the integrity of all RNA fractions. The left panel shows an aliquot of total RNA extracted from the cytolosic extract before loading on the sucrose gradient. **B.** Extracted RNAs were reverse transcribed using poly(dT) primers (specific for polyA-tails) and the obtained cDNAs were amplified with primers specific for both the full-length and exon 13-deleted mRNAs. Panels on the left correspond to total RNA loaded on gradient (1.), pooled polysome fraction (2.), pooled free 80S ribosome fraction (3.), pooled separated 60S and 40S ribosomal sub-units fraction (4.), and free or low density (non-translated) RNAs fraction (5.). Panel on the right corresponds to a representative control experiment were the reverse transcription step is omitted.

Supplementary Table I: list the 194 AspRSs sequences used for the Multiple Sequence Alignment. 185 are bacterial-type AspRSs. The outgroup is constituted by either archaeal or cytosolic-type AspRSs. For all sequence, accession numbers, name and lineage of the corresponding organism are given.

Accession number	Name	Lineage		
F9X8V8	Mycosphaerella graminicola	Eukaryota	Opisthokonta	Fungi
EKG15407	Macrophomina phaseolina	Eukaryota	Opisthokonta	Fungi
Q4WP80	Neosartorya fumigata	Eukaryota	Opisthokonta	Fungi
B6QS83	Penicillium marneffei	Eukaryota	Opisthokonta	Fungi
C5JIW2	Ajellomyces dermatitidis	Eukaryota	Opisthokonta	Fungi
J3K6B0	Coccidioides immitis	Eukaryota	Opisthokonta	Fungi
C5FSN6	Arthroderma otae	Eukaryota	Opisthokonta	Fungi
D4DDL9	Trichophyton verrucosum	Eukaryota	Opisthokonta	Fungi
G4N7B7	Magnaporthe oryzae	Eukaryota	Opisthokonta	Fungi
G2R3F1	Thielavia terrestris	Eukaryota	Opisthokonta	Fungi
EGS18066	Chaetomium thermophilum	Eukaryota	Opisthokonta	Fungi
CAE76417	Neurospora crassa	Eukaryota	Opisthokonta	Fungi
EGR45557	Trichoderma reesei	Eukaryota	Opisthokonta	Fungi
C7YJE7	Nectria haematococca	Eukaryota	Opisthokonta	Fungi
EFQ30809	Glomerella graminicola	Eukaryota	Opisthokonta	Fungi
XP_001560984.1	Botryotinia fuckeliana	Eukaryota	Opisthokonta	Fungi
D5GIM2	Tuber melanosporum	Eukaryota	Opisthokonta	Fungi
C4Y3N9	Clavispora lusitaniae	Eukaryota	Opisthokonta	Fungi
Q6BRH4	Debaryomyces hansenii	Eukaryota	Opisthokonta	Fungi
A3LXG1	Scheffersomyces stipitis	Eukaryota	Opisthokonta	Fungi
Q59QH1	Candida albicans	Eukaryota	Opisthokonta	Fungi
A5DWG2	Lodderomyces elongisporus	Eukaryota	Opisthokonta	Fungi
A5DBF3	Meyerozyma guilliermondii	Eukaryota	Opisthokonta	Fungi
C4R6V2	Pichia pastoris	Eukaryota	Opisthokonta	Fungi
Q6C0A5	Yarrowia lipolytica	Eukaryota	Opisthokonta	Fungi
Q6FRU1	Candida glabrata	Eukaryota	Opisthokonta	Fungi
C5DIL8	Lachancea thermotolerans	Eukaryota	Opisthokonta	Fungi
P15179	Saccharomyces cerevisiae	Eukaryota	Opisthokonta	Fungi
Q74ZW3	Ashbya gossypii	Eukaryota	Opisthokonta	Fungi
XP_003647747.1	Eremothecium cymbalariae	Eukaryota	Opisthokonta	Fungi
Q6CLY6	Kluyveromyces lactis	Eukaryota	Opisthokonta	Fungi
A7TGK1	Vanderwaltozyma polyspora	Eukaryota	Opisthokonta	Fungi
C5DYQ6	Zygosaccharomyces rouxii	Eukaryota	Opisthokonta	Fungi
B6K5W9	Schizosaccharomyces japonicus	Eukaryota	Opisthokonta	Fungi
O94242	Schizosaccharomyces pombe	Eukaryota	Opisthokonta	Fungi
EIE77274	Rhizopus delemar	Eukaryota	Opisthokonta	Fungi
B0CUT9	Laccaria bicolor	Eukaryota	Opisthokonta	Fungi
Q4P0F2	Ustilago maydis	Eukaryota	Opisthokonta	Fungi
CBQ71597	Sporisorium reilianum	Eukaryota	Opisthokonta	Fungi
CCF48425	Ustilago hordei	Eukaryota	Opisthokonta	Fungi
XP_002675253.1	Naegleria gruberi	Eukaryota	Heterolobosea	Schizopyrenida
XP_003388971.1	Amphimedon queenslandica	Eukaryota	Opisthokonta	Metazoa
B3RP59	Trichoplax adhaerens	Eukaryota	Opisthokonta	Metazoa

CCD58253	Schistosoma mansoni	Eukaryota	Opisthokonta	Metazoa
xp_002131496	Ciona intestinalis	Eukaryota	Opisthokonta	Metazoa
C3YTL2	Branchiostoma floridae	Eukaryota	Opisthokonta	Metazoa
XP_002931512.1	Xenopus tropicalis	Eukaryota	Opisthokonta	Metazoa
xp_002190760	Taeniopygia guttata	Eukaryota	Opisthokonta	Metazoa
Q6PI48	Homo sapiens	Eukaryota	Opisthokonta	Metazoa
Q8BIP0	Mus musculus	Eukaryota	Opisthokonta	Metazoa
G3VVB2	Sarcophilus harrisii	Eukaryota	Opisthokonta	Metazoa
E9QFP3	Danio rerio	Eukaryota	Opisthokonta	Metazoa
XP_002740304.1	Saccoglossus kowalevskii	Eukaryota	Opisthokonta	Metazoa
XP_003727977.1	Strongylocentrotus purpuratus	Eukaryota	Opisthokonta	Metazoa
NP_506019.2	Caenorhabditis elegans	Eukaryota	Opisthokonta	Metazoa
B7QK15	Ixodes scapularis	Eukaryota	Opisthokonta	Metazoa
EFX65518	Daphnia pulex	Eukaryota	Opisthokonta	Metazoa
B3MK45	Drosophila ananassae	Eukaryota	Opisthokonta	Metazoa
Q7Q3E7	Anopheles gambiae	Eukaryota	Opisthokonta	Metazoa
B0WGD1	Culex quinquefasciatus	Eukaryota	Opisthokonta	Metazoa
Q17LX8	Aedes aegypti	Eukaryota	Opisthokonta	Metazoa
xp_966901.2	Tribolium castaneum	Eukaryota	Opisthokonta	Metazoa
XP_001604266.2	Nasonia vitripennis	Eukaryota	Opisthokonta	Metazoa
E0VEP3	Pediculus humanus	Eukaryota	Opisthokonta	Metazoa
xp_001944516	Acyrthosiphon pisum	Eukaryota	Opisthokonta	Metazoa
EFZ11007	Solenopsis invicta	Eukaryota	Opisthokonta	Metazoa
EFN72741	Camponotus floridanus	Eukaryota	Opisthokonta	Metazoa
XP_002161330.1	Hydra magnipapillata	Eukaryota	Opisthokonta	Metazoa
EGD80713	Salpingoeca sp.	Eukaryota	Opisthokonta	Choanoflagellida
A9UTK6	Monosiga brevicollis	Eukaryota	Opisthokonta	Choanoflagellida
XP_002997757.1	Phytophthora infestans	Eukaryota	Stramenopiles	Oomycetes
EGG19450	Dictyostelium fasciculatum	Eukaryota	Amoebozoa	Mycetozoa
A8J1A1	Chlamydomonas reinhardtii	Eukaryota	Viridiplantae	Chlorophyta
EFN55507	Chlorella variabilis	Eukaryota	Viridiplantae	Chlorophyta
C1FGB7	Micromonas sp.	Eukaryota	Viridiplantae	Chlorophyta
Q012T8	Ostreococcus tauri	Eukaryota	Viridiplantae	Chlorophyta
A0YXF9	Lyngbya sp.	Eukaryota	Viridiplantae	Chlorophyta
B9GIN8	Populus trichocarpa	Eukaryota	Viridiplantae	Streptophyta
F4JJT9	Arabidopsis thaliana	Eukaryota	Viridiplantae	Streptophyta
C0P8Q1	Zea mays	Eukaryota	Viridiplantae	Streptophyta
XP_001753365.1	Physcomitrella patens	Eukaryota	Viridiplantae	Streptophyta
B1ZSW8	Opitutus terrae	Bacteria	Chlamydiae	Verrucomicrobia
C0A895	Diplosphaera colitermitum	Bacteria	Chlamydiae	Verrucomicrobia
zp_05056625.1	Verrucomicrobiae bacterium	Bacteria	Chlamydiae	Verrucomicrobia
Q1MPN6	Lawsonia intracellularis	Bacteria	Proteobacteria	delta/epsilon
zp_03632021.1	Pedosphaera parvula	Bacteria	Chlamydiae	Verrucomicrobia
A1V9B2	Desulfovibrio vulgaris	Bacteria	Proteobacteria	delta/epsilon
B8DMM5	Desulfovibrio vulgaris	Bacteria	Proteobacteria	delta/epsilon
Q317R3	Desulfovibrio alaskensis	Bacteria	Proteobacteria	delta/epsilon
B8J1H3	Desulfovibrio desulfuricans	Bacteria	Proteobacteria	delta/epsilon
zp_03311889.1	Desulfovibrio piger	Bacteria	Proteobacteria	delta/epsilon
C7LUJ9	Desulfomicrobium baculatum	Bacteria	Proteobacteria	delta/epsilon

zp_06232450.1	Desulfovibrio aespoeensis	Bacteria	Proteobacteria	delta/epsilon
YP_595362	Lawsonia intracellularis	Bacteria	Proteobacteria	delta/epsilon
C6E3W7	Geobacter sp.	Bacteria	Proteobacteria	delta/epsilon
A5G3L5	Geobacter uraniireducens	Bacteria	Proteobacteria	delta/epsilon
B9M1C9	Geobacter sp.	Bacteria	Proteobacteria	delta/epsilon
Q74D56	Geobacter sulfurreducens	Bacteria	Proteobacteria	delta/epsilon
B3E9R2	Geobacter lovleyi	Bacteria	Proteobacteria	delta/epsilon
A1ARG6	Pelobacter propionicus	Bacteria	Proteobacteria	delta/epsilon
zp_05711250.1	Desulfurivibrio alkaliphilus	Bacteria	Proteobacteria	delta/epsilon
zp_01292054.1	delta proteobacterium	Bacteria	Proteobacteria	delta/epsilon
Q6AQS4	Desulfotalea psychrophila	Bacteria	Proteobacteria	delta/epsilon
C0QFI1	Desulfobacterium autotrophicum	Bacteria	Proteobacteria	delta/epsilon
A0LNH3	Syntrophobacter fumaroxidans	Bacteria	Proteobacteria	delta/epsilon
Q2LTE0	Syntrophus aciditrophicus	Bacteria	Proteobacteria	delta/epsilon
D3P8H2	Deferribacter desulfuricans	Bacteria	Defferibacteres	Defferibacteres
Q6ME91	Protochlamydia amoebophila	Bacteria	Chlamydiae	Verrucomicrobia
B0K978	Thermoanaerobacter pseudethanolicus	Bacteria	Firmicutes	Clostridia
Q8RAI7	Thermoanaerobacter tengcongensis	Bacteria	Firmicutes	Clostridia
A4XHB1	Caldicellulosiruptor saccharolyticus	Bacteria	Firmicutes	Clostridia
B9MPF8	Anaerocellum thermophilum	Bacteria	Firmicutes	Clostridia
B8CXE5	Halothermothrix orenii	Bacteria	Firmicutes	Clostridia
zp_05335014.1	Thermoanaerobacterium	Bacteria	Firmicutes	Clostridia
A5D3E3	Pelotomaculum thermopropionicum	Bacteria	Firmicutes	Clostridia
B2V713	Sulfurihydrogenibium sp.	Bacteria	Aquificae	Aquificae
C1DUM5	Sulfurihydrogenibium azorense	Bacteria	Aquificae	Aquificae
C0QT74	Persephonella marina	Bacteria	Aquificae	Aquificae
A4J2J2	Desulfotomaculum reducens	Bacteria	Firmicutes	Clostridia
B1I364	Desulforudis audaxviator	Bacteria	Firmicutes	Clostridia
zp_01093975.1	Blastopirellula marina	Bacteria	Planctomycetes	Planctomycetia
D2R817	Pirellula staleyi	Bacteria	Planctomycetes	Planctomycetia
Q3SLL2	Thiobacillus denitrificans	Bacteria	Proteobacteria	Proteobacteria
zp_05105425.1	Methylophaga thiooxydans	Bacteria	Proteobacteria	Gammaproteobacteria
P21889	Escherichia coli	Bacteria	Proteobacteria	Gammaproteobacteria
Q0AIE4	Nitrosomonas eutropha	Bacteria	Proteobacteria	Betaprotobacteria
zp_02928149.1	Verrucomicrobium spinosum	Bacteria	Chlamydiae	Verrucomicrobia
B4CWE9	Chthoniobacter flavus	Bacteria	Chlamydiae	Verrucomicrobia
B2UKU7	Akkermansia muciniphila	Bacteria	Chlamydiae	Verrucomicrobia
D0MEK0	Rhodothermus marinus	Bacteria	Bacteroidetes	Bacteroidetes
B9KZD9	Thermomicrobium roseum	Bacteria	Chloroflexi	Thermomicroba
D1C339	Sphaerobacter thermophilus	Bacteria	Chloroflexi	Thermomicroba
YP_003322689	Thermobaculum terrenum	Bacteria	unclassified	Thermobaculum
Q3Z8K0	Dehalococcoides ethenogenes	Bacteria	Chloroflexi	Dehalococcoidetes
zp_06813154.1	Dehalogenimonas lykanthroporepellens	Bacteria	Chloroflexi	Dehalococcoidetes
A7NQH9	Roseiflexus castenholzii	Bacteria	Chloroflexi	Chloroflexales
B8G6W9	Chloroflexus aggregans	Bacteria	Chloroflexi	Chloroflexales
A9B392	Herpetosiphon aurantiacus	Bacteria	Chloroflexi	Chloroflexi
C8WM31	Eggerthella lenta	Bacteria	Actinobacteria	Actinobacteria
A4E8U2	Collinsella aerofaciens	Bacteria	Actinobacteria	Actinobacteria
C0GFV3	Dethiobacter alkaliphilus	Bacteria	Firmicutes	Clostridia
zp_06463648.1	Hydrogenobaculum	Bacteria	Aquificae	Aquificae
B5YJP5	Thermodesulfovibrio yellowstonii	Bacteria	Nitrospirae	Nitrospirae
D3DG67	Hydrogenobacter thermophilus	Bacteria	Aquificae	Aquificae

B5YDT0	Dictyoglomus thermophilum	Bacteria	Dictyoglomi	Dictyoglomi
B8E1C0	Dictyoglomus turgidum	Bacteria	Dictyoglomi	Dictyoglomi
Q3AA17	Carboxydothermus hydrogenoformans	Bacteria	Firmicutes	Clostridia
C9R8Y7	Ammonifex degensii	Bacteria	Firmicutes	Clostridia
C0WCD0	Acidaminococcus sp.	Bacteria	Firmicutes	Clostridia
D2RLB8	Acidaminococcus fermentans	Bacteria	Firmicutes	Clostridia
zp_06757276.1	Veillonella sp.	Bacteria	Firmicutes	Negativicutes
D3LVR5	Megasphaera genomosp	Bacteria	Firmicutes	Negativicutes
Q67LP0	Symbiobacterium thermophilum	Bacteria	Firmicutes	Clostridia
B8FQR6	Desulfitobacterium hafniense	Bacteria	Firmicutes	Clostridia
O32038	Bacillus subtilis	Bacteria	Firmicutes	Bacilli
Q65GR4	Bacillus licheniformis	Bacteria	Firmicutes	Bacilli
A8FFP5	Bacillus pumilus	Bacteria	Firmicutes	Bacilli
zp_06809319.1	Geobacillus thermoglucosidasius	Bacteria	Firmicutes	Bacilli
B7GFR1	Anoxybacillus flavithermus	Bacteria	Firmicutes	Bacilli
zp_01173653.1	Bacillus sp.	Bacteria	Firmicutes	Bacilli
D5DT18	Bacillus megaterium	Bacteria	Firmicutes	Bacilli
Q03F52	Pediococcus pentosaceus	Bacteria	Firmicutes	Bacilli
B9Y6V6	Holdemania filiformis	Bacteria	Firmicutes	Erysipelotrichi
Q5KWS9	Geobacillus kaustophilus	Bacteria	Firmicutes	Bacilli
zp_03227019.1	Bacillus coahuilensis	Bacteria	Firmicutes	Bacilli
zp_06362373.1	Bacillus cellulosilyticus	Bacteria	Firmicutes	Bacilli
D3FVX7	Bacillus pseudofirmus	Bacteria	Firmicutes	Bacilli
B9E706	Macrococcus caseolyticus	Bacteria	Firmicutes	Bacilli
Q833I2	Enterococcus faecalis	Bacteria	Firmicutes	Bacilli
C6D7W6	Paenibacillus sp.	Bacteria	Firmicutes	Bacilli
zp_03297292.1	Collinsella stercoris	Bacteria	Actinobacteria	Actinobacteria
A6Q324	Nitratiruptor sp.	Bacteria	Actinobacteria	Actinobacteria
D1AFX0	Sebaldella termitidis	Bacteria	Fusobacteria	Fusobacteria
Q8RGJ4	Fusobacterium nucleatum	Bacteria	Fusobacteria	Fusobacteria
D1BSL3	Xylanimonas cellulosilytica	Bacteria	Actinobacteria	Actinobacteria
Q5YTJ4	Nocardia farcinica	Bacteria	Actinobacteria	Actinobacteria
C7MCD1	Brachybacterium faecium	Bacteria	Actinobacteria	Actinobacteria
C7MR11	Saccharomonospora viridis	Bacteria	Actinobacteria	Actinobacteria
B2RHE0	Porphyromonas gingivalis	Bacteria	Bacteroidetes	Bacteroidetes
zp_01719781.1	Algoriphagus sp.	Bacteria	Bacteroidetes	Bacteroidetes
D5BDF5	Zunongwangia profunda	Bacteria	Bacteroidetes	Bacteroidetes
zp_02160712.1	Kordia algicida	Bacteria	Bacteroidetes	Bacteroidetes
zp_00949338.1	Croceibacter atlanticus	Bacteria	Bacteroidetes	Bacteroidetes
C7M924	Capnocytophaga ochracea	Bacteria	Bacteroidetes	Bacteroidetes
B3EUH5	Candidatus Amoebophilus	Bacteria	Bacteroidetes	Bacteroidetes
outgroup				
BAM80333	Cyanidioschyzon merolae	Eukaryota	Rhodophyta	Bangiophyceae
P14868	Homo sapiens	Eukaryota	Opisthokonta	Metazoa
Q8ZYM8	Pyrobaculum aerophilum	Archaea	Cranarchaeota	Thermoprotei
Q9Y9U7	Aeropyrum pernix	Archaea	Cranarchaeota	Thermoprotei
PDB_1EOV_A	Saccharomyces cerevisiae	Eukaryota	Opisthokonta	Fungi
PDB_3NEL_A	Thermococcus kodakarensis	Archaea	Euryarchaeota	Thermoccocci
Q9XYM1	Drosophila melanogaster	Eukaryota	Opisthokonta	Metazoa
O26328	Methanothermobacter thermautotrophicus	Archaea	Euryarchaeota	Mathanobacteria
Q8H104	Arabidopsis thaliana	Eukaryota	Viridiplantae	Streptophyta

Supplementary Method 1: Expression and purification of recombinant proteins. Numerous combinations of plasmids and *E. coli* strains were tested to express recombinant proteins. Transformants of *E. coli* TOP10 harboring pQE70-mt-AspRS-ΔExon13 were cultured in Lysogeny broth (LB) medium supplemented with 100 µg/ml ampicillin. Transformants of *E. coli* PGK JE7 (gift from M. Frugier, Strasbourg) harboring pQE70-mt-AspRS-ΔExon13 were cultured in LB medium supplemented with 100 µg/ml ampicillin, 100 µg/ml chloramphenicol and 2mg/ml L-arabinose. Transformants of *E. coli* PGK JE8 (gift from M. Frugier, Strasbourg) harboring pQE70-mt-AspRS-ΔExon13 were cultured in LB medium supplemented with 100 µg/ml ampicillin, 100 µg/ml chloramphenicol and 50 µg/ml tetracycline. Transformants of *E. coli* M15 (gift from L. Maréchal-Drouard, Strasbourg) harboring pQE70-mt-AspRS-ΔExon13 or pQE70-OmpA-mt-AspRS-ΔExon13 were cultured in LB medium supplemented with 100 µg/ml ampicillin and 25 µg/ml kanamycin. Transformants of *E. coli* BL21 derivative "Rosetta 2" (Novagen) and BL21(DE3)RIL harboring pQE70-mt-AspRS-ΔExon13, pDEST-$(His)_6$-MBP-mt-AspRS-ΔExon13 or pDEST-peri-$(His)_6$-MBP-mt-AspRS-ΔExon13 were cultured in LB medium supplemented with 100 µg/ml ampicillin and 100 µg/ml chloramphenicol. Transformants of *E. coli* ArcticExpress (gift from M. Mörl, Leipzig) harboring pDEST-$(His)_6$-MBP-mt-AspRS-ΔExon13 or pDEST-peri-$(His)_6$-MBP-mt-AspRS-ΔExon13 were cultured in LB medium supplemented with 100 µg/ml ampicillin and 20 µg/ml gentamycin. When the culture density reached an absorbance A_{600} of 0.4-0.6, 20 µM to 1 mM of IPTG was added to induce protein expression and the cultivation was then continued at several temperatures ranging from 12°C to 37°C for an additional time varying between 3 h to over-night. The cells were collected by centrifugation (1,700g) and frozen at -20°C until being analyzed.

Cells (from 20 ml cultures) were resuspended in 2 ml of one of the seven different buffers (**B1** to **B7**, see below). The suspension was sonicated on ice, three times for 30 seconds at 120 volts (Ultrasons Annemasse, France). The resulting crude extract was clarified by centrifugation (15,000g) at 4°C during 15 min. Aliquots from crude extracts, supernatants and pellets were analyzed on 12 % SDS-PAGE and were visualized by Coomassie staining. Diverse purifications were attempted on two liters of cell cultures by classical affinity chromatography and as previously described (7). When necessary, attempts to remove chaperon proteins were performed by increasing NaCl concentration to 0.5 M and/or by using 15% (v/v) isopropanol or 0.5% (v/v) Triton X-100. **B1**: 50 mM NaH_2PO_4, pH 7.5, 300 mM NaCl, 20 mM imidazole and 10 % glycerol; **B2**: 100 mM MES, pH 6.1, 250 mM NaCl, 0,1 %

Triton X-100, 10 mM β-mercaptoethanol and 10 % glycerol; **B3**: 100 mM MOPS, pH 7.0, 250 mM NaCl, 0,1 % Triton X-100, 10 mM β-mercaptoethanol and 10 % glycerol; **B4**: 100 mM NaH_2PO_4, pH 7.2, 250 mM NaCl, 0,1 % Triton X-100, 10 mM β-mercaptoethanol and 10 % glycerol; **B5**: 100 mM MOPS, pH 7.0, 250 mM NaCl, 10 mM $MgCl_2$ and 10 % glycerol; **B6**: 100 mM Tris-HCl, pH 8.1, 250 mM NaCl, 0,1 % Triton X-100, 10 mM β-mercaptoethanol and 10 % glycerol; **B7**: 50 mM NaH_2PO_4, pH 7.5, 300 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol and 10 % glycerol.

Supplementary Method 2. Mass spectrometry analysis

Samples preparation. Mitochondria purification was carried out on freshly prepared HEK293T or Myoblasts cells (optained from Myosix) as described before [40]. Frozen mitochondrial pellets were resuspended in 30 mM HEPES pH 7.5 with 200 mM potassium acetate. Mitochondrial matrix and membrane fractions were obtained after sonication and centrifugation at 100 000g. Membrane proteins were solubilized with 0.5 % Triton X-100. Gel electrophoresis of a 10 % SDS-PAGE were performed in a Protean II xi Cell (BioRad) for 2h at 25V and 5h at 15W. Proteins were stained with Simply Blue^{TM} Safe stain (Invitrogen). 5 µg of recombinant mt-AspRS or mt-AspRS-Δexon13 were loaded as size marker.

For 2-D Gel electrophoresis, mitochondrial pellet were solubilized in 7 M Urea, 2 M Thiourea, 4% CHAPS and 20 mM Tris pH 8.5 for 1h at 30°C. Samples were prepared by a reduction step with 50 mM DTT for 10 min at RT and alkylation step with 100 mM iodoacetamid for 15 min at 37°C, followed by a purification step using Sephadex G-25 purification step. First-dimension separation was performed on immobilized pH gradient (IPG) strip at pH 5–8 (BioRad) with 60000 Vh. Second Dimension separation was performed on a 10% SDS-PAGE.

In gel digestion. Each gel slice was cut into small pieces with a scalpel, washed and dehydrated two times with respectively 100 µl of 25 mM NH_4HCO_3 and with 100 µl of acetonitrile. Reduction was achieved by 1h treatment with 10 mM DTT at 57°C. Alkylation reaction was performed by incubation with 25 mM iodoacetamide for 45 min at room temperature, protected from light. Finally, gel spots were washed 3 times for 5 min alternately with 25 mM ammonium carbonate and acetonitrile. Gel pieces were completely dried with a Speed Vac before tryptic digestion. Tryptic digestion was performed using Trypsin Profile IGD Kit (PP0100 Sigma) at 35°C overnight. The gel pieces were centrifuged and 7 µl of 35% H_2O / 65% acetonitrile / 5% HCOOH were added for peptide extraction. The mixture was

sonicated for 5 min and centrifuged and directly used for MALDI-MS. For nanoLC-MS-MS, the supernatant solvent was completely evaporated to remove all acetonitrile from the sample. 10 µl of H_2O / 5% HCOOH were added prior to injection in the HPLC system.

MALDI mass spectrometry

Mass measurements were carried out on an Autoflex (Bruker-Daltonik GmbH, Bremen, Germany) Matrix Assisted Laser Desorption Time Of Flight mass spectrometer (MALDI-TOF-TOF). A saturated solution of α-cyano-4-hydroxycinnamic acid in 50 % water / 50% acetonitrile was used as a matrix. 0.5 µl of sample was deposited and 0.5 µl of saturated matrix solution was added. The preparation was dried. The sample was washed by applying 1 µl of formic acid (5 %) solution on the target and then flushed after few seconds. All mass spectra were calibrated with peptide mixture for calibration "Peptide Calibration Standard II" from Bruker (#222570). The resulting peptide mass fingerprints (PMFs) were searched against the non-redundant database SWISS-PROT (http://www.expasy.ch/sprot) using MASCOT search program. The parameters used in the search were as follows: peptide mass tolerance 50 ppm, 1 missed cleavage, methionine oxidation and cystein carboxymethylation. Each peak was also manually verified and attributed either to mt-AspRS or to mt-AspRS-Δexon13

Liquid chromatography tandem mass spectrometry. Samples were injected into an Ultimate 3000 nano HPLC (Dionex). 6.4 µl of each extract sample was injected via "microliter pickup" mode and desalted on-line through a 300 µm x 5 mm C_{18} trapping cartridge (Thermo Scientific, USA). The peptides were separated on a 75 µm x 15 cm, 3 µm C_{18} 100 Å, Acclaim® PepMap100 column (Thermo Scientific, USA) prior to the introduction into the mass spectrometer. Mobile phase A was 0.1% TFA in water and B was 0.08% TFA in acetronitrile. A linear gradient from 5% B to 65% B in 30 min, then to 80% B in 5 min, was used for separation at a flow rate of 300 nL/min. Mass calibration was performed using Tune Mix Low (Agilent Technologies # G1969-8500). Mass data were obtained using a MicrOTOF-Q II (Bruker-Daltonik GmbH, Bremen, Germany) fitted with Z-spray nanoflow electrospray ion source. The mass spectrometer was operated in positive ion mode with a potential of 3500V applied to the nanoflow. MS data was acquired from 300 to 2000 m/z and from 50 to 2000 m/z for MS/MS. Fragmentation was performed using argon as the collision gas and with a collision energy profile optimized for various mass ranges of precursor ions (only doubly and triply charge states are selected for fragmentation). An "inclusion list" (574.6/861.4/996.0/664.3 m/z) with m/z value corresponding to 3+ and 4+ of carboxymethylated peptides 350/364 and 344/366 -specific of mt-AspRS-Δexon13- was

added to MS/MS conditions aiming to increase chances to identify mt-AspRS-Δexon13. In recombinant sample, mt-AspRS-Δexon13 has been identified by peptide 350/364 carboxymethylated at RT = 24.5 min by its doubly-charged ion (m/z experimental 861.4) and triply-charged ion (m/z experimental 574.6). The data were processed by Data Analysis and Biotools softwares (Bruker Daltonik GmbH, Bremen, Germany) to generate ".mgf" lists which were subjected to NCBInr using Mascot search program. Searches were done with a large tolerance on mass measurement of 100 ppm in MS mode and 0.5 Da in MS/MS mode, to avoid any miss of interest peptides.

Supplementary reference:

[40] Messmer M., Florentz C., Schwenzer H., Scheper G.C., van der Knaap M.S., Maréchal-Drouard L., Sissler M., A human pathology-related mutation prevents import of an aminoacyl-tRNA synthetase into mitochondria., Biochem J. 433 (2011) 441-446.

3 Detection of the putative protein corresponding to mt-AspRSΔExon13

The efforts towards *in cellulo* and *in vivo* detection of a protein corresponding to mt-AspRSΔExon13 have been extensively discussed in the article (Article #1). For the sake of completeness, some additional facts are presented here. To recall, the aim of our efforts was the detection of the peptide N_{389}FAADHFNQ/CSLLGK$_{403}$ of mt-AspRS-Δexon13 by mass spectrometry (Supplementary Method 2). This peptide corresponds to the boundary of exon 12/exon 14 -encoded regions, is experimentally generated by trypsin digestion (as verified on recombinant protein, not shown), and is specific to mt-AspRS-Δexon13 (not found in any other human protein as verified when screening human sequences database).

A first series of experiments performed on 1D-SDS PAGEs and subsequent analysis by mass spectrometry led to the observation that the protein separation was not sufficient for the detection of low abundant proteins. Most of the time high abundant proteins masked the detection of other proteins (data not shown). Therefore, 2D-gel electrophoresis was performed in collaboration with the Proteomic Platform Esplanade. In a first attempt, 500µg of total mitochondrial proteins, extracted from HEK293T cells, were separated on pI gradient of 5-8 and resolved on 10% SDS-PAGE. In parallel, 5µg of recombinant mt-AspRS and mt-AspRSΔExon13 (containing high amount of chaperons) were loaded and separated in the same way on a 2D-Gel. Both gels were colored with colloidal blue and scanned. The overlay of both gels is shown in **Figure 8**. The pI of the recombinant protein was predicted to be 6.53 and 7.07 for the mt-AspRS and mt-AspRSΔExon13, respectively. However, under experimental conditions, the recombinant proteins were identified in the area corresponding to a pI of 8. The area of the gel, containing the separated mitochondrial proteins was sliced out and systematically analyzed by LC-MS/MS. Samples were injected into an Ultimate 3000 nano HPLC (Dionex). Mass data were obtained using a MicrOTOF-Q II (Bruker-Daltonik GmbH, Bremen, Germany) fitted with Z-spray nanoflow electrospray ion source. No specific peptide corresponding to the mt-AspRSΔExon13 was detected.

This method was repeated with a mitochondrial extract obtained from $2x10^8$ myoblast cells (generous gift from MYOSIX, France). The total mitochondrial proteins were separated on pI gradient of 6-11 and resolved on 10% SDS-PAGE. The area corresponding the pI of 7-10 and size of 50-90kD were systematically cut out and divided into 100 gel samples. The analysis of the 100 samples results in the identification of 318 proteins with a mascot score above 40. Among these proteins only the LysRS and GlyRS were identified. The mitochondrial and cytosolic forms of this proteins share the same sequence, thus we could not differentiate between the homologues. It has also been shown that the corresponding mRNAs are the most abundant ones in samples derived from 20 different human tissues (see Bookchapter #1 in the Introduction). This high mRNA abundance may reflect the detection here of only two aaRSs, which belong to a family of underrepresented species.

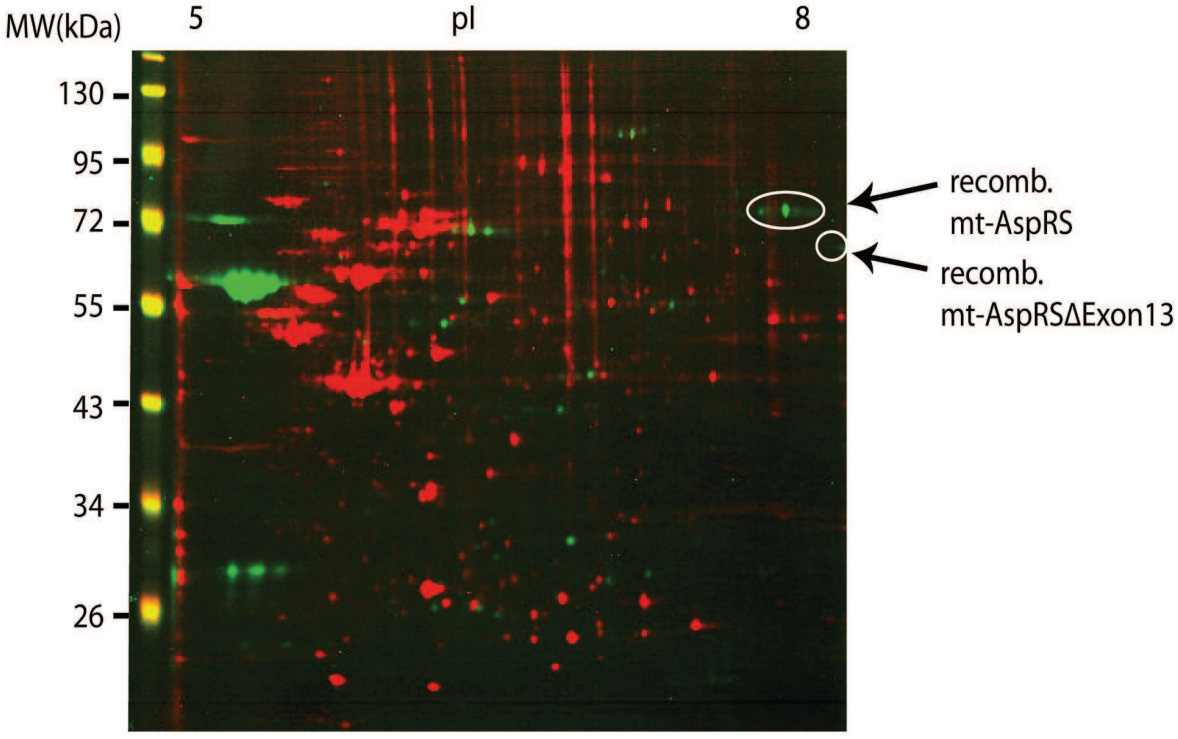


Figure 8: Human mitochondrial proteins separated by 2D-gel electrophoresis. Overlay of control experiment (green spots corresponding to recombinant mt-AspRS and recombinant mt-AspRSΔExon13) and 2D-gel separating the full mitochondrial protein extract, extracted from HEK293T cells (red spots). White circles indicate the area of the identified recombinant proteins. Green spots at approximately 57kDa were identified as GroEL chaperone proteins.

When we decreased the threshold for the mascot score from 40 to 10, we were able to identify additional 698 proteins. Among these proteins we could identify two additional cytosolic aaRSs (HisRS and IleRS), the mitochondrial TyrRS but no peptides corresponding neither to mt-AspRS nor mt-AspRSΔExon13.

In summary, in both analyses it was not possible to identify peptides corresponding to either the mt-AspRS or the mt-AspRSΔExon13 proteins. We cannot explain why we could not identify any peptides corresponding neither to mt-AspRS nor mt-AspRSΔExon13. The experimental determination of the theoretical detection limit of the nanoLC-MS/MS analysis was estimated to be 32 ng of injected protein (mt-AspRS) by a sensitivity test (discussed in article #1). We estimated that the loaded extract contains approximately 8µg of mt-AspRS and 400-800ng of protein produced from the alternative spliced transcript. However, despite good theoretical experimental conditions, the mt-AspRS could not be detected, suggesting that additional factors (such as *e.g.* separation condition, dilution in gel) may limit the detection of mt-AspRS and mt-AspRSΔExon13 on the 2D-Gels.

Of note, the determination of the detection limit was performed on 1-D SDS-PAGE with a limited amount of background proteins. It turned out, that the 2D-Gel electrophoresis improved the separation of the proteins but did not overcome the problem of detection limit caused by a high background.

The total absence of a protein corresponding to mt-AspRSΔExon13 could have several different explanations. The excision of the peptide sequence corresponding to exon 13 occurs at extremities of well-defined structural domains. Despite the fact that the excision does not hamper the overall structure, and that fact that we found the corresponding mRNA (missing exon 13) bounded to polysomes suggesting an active translation, we speculate that the protein is unstable and the nascent protein is directly degraded. This hypothesis could be tested through the experimental inhibition of the intrinsic proteasome. This machinery degrades unneeded, misfolded or damaged proteins. Several proteasome inhibitors exist e.g. Bortezomib (Adams and Kauffman, 2004).

4 RNAi knock-down of mt-AspRS and mt-AspRSΔExon13

The difficulty in producing an *in vitro* recombinant protein coupled with our inability to detect the protein *in vivo* prevents the full functional characterization of mt-AspRSΔExon13 (discussed in article #1). Nevertheless, knock down experiments were performed to analyze the possible relevance of the protein for the viability of the cells. To do so, HEK293 cells were transiently or stably transfected with different siRNA or shRNA against the mRNA encoding mt-AspRS, mt-AspRSΔExon13 or both. Following siRNAs were used:

siRNA-FL	→ (against full-length mt-AspRS mRNA)
siRNA-Δ13	→ (against mt-AspRSΔExon13 mRNA)
siRNA-(FL+Δ13)	→ (against both mRNAs)
siRNA-NR	→(against a non-target sequence)
siRNA-scramble	→(sequence of siRNA-Δ13 is randomly mixed)

The cell proliferation rate was analyzed by counting the cells or determined with the CellTiter 96® AQ_{ueous} Non-Radioactive Cell Proliferation (Promega). The latter method is a homogeneous, colorimetric assay for determining the number of viable cells in a time dependent manner.

4.1 Proliferation of transient transfected cells with siRNA targeting mt-AspRS or mt-AspRSΔExons13

To determine a possible effect on cell proliferation after the depletion of mRNA coding for mt-AspRS or mt-AspRSΔExon13 by siRNA, a cell viability assays were performed by counting the number of living cells 24h, 48h, 96h and 120h after transfection. These experiments were performed under restrictive conditions. The use of galactose media helps to distinguish healthy cells from those with defective mitochondrial functions, as cells

with defective mitochondrial function cannot proliferate in a non-fermentable growth medium.

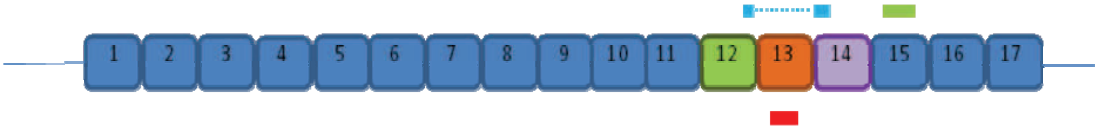


Figure 9: Schematic illustration of the hybridization sites of siRNA. Green bar displays the site of siRNA-(FL+D13); Red bar illustrates the site of siRNA-FL and blue bar indicates the exon junction spanning site of siRNA-Δ13.

Cells were transiently transfected with siRNA against the full-length transcript (siRNA-FL), the alternative spliced transcript (siRNA-Δ13), both transcripts (siRNA-(FL+ Δ13)) or with a control siRNA (siRNA-NR) (**Figure 9**). The mRNA knock down was monitored by RT-PCR on total RNA extracted from transfected cell lines after 96h and 120h. Messenger RNA depletion was clearly observed as a result of RNAi knockdown (**Figure 10**).

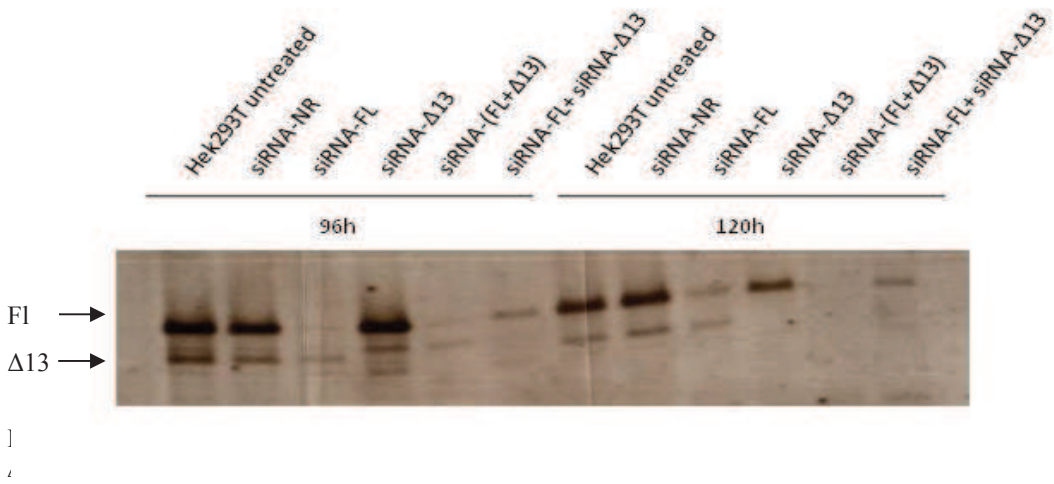


Figure 10: Verification of siRNA-mediated knockdown of mRNA encoding mt-AspRS and mt-AspRSΔExon13 in HEK293T cell by RT-PCR. Total RNA was extracted 96h and 120h post-transfection. RT-PCR with specific primers was performed and products were loaded on agarose gel. Two amplicons representing mt-AspRS (FL) and mt-AspRSΔExon13 (Δ13) are visible in the non-transfected (untreated) HEK293 lane. Reduced mRNA expression are observed when cells are treated with the corresponding siRNA.

Figure 11 shows growth curves of transfected HEK293T cells. Of note, non-transfected and siRNA-NR transfected cells showed similar growth behaviors. Transfection with siRNA-FL resulted in a 70% reduced cell number compared to siRNA-NR transfected cells, indicative for a reduced proliferation (or enhanced apoptosis). Conversely, transfection with siRNA-Δ13 resulted in a 30% reduced cell number, which corresponds roughly to the level seen when the cells were transfected with siRNA-scramble. Cells transfected with siRNA-FL and siRNA-Δ13 displayed the same phenotype as cells transfected with siRNA-Δ13. In contrast, cells transfected with siRNA against both mRNAs showed a milder growth defect compared to cells transfected with siRNA-FL alone. Those experiments were kindly performed by L. Echevarria (Madrid, Spain).

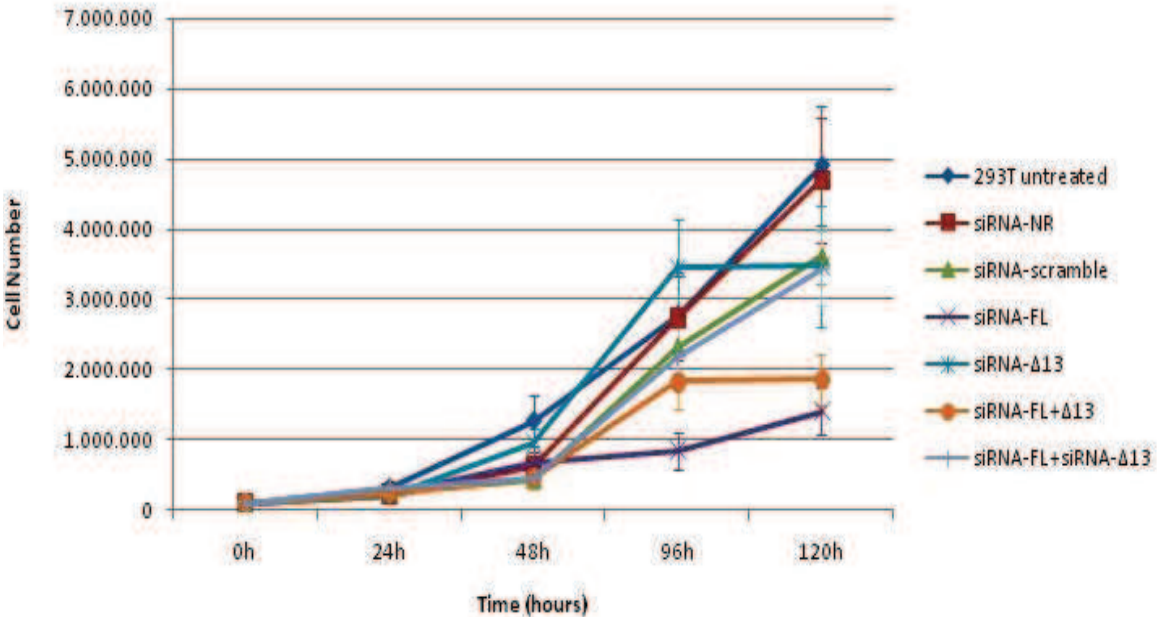


Figure 11: Growth curves of HEK293T cells transiently transfected with siRNAs. Cells were transfected with different siRNAs and have grown under limited conditions (galactose medium) to enhance the selection against mitochondria deficit cells. Number of cells was manually counted at indicated time points post-transfection (Data were kindly provided by L. Echevarria, Madrid, Spain).

4.2 Proliferation of stably transfected cells with shRNA targeting mt-AspRS or mt-AspRSΔExons13

To confirm the previous results and to investigate possible long-term effects of the knockdown of mt-AspRS and mt-AspRSΔExon13, we produced stable cell lines expressing shRNAs targeting the mRNA corresponding either to the full length or the alternatively spliced transcript (prepared and kindly provided by L. Echevarria, Madrid, Spain). shRNA expression was under the control of the tetracycline repressor protein (TetR) binding to its cognitive operator sequence (TetO). In the absence of Dioxycline, TetR binds to the TetO and inhibits the binding of the Polymerase III. In the presence of Dioxycycline, TetR does not bind to TetO and the expression of the target gene occurs (Tet-on expression) (Kappel et al., 2007). The principle of this promotor system is given in **Figure 12**. The proliferation rate of shRNA expressing cells was analyzed using a proliferation assay measuring the metabolic activity of vital cells.

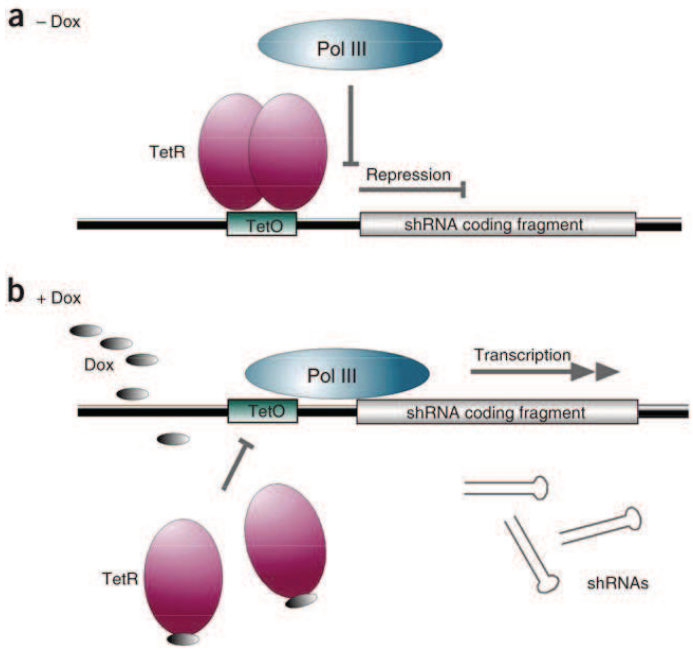


Figure 12: Regulation of the expression of shRNA by Dioxycycline. Figure has been taken from Kapple et al., 2007 (a) In the absence of Dioxycycline (Dox) the tetracycline repressor protein (TetR) binds to its

cognitive operator sequence (TetO) and represses the shRNA transcription. (b) In the presence of Dox, TetR is no able to bind TetO and the Polymerase III can transcribe the shRNA.

In a first set of experiments the conditions for the cell proliferation assay was optimized. To do so, different amounts of cells were seeded on a 96-well plate and the proliferation rate was measured over 6 days. The optimal ratio between number of seeded cells and the intensity of the signal was determined with 2000 cells per well in a 96-well plate. This amount of cells ensures a signal sufficiently strong at the beginning of the experiment whilst remaining unsaturated after 6 days of incubation. In a second phase of optimization, the knock down of the mRNA was controlled at the protein level. To do so, same amount of cells were seeded on a 6-well plates and the expression of shRNA was induced by Dioxycycline. Cells were harvested and lysed after 12h, 24h, 36h, 48h, 72h, 96h and 120h post-induction and the protein lysates were analysed by western blotting (**Figure 13**). The mt-AspRS was detected with a protein specific antibody. Prohibitin was detected as a loading control. A significant knockdown was observed after 48h post-induction for the mt-AspRS. The knock down of the splice variant could not be verified by western blot detection. For that reason, total RNA was extracted from shRNA-Δ13 induced cells and cDNA synthesized. With gene specific primers for the gene encoding mt-AspRSΔExon13, a knock down was confirmed on the mRNA level. The corresponding mRNA was not detectable after 24h, while the mRNA of the full-length transcripts retained.

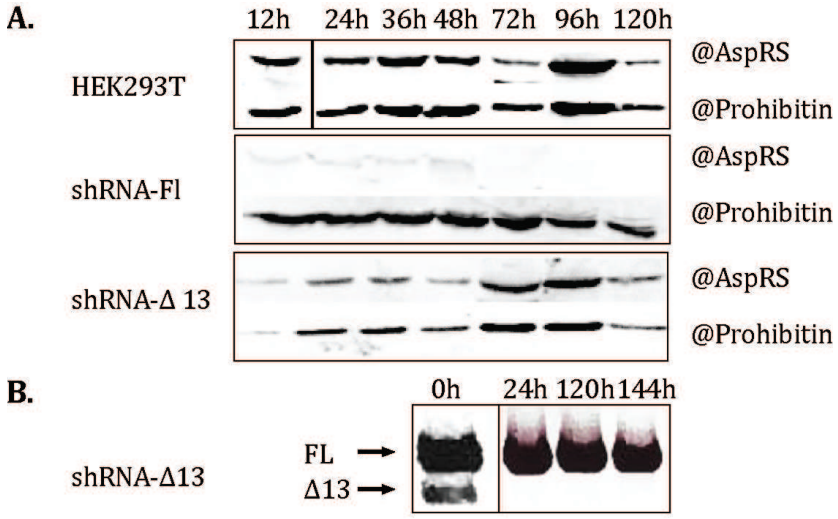


Figure 13: shRNA-mediated knock down of mt-AspRS and mt-AspRSΔExon13. Expressions of shRNAs were induced using Dioxycyline. Cells were harvested and lysed at the indicated time points. A) Crude protein lysates were analysed by western blotting. Panels show western blot identification of cells expressing shRNA against mt-AspRS and mt-AspRSΔExon13 as well the control cell line (HEK293T) expressing no shRNA (non-induced stable transfected cell lines). Specific antibodies were used to detect mt-AspRS or Prohibitin. The latter serves as loading control. B) Total RNA from cells expressing shRNA against mt-AspRSΔExon13 were extracted and subjected to RT-PCR using gene specific primers. PCR products are separated on 1% agarose gel and stained with EtBr.

This experiment was performed to control the knockdown of the protein corresponding the full-length and alternatively spliced transcript. In addition, we also suggest that it might be possible to detect the splice variant by western blotting, when we knock down the full-length protein. Of note, the size difference of both proteins is only 5kDa. Unfortunately, these experiments did not lead to a detection of a protein corresponding to the splice variant despite the knockdown of the full-length protein and the assumption that the mt-AspRSΔExon13 is expressed. But in contrast, these experiment also showed that the knockdown of the mRNA corresponding to the splice variant did not impact the translation of the full-length mt-AspRS. Thus, it can be assumed that mRNA corresponding to mt-AspRSΔExon13 does not play any regulatory role (as proposed as possible function of splice variants in e.g.(Kelemen et al., 2013; Keren et al., 2010), at least for the expression of mt-AspRS (under our experimental conditions).

Taken all optimized parameters together, the starting point for the proliferation measurement was defined at 48h post-induction and the proliferation rates of cells expressing either shRNA against full-length transcript or the alternative spliced transcript were analyzed over 6d post- induction.

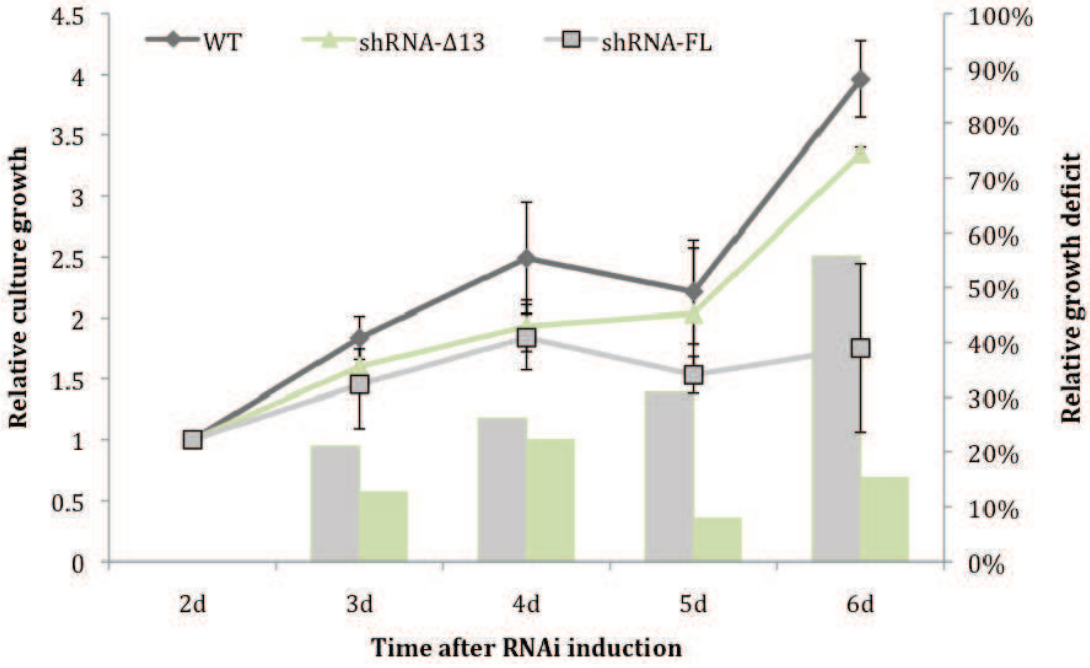


Figure 14: Proliferation rate of cells expressing shRNA-FL and shRNA-Δ13 over 6 days. Expressions of shRNA were induced with Dioxycyline. One day after induction, cells were seeded on 96 well plates. Measurements of vital cells were performed on day two post-induction. Absorbance at 490nm was measured and normalized against the value measured at 2d. The relative culture growth (lines) is displayed in the left vertical axis. The relative growth deficit (bars) compared to WT is displayed in the right vertical axis. WT correspond to the growth of non-induced cells. Values are mean values of three independent experiments. Error bars indicate standard variations.

Each value was normalized against samples taken at day 2 (**Figure 14**) post-induction. The calculated value represents the relative culture growth, describing the doubling time of the cells.For an easier illustration, the growth curves were further normalized against the non-induced cells being displayed as the wild type growth curve. The mean growth curves of three independent experiments are displayed in **Figure 14**. In addition, the relative growth deficit compared to the wild type cells are displayed. Knock-down of the full-length mRNA showed a significant effect on the growth of the cells

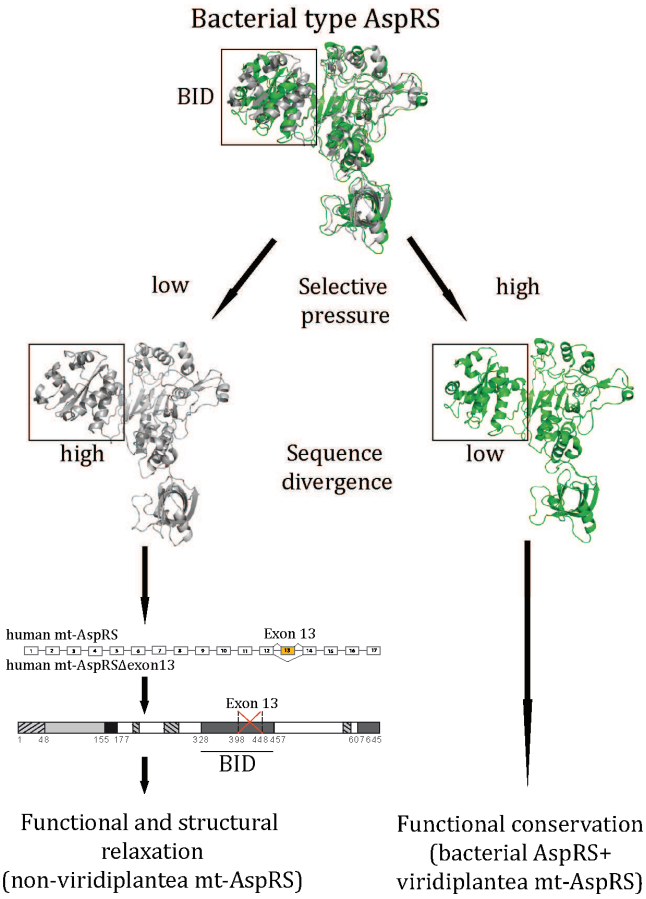
(deficit of up to 60% at day 6) while the knock down of the splice variant mRNA showed a growth deficit of only 15 % after the same period of time. To extend this investigation to a longer period of time, cells should be grown in 10cm plates and split at 6 days post-induction. Preliminary, experiments measuring the cell proliferation indicate no further decreased proliferation rates for neither the cells expressing shRNA-FL nor shRNA-Δ13.

In summary, these experiments confirm the previously data obtained from transiently transfected cells expressing siRNA-FL or siRNA-Δ13. Altogether, they suggest that the knockdown of the mRNA encoding the mt-AspRS leads to decrease in growth rate up to 70%, while the knock down of the mRNA encoding the alternative spliced protein does not seem to have a significant effect on the proliferation rate of HEK293T cells. We can hypothesize that the alternative transcript has no regulatory role or essential function for the expression of the mt-AspRS or survival of the cell, respectively. Thus, these data may further support our hypothesis that there is an ongoing decline of the bacterial insertion domain, a domain that is already missing in AspRS of cytosolic location in eukaryotes, or from archaea and which have likely no function in mt-AspRS of non-viridiplantea.

5 Achievements at a glance

- We show the discovery of alternatively spliced forms of mt-AspRS
- We show that a different selective pressures occurs within bacterial insertion domains of mt-AspRS
- We speculate about a functional relaxation of mt-AspRS in non-viridiplantea
- We hypothesized about a evolutionary mechanisms for adaptation of nucleus-encoded proteins to mitochondrial enviroment

6 Graphical summary



Chapter II: Sub mitochondrial localization of aaRSs

1 Introduction

Mitochondria are cellular organelles separated from the cytosolic compartment by a double membrane. Mitochondrial proteins are located either in the outer membrane (e.g. porine), inter membrane space (e.g. cytochrome C), inner membrane (e.g. proteins of respiratory chain complexes) or matrix (e.g. SOD2). Proteins found in the membrane can be either deeply embedded into the membrane (integral proteins) or more loosely attached to the membrane (peripheral proteins). This classification was primarily defined within the fluid mosaic model (Singer and Nicolson, 1972) (**Figure 15**).

Integral proteins are embedded into the lipid bilayer *via* one or more membrane anchoring alpha helices or *via* protein channels built from beta-barrel structures. Peripheral proteins can be attached through direct or indirect interactions with the membrane, and these interactions may be specific or non-specific. For example, single amphiphilic helixes or hydrophobic turns in the structure of the protein can interact with the membrane through non-specific hydrophobic interactions. Positively charged residues can also interact non-specifically with negative charged phospholipids heads. Alternatively, proteins can be attached by covalently bound lipid anchors (e.g. farnesyl and palmitoyl groups) or by specific protein-lipid binding of certain proteins domains (C1 or annexin type domain). Finally, proteins can be indirectly associated with the membrane by interaction with membrane bound integral proteins. The most prominent example of such an interaction is the complex formed by the F(1) and F(0) subunits of the ATP-Synthase. Of note, the classification in peripheral and integral protein is not always clear e.g. farnesyl groups can lead to a strong anchorage to the membrane similar to integral proteins.

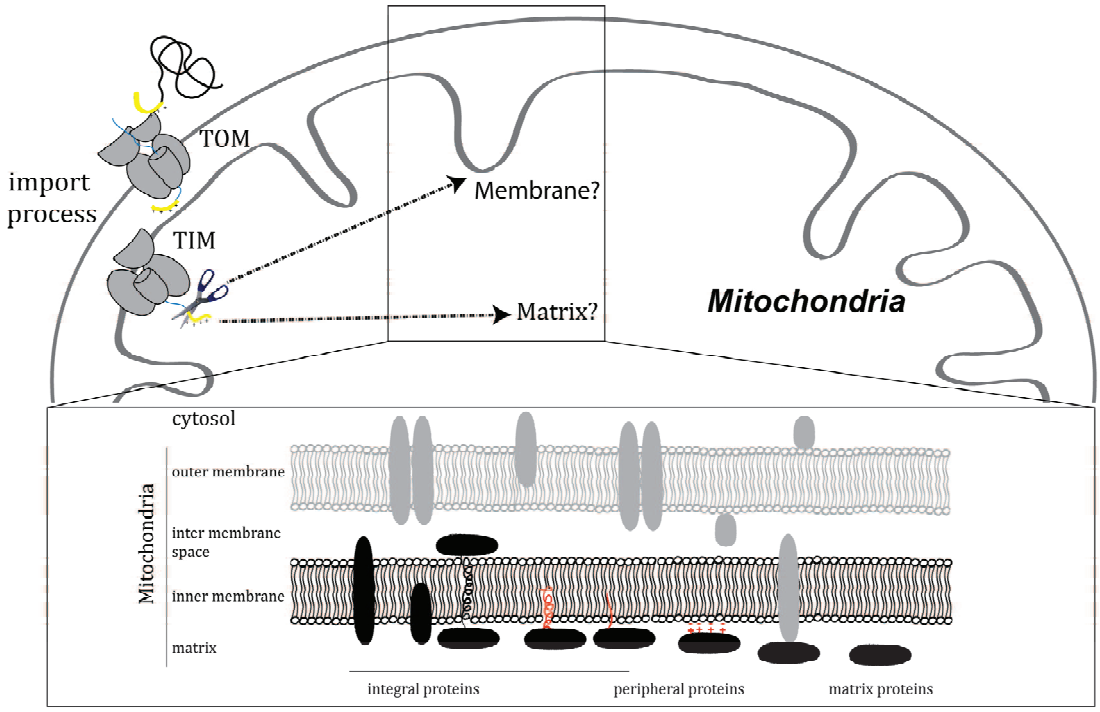


Figure 15: Scheme of the possible distribution of aaRS proteins within the mitochondrial matrix or along the inner membrane. After cytosolic expression, mitochondrial proteins are internalized into the mitochondria thanks to the mitochondrial targeting sequence (yellow). The hypothesized location of the mt-aaRSs in matrix and along the membrane is illustrated. Proteins can be differently attached as "integral" or "peripheral" protein to the inner membrane. Integral proteins are completely or partially embedded into the lipid double layer. Peripheral proteins are localized along the matrix side of the inner membrane and attached by electrostatic, ionic interaction or via single hydrophobic domain or posttranslational modifications to the membrane. In red the important interaction for membrane localization are highlighted.

Our goal was to investigate the localization profile of the full set of human mt-aaRSs and to decipher the possible links between the localization, function and the impact of pathology-related mutations (see chapter V). To address these goals, purified mitochondria from HEK293 cells were separated into fractions containing matrix-located proteins, integral membrane proteins or proteins peripheral to the inner membrane. The fraction were analyzed by:

I. Western blot detection for their specific protein content
II. Northern blot for their tRNA content
III. Mass spectrometry for their total protein content

A second approach connects the fractionation with an engineered and recombinant expression strategy (Chapter III). A BHK21 (Baby Hamster Kidney) cell/Vaccinia expression platform has been developed in the lab that allows the modulated expression, purification and detection of recombinant proteins in mammalian cells (Jester et al., 2011). We validated this method by expressing mitochondrial recombinant proteins for *in vivo* and *in vitro* characterizations (article#2). We also used it to investigate possible links between the presence or absence of motifs within the MTS and the sub-mitochondrial localization.

2 Sub-mitochondrial fractionation and detection of aaRSs

2.1 Validation of the mitochondrial fractionation protocol

At first, we established a valid and robust protocol to purify and fractionate mitochondria (extracted quantitatively from HEK293F cells) into highly pure matrix, peripheral (membrane associated) and integral (membrane embedded) protein fractions (**Figure 16**) (the protocol is adapted from (Suzuki et al., 2007).

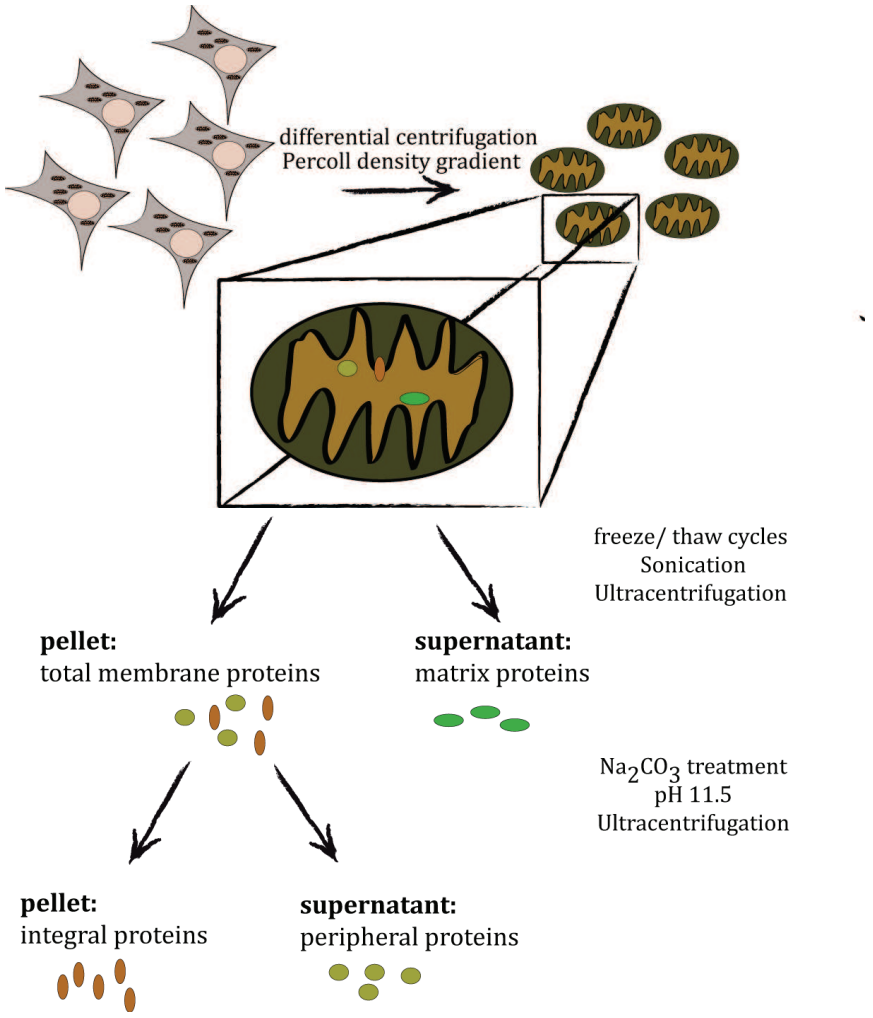


Figure 16: Flowchart of mitochondria purification and fractionation. After cell lysis, mitochondria are recovered by ultracentrifugation and purified on a density gradient. Purified mitochondria are fractionated into matrix and total membrane protein containing fractions. The total membrane proteins fraction is further fractionated into integral and peripheral proteins.

The standard protocol for mitochondrial purification includes sub-cellular fractionation by differential centrifugations and a subsequent discontinuous *Percoll* density *gradient centrifugation* step to obtain highly purified mitochondria. Purified mitochondria were then subjected to a combination of freeze/thaw cycles, followed by a sonication step. Soluble matrix proteins were separated from the total membrane proteins by ultracentrifugation. The obtained pellet (containing total membrane) was treated with Na_2CO_3 to strip peripheral membrane proteins from the lipid layer. The overall protein composition of the different fractions was visualized by separating the samples on a 10% SDS-Gel and coomassie staining (**Figure 17** A). The different protein composition in each fraction is shown as a diverse protein running pattern in the gel.

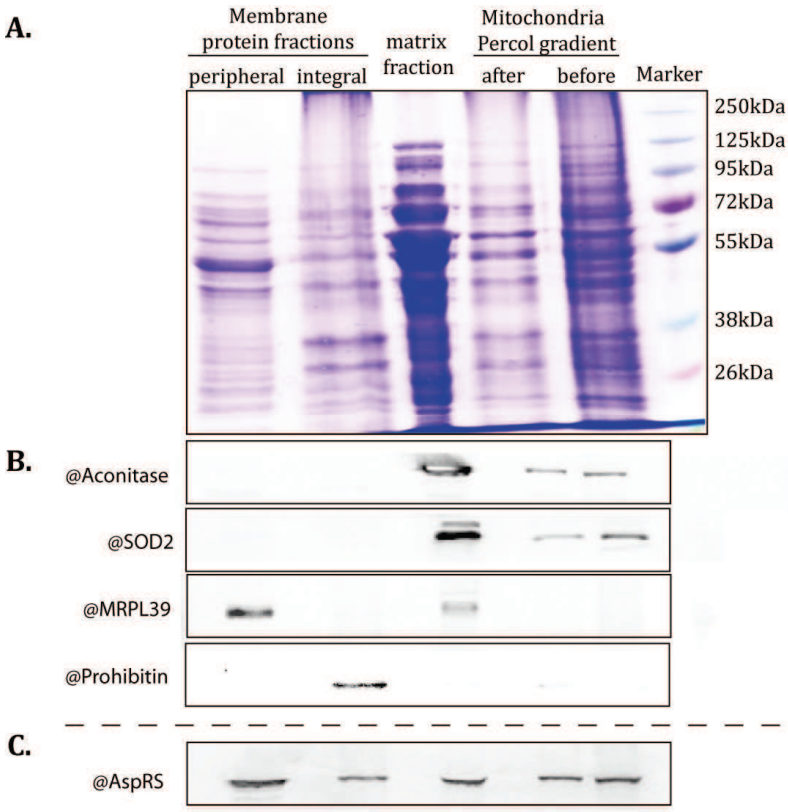


Figure 17: Quality control of mitochondrial fractions A: Separation of sub-fractions on a 10% SDS-PAGE. Proteins are stained with coomassie staining. **B:** Proteins from mitochondrial fractions, were separated on 10% SDS-PAGE and transferred on PVDF membrane. The quality of fractionation was analyzed by detection of marker proteins with antibodies specific for Aconitase (matrix), SOD2 (matrix), MRPL39 (peripheral protein) and Prohibtin (integral protein). **C:** The sub-mitochondrial localization of the mt-AspRS was analyzed with a specific antibody (see paragraph below).

The quality of sub fractionation was verified by western blot detection of marker proteins (**Figure 17** B) using antibodies, specific to proteins representative to each sub-fraction (SOD2 and Acconitase for matrix, Prohibitin for integral proteins, and mt-ribosomal protein MRPL39 for peripheral proteins). The marker proteins specific for each fraction are solely detected in the expected fraction. Of note, MRPL39 and Prohibitin are not detected in the M1 and M2 fraction but this is due to the high dilution factor of the mitochondria (10µl out of 400µl sample) compared to the sub-mitochondrial fraction (10 out of 100µl).

2.2 Establishment of the sub-mitochondrial localization of mt-AspRS

After validation of the sub mitochondrial fractionation protocol, our aim was to localize first our protein of interest, the mt-AspRs, within the separated fractions (**Figure 17** C). This analysis was performed by western blot detection using an antibody specific for mt-AspRS. We reproducibly, detected the mt-AspRS (out of >3 independent experiments) within all three fractions. The fact that the mt-AspRS was found both in the soluble and in the membrane-associated/embedded protein fractions raised the question about the mode of anchorage. As recalled in the introduction, a protein can be embedded or associated to the membrane following various ways. The crystallographic structure of the mt-AspRS was recently established (Neuenfeldt et al., 2013) but did not reveal any potential transmembrane domain or helices. However, we then used a software prediction tool, TMHMM Server v. 2., to test for the presence of hydrophobic sequences that could explain the anchorage to or inside the membrane (**Figure 18**). A known membrane protein, prohibitin, was used as a control. A well-defined N-terminal transmembrane region was predicted within the 40 N-terminal amino acid of prohibitin sequence (**Figure 18**A) with a probability >50%. Conversely, no similar region could be predicted for mt-AspRS (**Figure 18**B). It is thus concluded, that secondary structure elements are unlikely to be responsible for the localization.

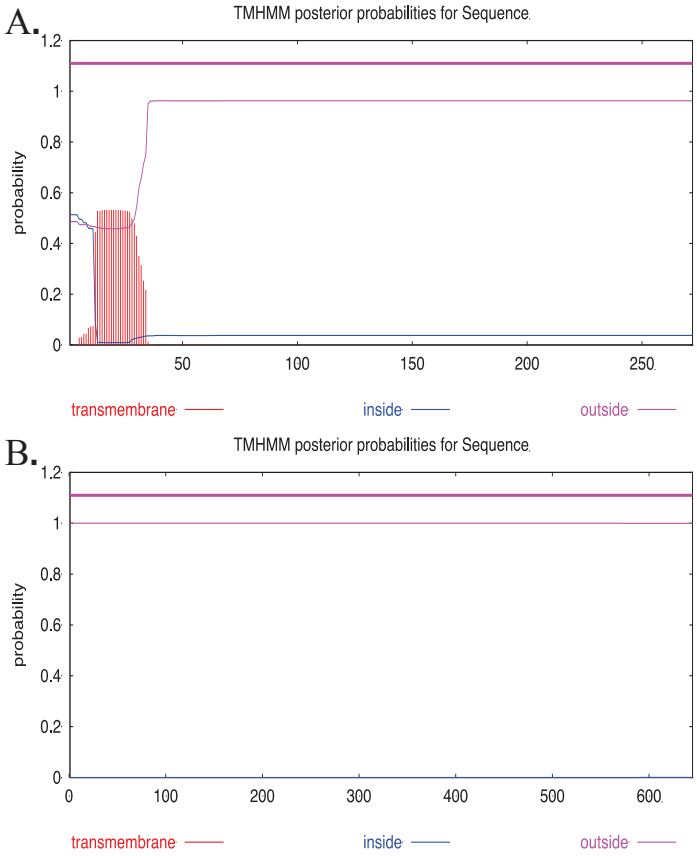


Figure 18: Transmembrane helices prediction using the TMHMM server v 2.0 (http://www.cbs.dtu.dk/services/TMHMM/). A) Full-length protein sequence of the Prohibitin (Uniprot:P35232). **B)** Full length protein sequence of mt-AspRS (Uniprot: **Q6PI48**).

2.3 Deciphering the nature of the interaction of mt-AspRS to or within the membrane

2.3.1 Sensitivity to alkaline treatment

The presence of a transmembrane domain/helix being excluded (see above), the biophysical properties of attachment of mt-AspRS was now tested. We know from the structural comparison of the mt-AspRS with *E. coli* homolog, that the mt-AspRS has a more electropositive surface potential (Neuenfeldt et al., 2013). The surface of the inner membrane is negatively charged, due to the phosphate groups at the hydrophilic heads of phospholipids. Thus, an electrostatic interaction between the electropositive surface of the mt-AspRS and the negative charged inner membrane is hypothesized. To test this, we performed a fractionation protocol that includes an alkaline treatment (Na_2CO_3

pH>11) to disturb all electrostatic interactions between protein:protein or protein:membrane. After alkaline treatment, a fraction of mt-AspRS is released in the soluble fraction, while a significant one remains in the membrane fraction (**Figure 17** C).

2.3.2 Sensitivity to high salt concentration and to chaotropic agents

Another way o investigate electrostatic interactions is to use a high concentration of salt (as described in Liu and Spremulli 2000) or a chaotropic agent to disturb ionic interactions. Thus, aliquots of mitochondrial total membrane fractions were treated with either high KCl concentration (250mM and 500mM) or chaotropic agent (6M urea) both independently of (**Figure 19**A), and following Na_2CO_3 treatment (**Figure 19**B). After each treatment, the soluble proteins were separated from the membrane protein fraction by ultracentrifugation. We observed that the mt-AspRS was not released from the membrane fraction by treatment with KCl, while part of it was released by treatment with 6M urea (**Figure 19**A). The same observations were made when samples were sequentially treated with NA_2CO_3 and KCl (**Figure 19**B).

These experiments indicated that the mt-AspRS is not anchored to the membrane by electrostatic/ionic interactions. A part of the mt-AspRS is sensitive to Na_2CO_3 treatment but major fraction of mt-AspRS remains inside the membrane. It has been reported that mitochondrial disruption by sonication could lead to membrane vesicles containing matrix proteins (Speers and Wu, 2007). These vesicles, if any, would co-purified with the membrane fraction and the trapped proteins would appear as membrane proteins. To consider this hypothesis, additional alkaline washing steps were combined with a brief sonication episodes so that to destroy possible vesicles. As shown in **Figure 19**C, no further mt-AspRS is released after such a treatment.

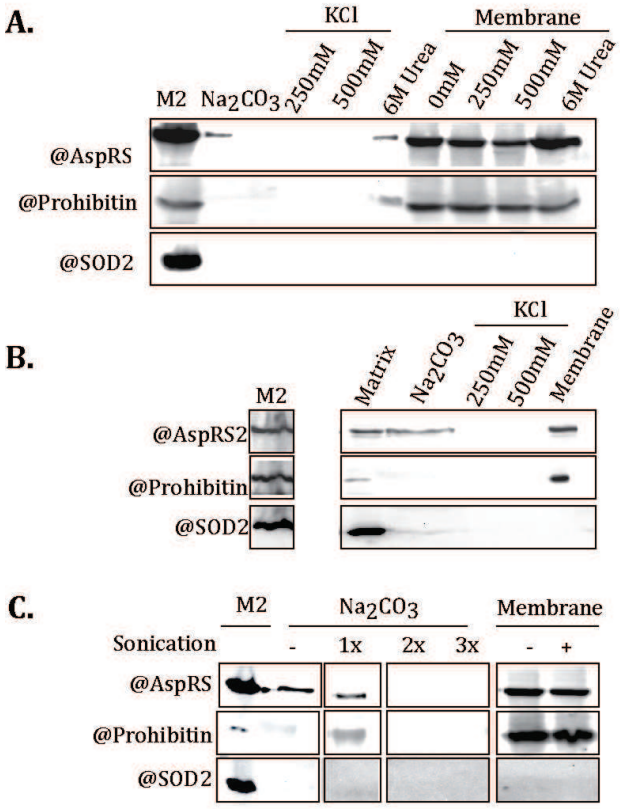


Figure 19: Sensitivity of the membrane bound mt-AspRS to different salt concentration and chaotropic agent. Samples of each fraction were loaded and analyzsed by western blot detection using specific antibodies against mt-AspRS (and prohibitin and SOD_2 as controls of the quality of the fractionation). A) Aliquots of mitochondrial membrane were treated with different concentration of salt. The supernatant containing soluble proteins were removed and remaining pellets were treated with protein loading dye. B) Mitochondrial membrane was sequencially treated with Na_2CO_3 and an increasing amount of KCl. Each time the supernatant containing the soluble proteins were removed. The remaining membrane pellet was treated with protein loading dye. C) Mitochondrial membrane was sequencially washed with Na_2CO_3. At each treatment membrane was solubilzed using sonication (indicated as 1x, 2x and 3x for the first, second and third time of washing respectively). Remaining membrane pellets were treated with protein loading dye.

2.3.3 Sensitivity to the non-ionic detergent digitonin

To further investigate the fraction of proteins found in the membrane we used digitonin titration to sequentially release the membrane-anchored proteins. Digitonin is a non-ionic detergent used for water solubilization of lipids and is widely used for solubilization mitochondrial membranes (Nabholz et al., 1999). To do so, we performed mitochondrial fractionation and separated matrix proteins from the membrane proteins

fraction. The membrane proteins fraction was treated sequentially with an increasing concentration of digitonin ranking from 0.05 to 2% (**Figure 20**).

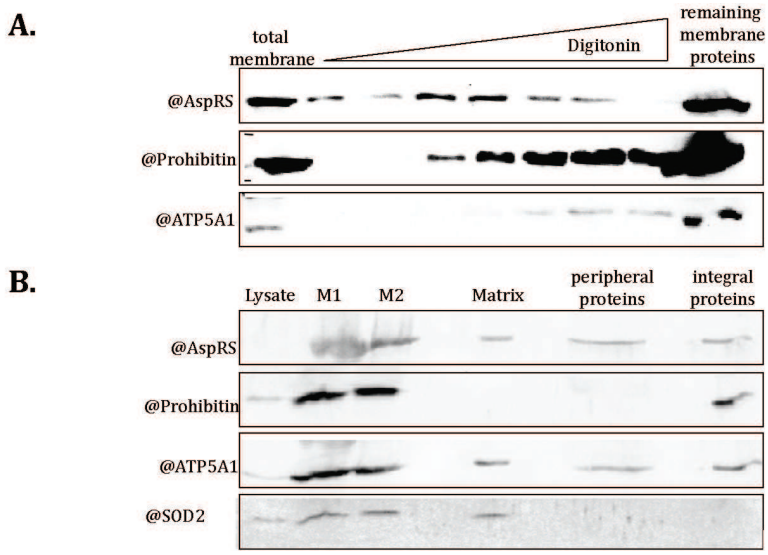


Figure 20: Digitonin treatment of mitochondrial membrane. Sequential treatment of total membrane with increasing concentrations of digitonin (0.05-2%). **A)** Aliquot of total membrane was treated whether with increasing amount of digitonin or **B)** with Na_2CO_3. Western blot detection of marker proteins with specific antibodies confirms the quality of fractionation.

After each treatment we collected the supernatant and submitted the pellet to the next treatment. In a final step, the remaining membrane was treated with SDS-Page loading dye and incubated at 95°C for 8min. All samples were separated on a 10% SDS-PAGE and proteins were transferred to a PVDF membrane and proteins were detected using specific antibodies. Prohibitin and ATP5A1 were used as fraction specific markers. Prohibitin is anchored to the membrane *via* its transmembrane N-terminal domain (Berger and Yaffe, 1998). ATP5A1 is part of the soluble catalytic F1 portion of the ATP synthase and is located at the mitochondrial inner membrane (matrix side) *via* an interaction with the membrane embedded proton channel F0. Both F1 and F0 form the complex V of the respiratory chain, also called ATP synthase.

The digitonin treatment was performed to further estimate the deepness of anchorage of mt-AspRS and to compare it with the control proteins. Results are displayed in **Figure 20**. We show that the mt-AspRS elutes from the membrane at lower digitonin concentration than the prohibitin and ATP5A1. Surprisingly, the prohibitin is

more sensitive to digitonin than ATP5A1, while the latter shows sensitivity to alkaline treatment and prohibitin does not. This implies that electrostatic interaction between two proteins can be really tackled by alkaline treatment but not by digitonin (since it is a non ionic detergent). On the other hand, our results suggest that mt-AspRS is attached to the membrane through an "anchor" interaction that is sensitive to both digitonin and alkaline treatment.

Of note, experiments performed until here were made with total mitochondrial membranes, meaning they contained both inner and outer membranes. This led us to hypothesize that the membrane fraction may contain mt-AspRS proteins that are in the process of being translocation through the membrane, thus contaminating the membrane fraction with mt-AspRS. In order to exclude this possibility we prepared mitoplasts (mitochondria without the outer membrane) by digitonin treatment of intact mitochondria.

2.3.4 Preparation of mitoplasts

Two common strategies to prepare mitoplasts are reported in the literature. Due to the different constitution of the mitochondrial outer membranes compared to the inner membranes, the outer membrane can be specifically removed by treatment with a specific amount of digitionin or by osmotic shock in a hypotonic buffer.

Both strategies were tried. First, aliquots of purified mitochondria were treated with different amount of digitonin (50, 100, 200, 500 µg digitonin per mg mitochondrial protein). Mitoplasts were harvested by centrifugation and subsequently fractionated into matrix, peripheral and membrane protein fractions. The localization of mt-AspRs, SOD2 (matrix protein) and VDAC (a porine from the outer membrane) were determined by western blot (**Figure 21**). However, we found that VDAC was still visible in the membrane fraction at all digitonin concentrations. Furthermore, SOD2 was released in the supernatant at a digtonin amount of >200µg digitonin per mg mitochondrial protein. This indicates that we could not optimize a compromised digitonin amount that would keep the inner membrane intact and be sufficient to remove the outer membrane. Nevertheless, the distribution of the mt-AspRS did not significantly change.

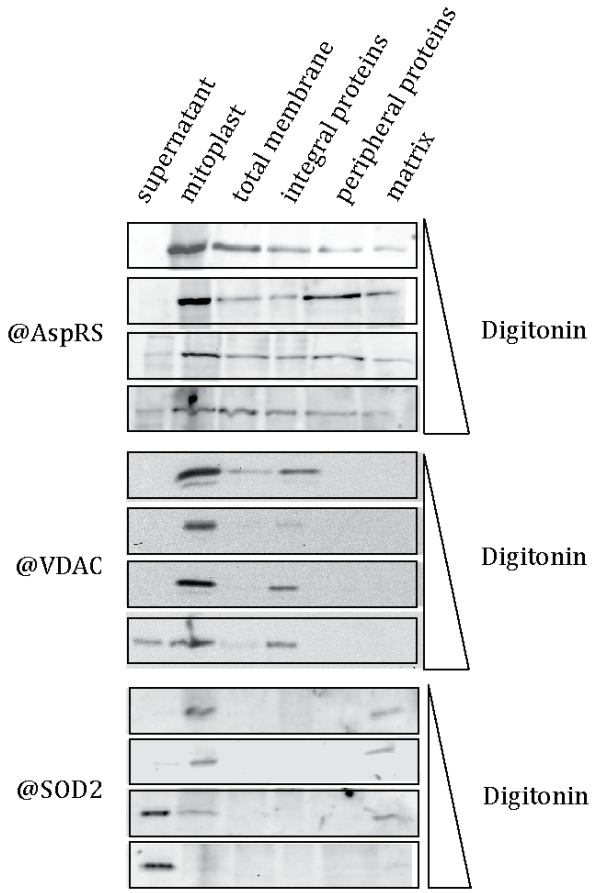


Figure 21: Mitoplast preparation using increasing amount of digitonin. Purified mitochondria were treated with 50µg, 100µg, 200µg and 500µg of digitonin per mg mitochondrial protein. Quality of mitoplast preparation and fractionation were confirmed using specific marker proteins. VDAC was used as marker for outer membrane and SOD2 for matrix.

The second method for mitoplast preparation was then tested. To do so, we incubated purified mitochondria in a hypotonic 10mM Tris-HCl buffer (pH=7,5). Mitoplasts were recovered by centrifugation, and fractionated as usual. The quality of the preparation was controlled by western blot detection using antibodies against prohibitin, SOD2 and cytochrome C (**Figure 22**). The latter one is known to be a soluble protein of the inter membrane space. We observed a small fraction of cytochrome C in the mitoplast before fractionation, but also in the matrix. This result indicates that the hypotonic treatment also impacts the inner membrane permeability causing the release of cytochrome C from the inter membrane space through the matrix. The total membrane fraction does not contain any cytochrome C indicating, that we have no cross contamination with inter membrane proteins, which are may trapped in membrane-enclosed vesicles. The other controls, SOD2 for matrix and prohibitin for

membrane indicate clean fractions. Unfortunately, the detection of VDAC in the mitochondria fraction failed (data not shown), which prevents us from concluding that the membrane fraction is free of outer membrane proteins. Nevertheless, the mt-AspRS was detected in matrix and membrane fraction indicating that membrane fraction does not contain proteins contaminations from the inter membrane space.

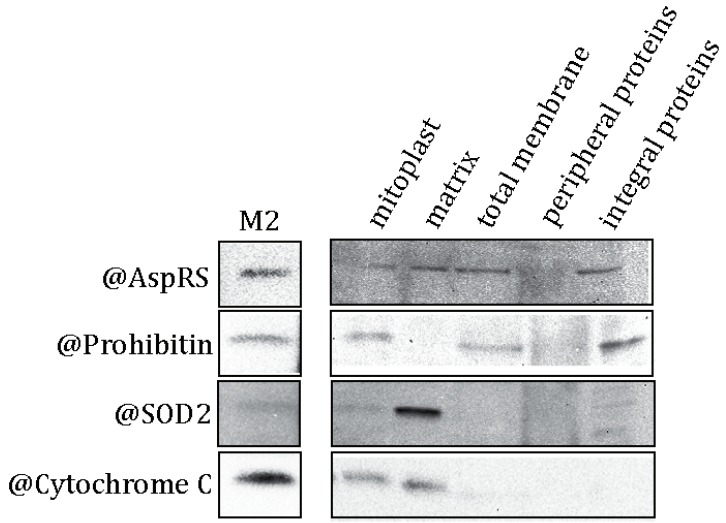


Figure 22: Mitoplast preparation using osmotic shock. Purified mitochondria were incubated in hypotonic solution. Obtained mitoplast were fractionated and the quality of fractionation was controlled by western blot detection of marker proteins with specific antibodies. Prohibitin is a marker for inner membrane, SOD2 is a marker for matrix and Cytochrome C is a marker for intermembrane space.

2.3.5 Summary

We demonstrated a double distribution of the mt-AspRS inside mitochondria. Surprisingly, we found a fraction of mt-AspRS anchored to the membrane. We characterized the anchorage by alkaline stripping, high salt and digitionin treatment. The sensitivity to alkaline stripping indicated that mt-AspRS interacts with the membrane in partially electrostatic manner. We excluded ionic interactions (shown by KCl treatment) and secondary structure elements as cause for the attachment to the membrane. In addition, we sequentially treated the membrane with digitonin and confirmed that the anchorage is not as deep as the anchorage of prohibitin. The different behavior of mt-AspRS and ATP5A1 suggests that the mt-AspRS does not interact via an integral protein. Our experiments with mitoplasts suggest that the membrane fraction of mt-AspRS do not contain proteins on the way to be imported. This is also supported

by the fact that we do not observe intermediate products for other proteins as e.g. SOD2. Altogether the list of possibilities explaining the localization at the membrane is quite short and will be discussed in the general discussion and perspectives at the end of this chapter.

3 Deciphering the distribution of the full set of mt-aaRSs inside mitochondria

The mt-AspRS is one out of 19 mitochondrial aaRSs. In order to achieve a complete picture of the sub mitochondrial localization of all aaRSs, we fractionated mitochondrial proteins, separated them using a precast TGX-Gel system (for the reason of time-saving and reproducibility of results) and transferred the proteins to PVDF membranes. We systematically analyzed the fraction using commercially available antibodies specific against each of the 19 mt-aaRSs. Each PVDF membrane was reused for a maximum of three detections. We search for each aaRS in at least three independent fractionation experiments obtained from independent mitochondria purifications. The summary of all western blot detections is displayed in **Table 3**. The confidence of the results, meaning the clear absence or presence of a signal, is displayed by a color code: Dark red and dark green represent the clear absence or the clear presence of a mt-aaRS in the fraction, respectively.

Protein	Prediction		Experimental data		Result
	Motif	Localization	Matrix	Membrane	
AlaRS	R-10	membrane	?		unclear
ArgRS	no prediction	no prediction			matrix/membrane
AsnRS	R-10	membrane			unclear
AspRS	R-10	membrane			matrix/membrane
CysRS	R-2	matrix			matrix
GluRS	no prediction	no prediction			matrix/membrane
GlyRS	R-10	membrane			matrix
HisRS	R-3	matrix			matrix/membrane
IleRs	R-10	membrane			matrix
LeuRS	R-10	membrane			matrix/membrane
LysRS	R-10	membrane			matrix/membrane

MetRS	R-10	membrane		matrix/unclear
PheRS	R-3	matrix		matrix/membrane
ProRS	R-10	membrane		matrix/membrane
SerRS	no predicition	no prediction		matrix/membrane
ThrRS	R-3	matrix		matrix/membrane
TrpRS	R-10	membrane		membrane
TyrRS	R-10	membrane		matrix/membrane
ValRS	R-2	matrix	?	matrix/unclear

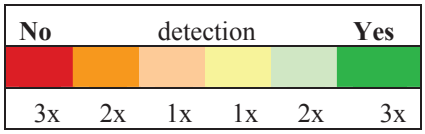


Table 3: Summary of experimentally determined and *in silico* predicted sub-mitochondrial distribution of aaRSs in membrane and matrix fractions. At least three independent fractionations were performed and the aaRSs were detected by western blot detection using specific antibodies. Color code indicates the presence or the absence of specific signals in western blot detection. Presence of a signal was counted as true if a significant signal of expected size was observed in the specific fraction and in at least in one other fraction (e.g. lysate or purified mitochondria). The absence of a signal was counted as true if a significant signal of expected size was observed in the other fraction (matrix or membrane) and in at least in one other fraction (e.g. lysate or purified mitochondria). In addition to experimental data, the theoretical distribution dependent from possible predicted motifs is given (prediction performed by Marie Messmer, Strasbourg). Red letters indicates clear discrepancies between prediction and experimentally determined localization.

Ten mt-aaRSs (AspRS, GluRS, HisRS, LeuRS, LysRS, PheRS, ProRS, SerRS, ThrRS and TyrRS) were clearly identified in both the matrix and in the membrane fraction. For three synthetases, we are confident about the absence of the protein whether in the matrix fraction (TrpRS) or in the membrane fraction (CysRS, IleRS). Three aaRSs (ArgRS, AsnRS and MetRS) were detected in both fractions, but with low confidence since the detection could not be reproduced. Also two aaRSs (GlyRS and ValRS) were clearly identified within one fraction (the matrix) but the presence or absence in the other one (the membrane) could not be clearly verified. For the mt-AlaRS, no clear result could be obtained (**Figure 23**). Obviously, no common rule for mt-aaRS distribution is emerging from our data.

The distribution pattern, established here, was then compared to a predicted distribution previously proposed by Marie Messmer, PhD, Strasbourg.

Box 1: Predicted MTS sequences in mt-aaRS (work was performed by Marie Messmer, Strasbourg, France)

R-2 motif: xRx⇓x(S/x)

mt-CysRS MLRTTRGPGLGPPLLQAALGLGRAGWHWPAGRA⇓ASGGRGRAWLQPTGRETGVQVYNSLTGRKEPLI
mt-ValRS MPHLPLASFRPPFWGLRHSRGLPRF⇓HSVSTQSEPHGSPISRRNREAKQKRLREKQATLEAEIAGESKSPA

R-3 motif: xRX(Y/x)⇓(S/A/x)x

mt-HisRS MPLLGLLPRRAWA⇓SLLSQLLRPPCASCTGAVRCQSQVAEAVLTSQLKAHQEKPNFIIKTPKGTRDLSPQH
mt-PheRS MVGSALRRGA⇓HAYVYLVSKASHISRGHQHQAWGSRPPAAECATQRAPGSVVELLGKSYPQDDHSNLT
mt-ThrRS MALYQRWRCLRLQGLQACRLHTAVVSTPPRWLA⇓ERLGLFEELWAAQVKRLASMAQKEPRTIKISLPGG

R-10 motif: xRx⇓(F/L/I)xx(S/T/G)xxxx⇓

mt-AlaRS MAASVAAAARRLRRA⇓IRRSPAWR⇓GLSHRPLSSEPPAAKASAVRAAFLNFFRDRHGHRLVPSASVRPR
mt-AsnRS MLGVRC⇓LLRSVRFC⇓SSAPFPKHKPSAKLSVRDALGAQNASGERIKIQGWIRSVRSQKEVLFLHVNDGS
mt-AspRS MYFPSWLSQLYRGLSRPIRRTTQPIWGSLYRS⇓LLQSSQRR⇓IPEFSSFVVRTNTCGELRSSHLGQEVTLCG
mt-GlyRS MLRGAGGARRFRRHPLWTAQGRR⇓LMPSPRPV⇓LLRGARAALLLLLPPRLLARPSLLLRRSLSAASCAPIS
mt-IleRS MRWGLRPRGPGAAALATARSLWGTPRLPCSPGWQGATKRL⇓LVRSVSGA⇓SNHQPNSNSGRYRDTVL
mt-LeuRS MASVWQRLGFYASLLKRQ⇓LNGGPDVI⇓KWERRVIPGCTRSIYSATGKWTKEYTLQTRKDVEKWWH
mt-LysRS MLTQAAVRLVRGSLRKTSWAEWGHRE⇓LRLGQLAP⇓FTAPHKDKSFSDQRSELKRRLKAEKKVAEKEAK
mt-MetRS MLRTSVLRL⇓LGRTGASR⇓LSLLEDFGPRYYSSGSLSAGDDACDVRAYFTTPIFYVNAAPHIGHLYSALLAD
mt-ProRS MEGLLTRCRALPALATCSRQLSGYVPCRFHHCAPRRGRR⇓LLLSRVFQ⇓PQNLREDRVLSLQDKSDDLTC
mt-TrpRS MALHSMRKARERWSFIRA⇓LHKGSAAA⇓PALQKDSKKRVFSGIQPTGILHLGNYLGAIESWVRLQDEYD
mt-TyrRS MAAPILRS⇓FSWGRWSG⇓TLNLSVLLPLGLRKAHSGAQGLLAAQKARGLFKDFFPETGTKIELPELFDRG

Without motif

mt-ArgRS MACGFRRAIACQLSRVLNLPPENLITSISAVPISQKEEVADFQLSVDSLLEKDNDHSRPDIQVQAKRLAE
mt-GluRS MAALLRRLLQRERPSAASGRPVGRREANLGTDAGVAVRVRFAPSPTGFLHLGGLRTALYNYIFAKKYQ
mt-SerRS MAASMARRLWPLLTRRGFRPRGGCISNDSPRRSFTTEKRNRNLLYEYAREGYSALPQLDIERFCACPEE

In previous work, the sub-mitochondrial localization was predicted on the basis of the presence of the possible cleavage motifs R-2, R-3 and R-10 (Gakh et al., 2002; Nett and Trumpower, 1996) within the MTS sequences. MTS with the R-10 motifs are e.g. found in proteins of inner membrane localization. It is cleaved sequentially by MPP in the mitochondrial matrix and by MIP in the mitochondrial inner membrane.

Motifs R-2 and R-3 are found in proteins localized in the mitochondrial matrix. We confirmed experimentally the predicted localization of the proteins that harboring the R-2 motifs (CysRS and ValRS) are solely localized in matrix proteins. All proteins harboring the R-3 motifs (ThrRS, HisRS and PheRS) were localized in the matrix, but also found in membrane fraction of the mitochondria. For 9 out of 11 aaRSs harboring the R-10, membrane localization was confirmed, while at the opposite the two remaining aaRSs (GlyRS and IleRS) were found only in the matrix. Three synthetases (ArgRS, GluRS, SerRS), for which no motifs could be predicted, were experimentally found both in the matrix and the membrane.

Thus, the prediction of the sub-mitochondrial localization based on solely the presence of putative sequence motifs appears to be too simplistic. The absence of a clear consensus sequences for cleavage sites, but also the discoveries of multiple proteolytic steps (in addition to MBB and MIP, Chacinska et al., 2009) render the prediction difficult. Experimental evidences, presented in the chapter III, indicate that the sequence outside MTS may be of importance for the final sub-mitochondrial location.

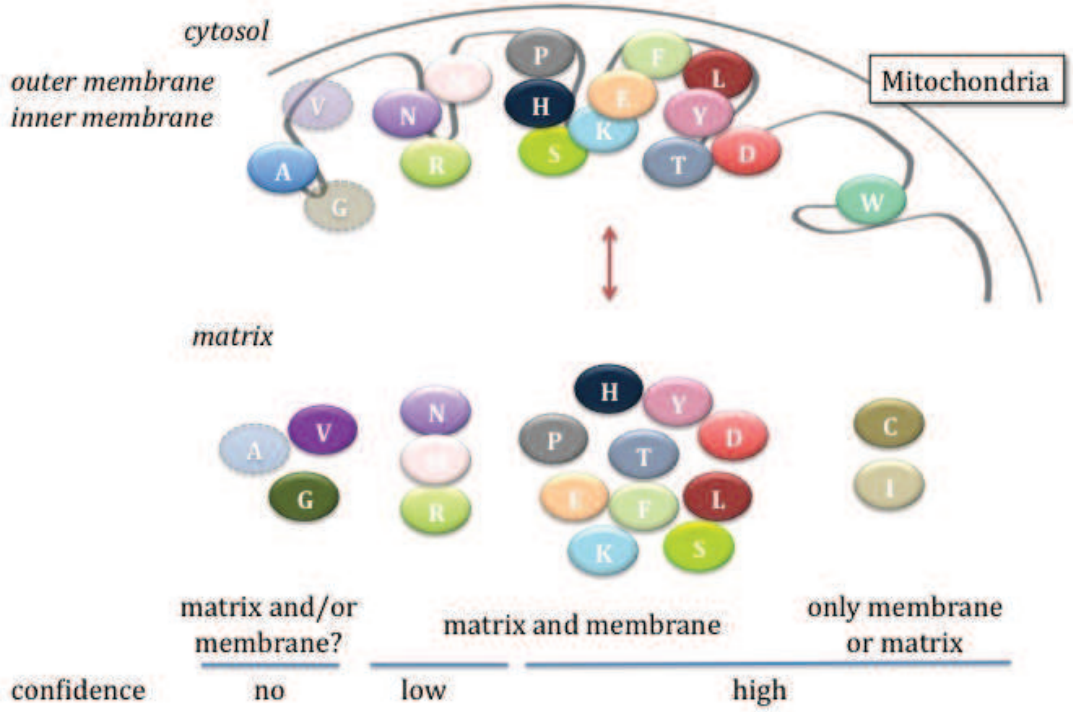


Figure 23: Schematized distribution of mt-aaRSs in matrix and/or along the membrane of mitochondria. AaRSs are displayed by the single letter code of the corresponding amino acid. The illustration of the aaRSs is ordered concerning the confidence about their localization. Dashed circle indicates that localization on membrane or matrix of the aaRSs cannot be excluded.

4 Sub-mitochondrial localization of tRNAs

In the previous paragraph, we established the wild type distribution of all mt-aaRSs. We discovered that many of them are double distributed within both the membrane and matrix.

Investigations into the cytosolic aaRSs have revealed that alternate locations could be associated with non-translational functions (see introduction). Following this

idea, our hypothesis is that mt-aaRSs situated in the matrix may not have the same function as those situated in the membrane. One possible way to test this hypothesis is to decipher the relative tRNA and aminoacylated-tRNA abundance in the sub-mitochondrial fractions and correlate this with translational function. To do so, we analyzed the distribution of tRNA within the different fractions by northern blot experiments. After mitochondria purification and fractionation, the total RNA was purified from aliquots of each fraction and quantified. The quantification indicated a ratio 17±3%/83±3% (n=3) for matrix RNA/membrane RNA. This repartition is in agreement with the presence of mito-ribosme (high RNA content) in the membrane fraction. The total RNAs from each fraction were loaded and separated on a 12% Polyacrylamide gel. RNAs were further transferred to a Zeta-Blot transfer membrane and UV-cross linked. The membrane was hybridized with 5'-radiolabeled DNA oligonucleotides complementary to the 3'-end of the tRNA. We sequentially analyzed 6 tRNAs ($tRNA^{Asn}$, $tRNA^{Asp}$, $tRNA^{Cys}$, $tRNA^{Leu}$UUR, $tRNA^{Lys}$ and $tRNA^{Glu}$) for their presence in the different fractions. A representative northern blot is displayed in **Figure 24**. The high quality of fractionation was confirmed by western blot detection of the marker proteins SOD2, VDAC and mt-AspRS. Transfer RNAs were observed in both the matrix and membrane fractions of mitochondria. However, while single bands were visible for $tRNA^{Asn}$, $tRNA^{Asp}$ and $tRNA^{Leu}$, multiple bands were visible for the three other tRNAs. These additional bands could be either artifacts (e.g. partly degradation), or could result from post-transcriptional hypermodifications (as e.g. in mt-$tRNA^{Glu}$ and mt-$tRNA^{Lys}$; (Umeda et al., 2005) that were known to alter migration. Thus, we speculate that the additional bands (of faster migration) may also represent modified tRNAs.

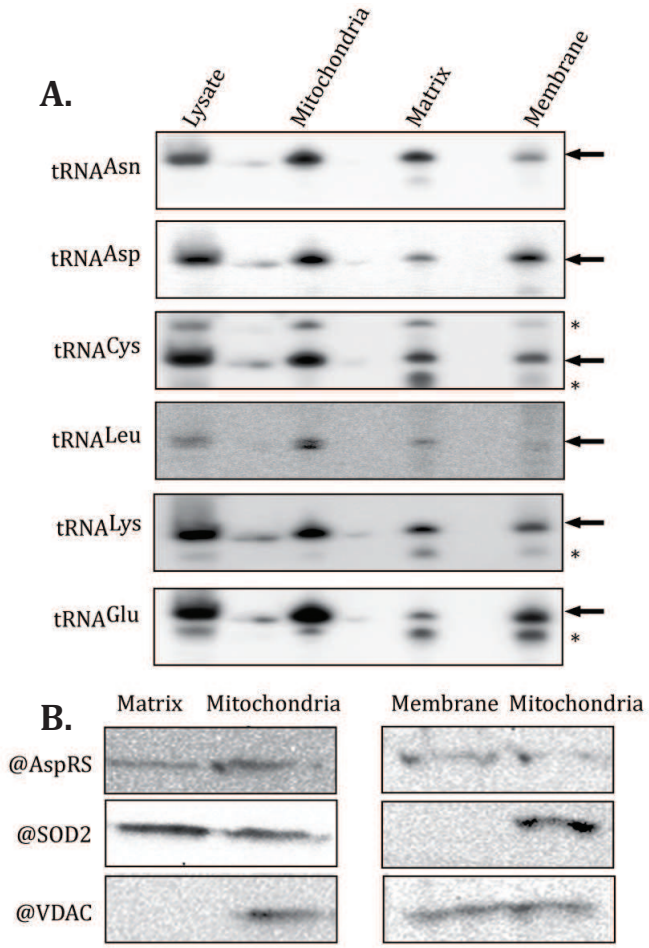


Figure 24: Distribution of tRNAs inside mitochondria. A) Total RNA of each fraction was isolated and loaded on a 12% PAGE. RNA was transferred to Zeta-Blot membrane and tRNA was detected with radiolabeled oligo-probes. An arrow indicates bands corresponding to tRNA. An asterisk indicated observed artifacts (e.g. putative modified tRNAs). **B)** Western blot detection of marker proteins shows quality of fractionation.

Northern blot experiments were reproduced three times and quantified using the ImageJ software package. **Figure 25** displays the mean values obtained from three experiments for $tRNA^{Asp}$, $tRNA^{Cys}$, $tRNA^{Asn}$ and $tRNA^{Lys}$. The amount of $tRNA^{Glu}$ was calculated for two quantifications. The quality of detection of the $tRNA^{Leu}$ was not sufficient for quantification. In general, we observed a high variability in the tRNA distribution (symbolized by the high deviation) from one experiment to the other.

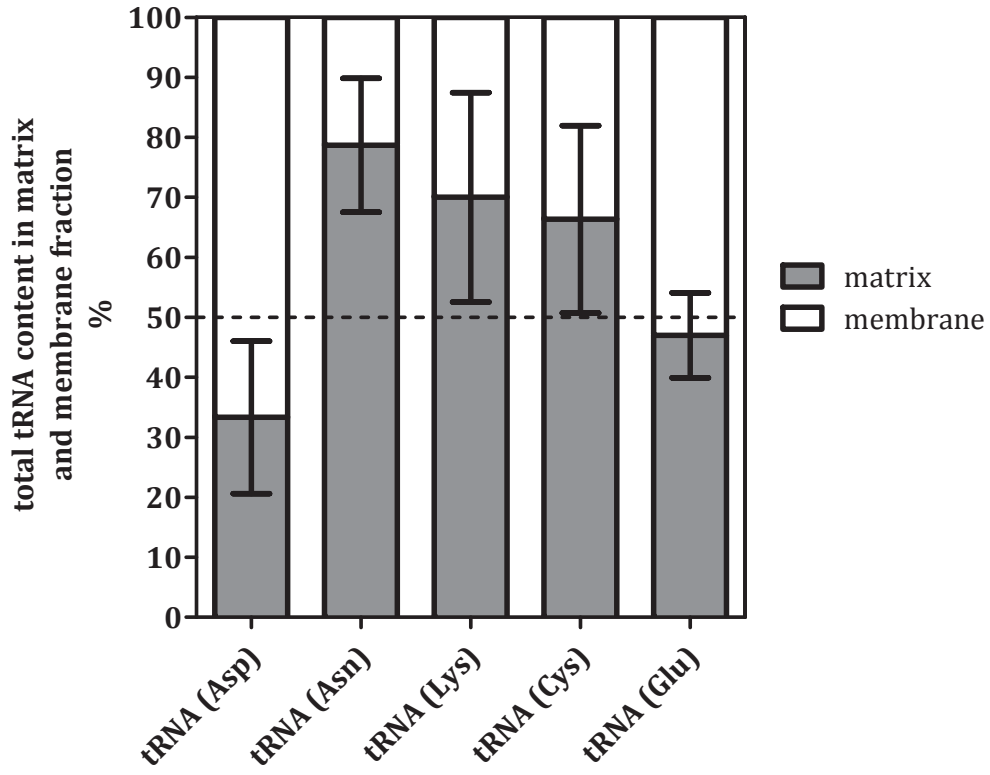


Figure 25: Quantification of tRNA amounts in membrane and matrix fractions. The tRNA amount in membrane and matrix was determined by quantification of the grey level of the corresponding tRNA band in northern blot detection. The relative distribution was calculated and is displayed as percentage of tRNA per fraction. Arrow bars indicate the variation in three independent experiments.

This large variability may arise from technical parameters (e.g. sample preparation, northern detection) or biological parameters (e.g. translational activity of mitochondria, state of cells, culture condition). Altogether, we see only a mean picture of the distribution of the tRNAs. Nevertheless, some tendencies are seen: $tRNA^{Asp}$ is more present in membrane than in matrix whereas $tRNA^{Asn}$ is more present in matrix than in the membrane. At the contrary, no statistically relevant differences are observed for $tRNA^{Glu}$, $tRNA^{Lys}$ or $tRNA^{Cys}$, despite a slight higher abundance in the matrix for the two latter.

These results correlate with the sub-mitochondrial distribution of the corresponding aaRSs. For example, $tRNA^{Cys}$ was mostly found in the matrix, as was the mt-CysRS. In addition, it is interesting to note that the $tRNA^{Asp}$ is significantly enriched at the mitochondrial inner membrane, while other tested tRNAs are enriched in the matrix.

Thus, tRNAs may not be distributed randomly within mitochondria. Our results are however too preliminary to figure out possible rules for the distribution. The set of experiments should also include the detection of tRNAs, which correspond to the aaRSs that are not dual localized (as e.g. $tRNA^{Ile}$ or $tRNA^{Trp}$). As already mention, it would be interesting to investigate the level of aminoacylated-tRNAs in the different fractions. This investigation should give further clues regarding the functional implication (translational or non-translational) of the aaRSs in the different fraction. In addition to the characterization of the levels of mt-aaRSs and tRNAs in the different sub-mitochondrial fractions, we were also interested in the proteins partners of the mt-aaRSs. This work is fully described in chapter 4.

5 Preliminary analysis of the sub-mitochondrial proteome

The total human mitochondrial proteome has been estimated to be ~1500 proteins (Pagliarini et al., 2008). Since we obtained fractions of the mitochondrial proteome, we decided to analyze these fractions by mass spectrometry:

I. To further confirm the distribution of the full set of mitochondrial aaRSs
II. To further estimate the quality of the fractionation methods and/or the quality of mitoplast preparation
III. To analyze if the double localization is a common feature in mitochondria, or not.

5.1 Proteomic analysis of the sub-mitochondrial localization of aaRSs

Purified mitochondria were either kept whole or fractionated into matrix and membrane protein fractions. The total protein content of all was analyzed by LC-MS/MS (10µg of total protein extract were *in-liquid* tryptic digested). We identified 1548 proteins in mitochondria, 1447 proteins in matrix and 1595 proteins in membrane. The number of proteins in mitochondria is in the expected range, but the number of identified proteins in the matrix and membrane fraction raises questions. It was

expected that the amount of identified proteins in matrix and membrane would approximate a summation of the total protein in mitochondria. Such high amounts of proteins were not expected and may be accounted by either a frequent double distribution or by contamination. In addition, abundant proteins, which may mask the detection of low abundant proteins in mitochondria, are fractionated in matrix or membrane and may allow the detection of the less abundant protein in one of the fractions. To take these facts into account, we refined our obtained data set.

The first aim of the mass spectrometry of sub-mitochondrial fractions was to confirm the distribution scheme of mt-aaRs established by western blot. The sequence coverage of all detected mt-aaRSs in the three samples (mitochondria, membrane, matrix) is given in **Figure 26**.

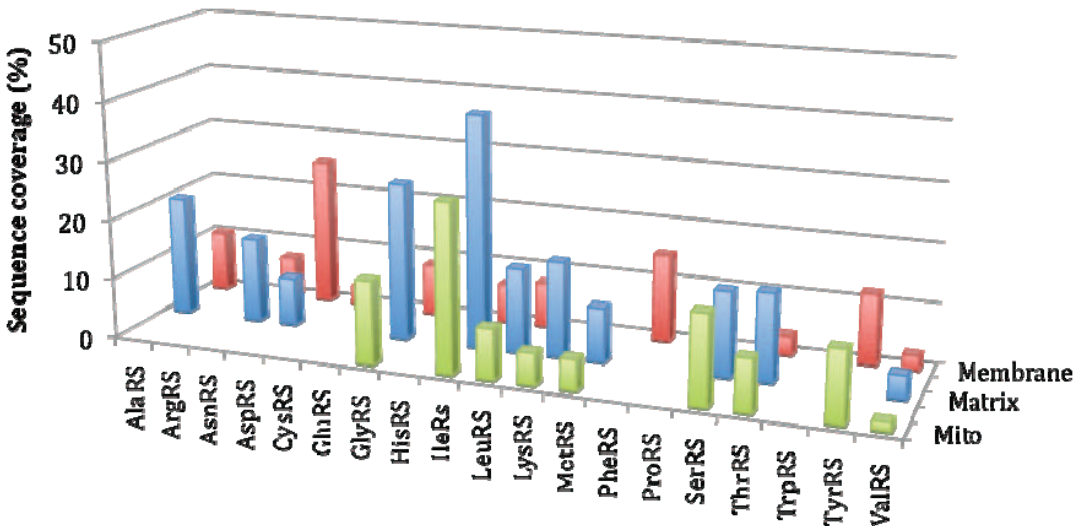


Figure 26: Proteomic analysis of the sub-mitochondrial distribution of aaRSs. The sequence coverage (%) obtained by mass spectrometry is displayed for each aaRSs depending on the sub-mitochondrial fraction.

Since we loaded similar amounts of proteins in each lane, the distribution of the aaRSs between different fractions is comparable. We found that the pattern of distributions approximates the one established by western blot detection. One advantage of the proteomic approach is that it also allowed us to analyze the ratio of the aaRSs between the fractions. We found eight mt-aaRSs (AlaRS, AsnRS, GlyRS, IleRS, LeuRS, LysRS, MetRS, SerRS and ThrRS) enriched in matrix and three mt-aaRSs

(AspRS, PheRS and TyrRS) enriched in the membrane. Five mt-aaRSs (ArgRS, GluRS, HisRS, ProRS and TrpRS) could not be identified by mass spectrometry. Although globally, we observe a similar pattern by western and mass spectrometry, some discrepancies need to be further investigated. First, some synthetases were not identified by mass spectrometry despite their identification by western blot (e.g. TrpRS, ProRS). Second, some aaRS were identified by mass spectrometry only in one fraction (TyrRS or PheRS) while we observed signals in two fractions by western blot analysis.

5.2 Advanced proteomic analysis of fractionated mitochondria

As mentioned previously, the number of proteins found by mass spectrometry raised some questions. To answer these questions, we set up a Venn diagram to illustrate number of shared proteins of each fraction (**Figure 27**) For sake of simplicity, all proteins found solely in matrix or membrane (and not in the mitochondria) were not further considered but we are aware that ignoring them might introduce a bias in the interpretation. These proteins might be contaminants (from cytosol or nucleus) or the might correspond to undetected proteins in the mitochondria sample (most abundant proteins may mask less abundant one that are unveil only after fractionation).

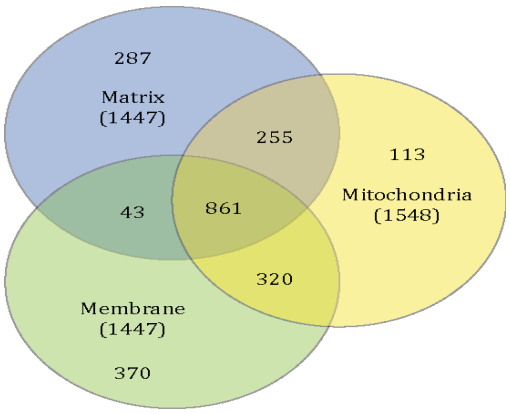


Figure 27: Venn diagram of identified proteins by LC-MS/MS in different sub mitochondrial fractions. Raw data were refined by identifying redundant proteins in mitochondria/membrane, mitochondria/matrix and mitochondria/membrane/matrix fraction.

Only 37% of the 1436 identified proteins were mitochondrial annotated using the GeneOntology database (Ashburner et al., 2000). Proteins were further sorted according their sub-mitochondrial location (**Figure 28**). As result, only 65 (34%) of the proteins in the matrix fraction, 120 (59%) of proteins in the membrane fraction and 205 (31%) of the proteins in the membrane/matrix fraction were annotated as mitochondrial proteins.

With the experiment performed we cannot decipher between false-positive identified proteins (cytosolic contamination) and false-negative annotated proteins. However, despite the large numbers of proteins that were not annotated of mitochondrial origin, we saw a clear enrichment of matrix proteins in matrix fraction and membrane proteins in membrane fraction. In addition, 15% of matrix annotated proteins, 38% of the membrane-annotated proteins were found in both fractions, 35% were not annotated for certain sub mitochondrial localization before

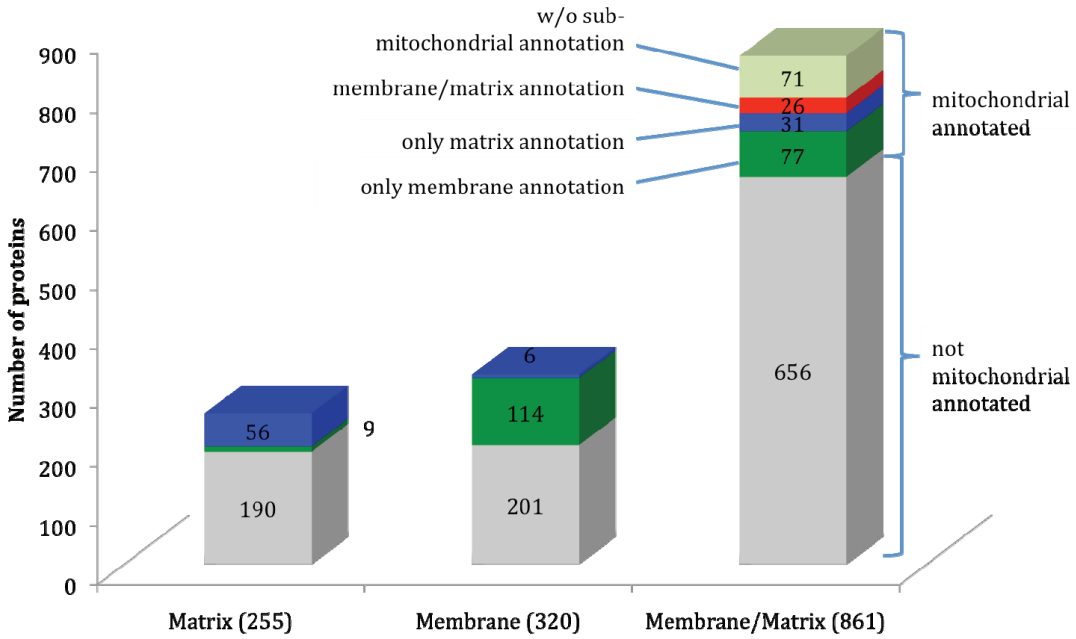


Figure 28: Mitochondrial annotation of identified proteins. Proteins found in the intersection mitochondria/membrane, mitochondria/matrix and mitochondria/membrane/matrix were annotated using GeneOntolgy database. Proteins are sorted according their annotation. Mitochondrial annotated proteins were further characterized concerning their known sub-mitochondrial localization. Numbers in brackets are the complete number of identified proteins.

And indeed, the marker proteins chosen for the quality control by western blot detection were found exclusively in matrix (SOD2) and membrane (VDAC). Of note,

prohibitin has been found in both fractions. This was expected, because we always observed a slight cross contamination within the matrix fraction.

Altogether, these results indicate that the localization of many proteins still remains unclear but analyses validate the quality of the sub-mitochondrial fractionation.

5.3 Proteomic analysis of mitoplast fractionation

As mentioned before, and as confirmed in the set of data presented in the previous paragraph, “purified” mitochondria are significantly contaminated by cytosolic proteins. Also the inaccuracies of annotation make impossible the distinction between false-positive and false-negative proteins. We decide thus to apply the mass spectrometry detection on fractions obtained from mitoplasts. The mode of preparation so far did not allow the production of pure mitoplast fully deprived of outer membrane proteins.

Nevertheless, the mitoplast fractions were analyzed with the aim of functionally clustering of the identified proteins. To do so, 10 µg of proteins (from whole mitoplast and mitoplast fractions) were loaded on a pre-casted TGX-gel and subjected to systematic analysis by mass spectrometry. A significant lower number of proteins were identified from each sample (640 mitoplast proteins, 520 matrix proteins and 400 peripheral proteins). In contrast, 1200 proteins were identified in the integral protein fraction. Of note, for technical reasons we loaded 40% more proteins of the integral fraction. Data were clustered using the **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery (**DAVID**) v6.7 (**Figure 29**) allowing the clustering of proteins according to function, localization, and involvement in a given pathway. We observed an enrichment of the matrix proteins in matrix, and an enrichment of membrane proteins in the membrane. Proteins identified in the peripheral fraction were clustered as whether as membrane or matrix proteins. Of note, the enrichment score is the overall enrichment based on the EASE (Expression Analysis Systematic Explorer) score of each term member (Huang et al., 2009a, 2009b). The higher the enrichment score, the more proteins can be dedicated to one term (e.g. mitochondrial matrix).

Next, we took the enriched proteins extracted from the biggest clusters (asterisk) obtained in **Figure 29** and re-cluster these proteins against their functional annotation (**Figure 30**).

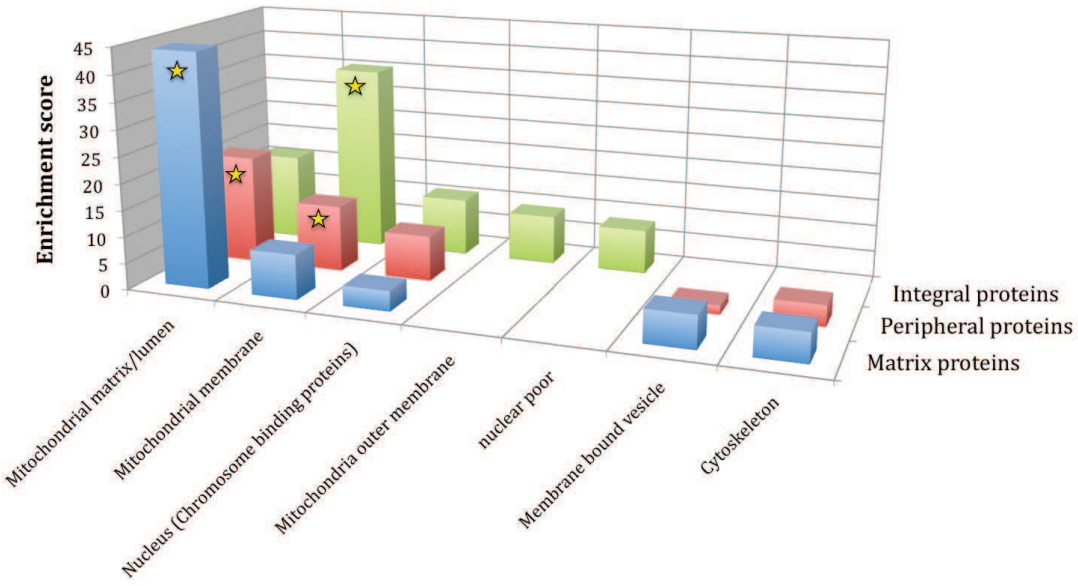


Figure 29: Clustering of proteins identified in matrix, peripheral and inner membrane fraction in regard to their sub-mitochondrial localization. The seven clusters with the highest enrichment factor are displayed. Clustering was performed using DAVID ABCC analysis software. Yellow asterisks indicate the highest enriched clusters. Of note, for subsequent analysis of the peripheral protein fraction the two highest enriched clusters were used

We observed a functional separation of the proteins in the different fractions suggesting: For example, membrane transporter or proteins of the respiratory chain were enriched in the membrane (peripheral and integral protein fraction) whilst proteins having a chaperone activity or involved in amino acid metabolism are enriched in the matrix. Interestingly, a proteins with aaRSs activity was enriched in all three fractions (matrix, peripheral and integral). Thus, we can assume that the particular distribution of the aaRSs in membrane and matrix fraction is not an artifact of western blot detection or fractionation.

In addition, this result offers new insight into the function of peripheral bound proteins, which were found to be mainly involved in RNA splicing/binding, DNA repair and citrate cycle. These results are in agreement with previous reports suggesting the mt-DNA (Spelbrink 2010) and mitochondrial ribosomes are bound to the inner

membrane (Suzuki et al.,2007) and also in agreement with the fact that the citrate cycle is closely connected to the respiratory chain complex.

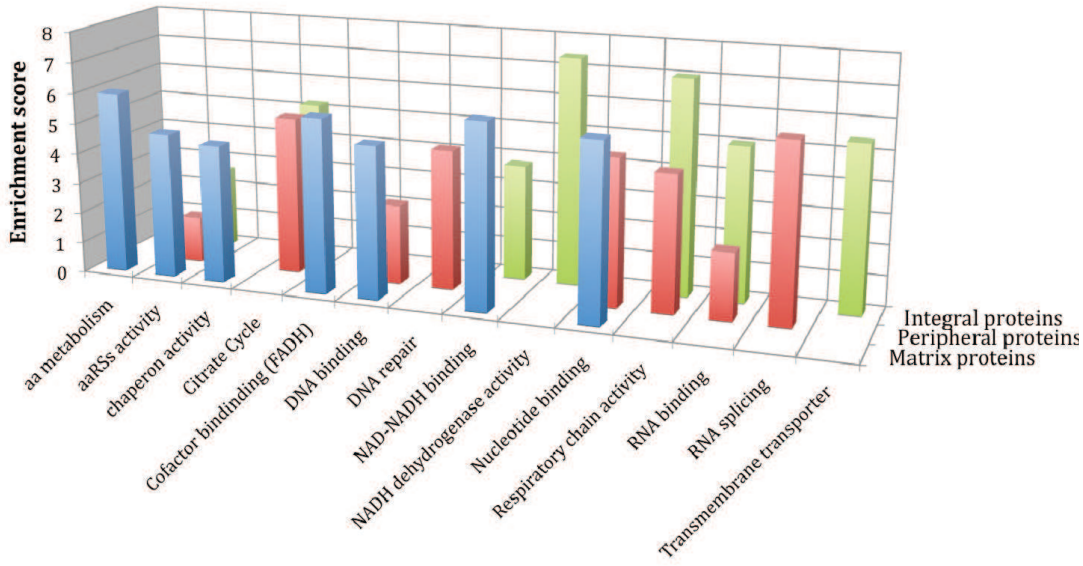


Figure 30 Functional clustering of proteins found in different sub-mitochondrial fractions. The highest enriched protein clusters of each fraction were used for the functional clustering (see **Figure 29**; yellow asterisk). The first seven clusters with the highest enrichment factor were chosen for each fraction and are displayed.

5.4 Future power of the proteomic analysis of the sub-mitochondrial fractions

In summary, the proteomic analysis of the sub-mitochondrial fractions seems to be a powerful tool to investigate the distribution of a family of proteins and to further correlate location with functions inside mitochondria. Despite the fact that our experiments remain preliminary, we can however immediately extract some technical and scientific benefits:

- We confirm the distribution of the aaRSs within the different fractions as a similar distribution of the aaRSs was observed by mass spectrometry as by western blot analysis.
- We confirmed the quality of fractionation in terms of the distribution of membrane and matrix proteins.

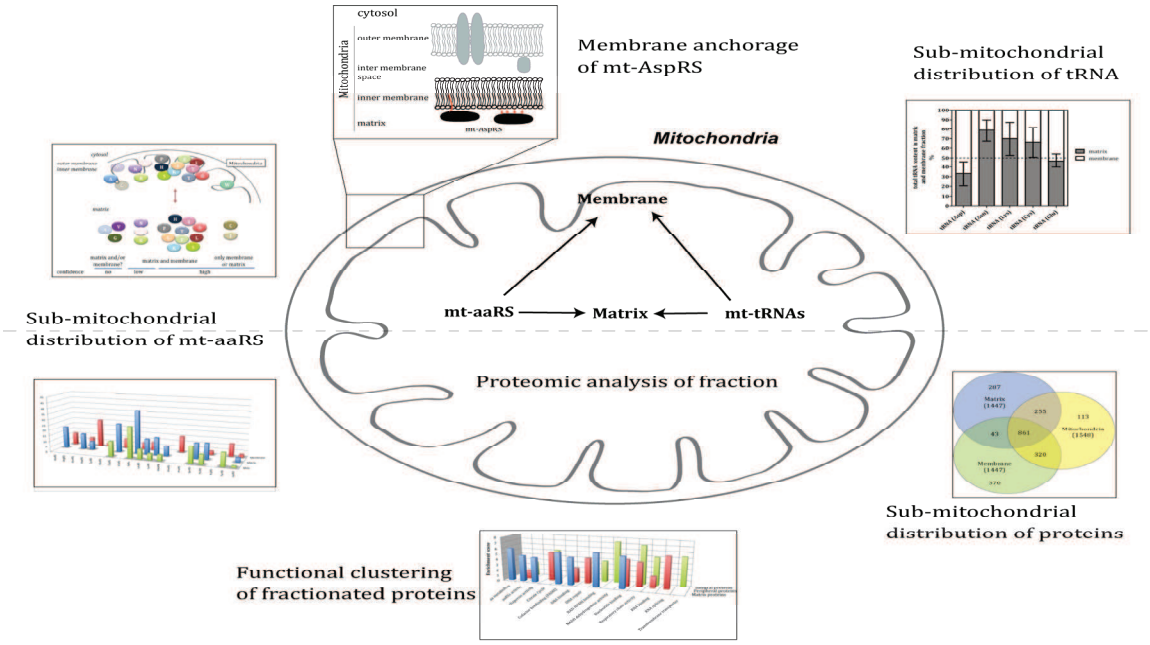
- It appears also that the double distribution of the mt-aaRS in matrix and membrane is, so far analyzed, specific to this family of proteins
- The *in-liquid* tryptic digestion method offers a powerful and sensitive tool to simultaneously identify a large number of proteins.
- The major difficulty we faced was the quality and purity of mitochondria. Significant numbers of cytosolic proteins were detected and it remains difficult to distinguish whether they are true contaminants or whether they have been wrongly annotated.

All of those technical and analytical drawbacks need to be solved so that to successfully reassign the mitochondrial proteome into the constitutive fractions. Of note, among the 861 proteins identified in the matrix, membrane, mitochondria intersection of the Venn diagram, only 205 proteins were annotated with a mitochondrial localization and only 26 proteins were annotated as membrane and matrix protein. For a valid interpretation the experiments need to be reproduced so that a statistical analysis is possible. For future experiments, it should be of interest to develop a well-defined pipeline to facilitate comparisons between experiments. This pipeline should include a optimized and adapted mitochondrial purification protocol, comparable LC-MS/MS analysis conditions, and a strategy to refine the raw dataset to achieve cleaned protein list of each fraction. Once this is established, comparative mitochondrial proteome analysis could be performed for many experimental conditions, such as between different cell lines or wild type *versus* pathology-related mitochondrial proteome.

6 Achievements at a glance

- We reveal a wild-type picture about the distribution of the full set of mt-aaRS in mitochondria. Some are double distributed in membrane/matrix and some are solely in matrix or membrane
- We investigate the membrane-anchorage of mt-AspRS, one of the double distributed mt-aaRSs.
- We systematic analysis the tRNA content in the sub-mitochondrial fractions for a subset of tRNAs indicating that they are not randomly distributed
- We analyze the sub-mitochondrial proteome of mitochondria and mitoplast and suggest a pipeline to further investigate the sub-mitochondrial proteome

7 Graphical summary



Chapter III: Towards the definition of N-terminal maturation site of mt-aaRSs

1 Introduction

As already mentioned at several instances, a challenging task is to understand the constitutive proteolytic cleavage of N-terminal MTS occurring after the internalization of the precursor protein into mitochondria, and its influence on the sub mitochondrial distribution of proteins. Several mitochondrial peptidases have been identified so far, such as the mitochondrial processing peptidase (MPP) that removes N-terminal MTSs (e.g. Gakh et al., 2002). Different additional cleavages can occur subsequently to remove newly exposed N-terminal hydrophobic sorting peptides (inner membrane peptidase IMP) or destabilizing residues (intermediated cleaving peptidases Icp55) (Vögtle et al., 2009). Several experiments demonstrated that a precise N-terminal definition of mitochondrial proteins is important for the expression, stability, solubility or activity of recombinant proteins into host (bacterial) strain (Gaudry et al., 2012). The following chapter is mainly dedicated to article #2, in which we validated and exploited a method, allowing for the experimental determination of the N-terminal processing site of imported mitochondrial proteins. In additional paragraphs we present our studies made to decipher some rules underlying the sub mitochondrial localization of aaRSs. We wish to understand the influence of physical-chemical properties of aaRSs on the sub-mitochondrial localization and of the MTS sequence on the sub-mitochondrial localization of the aaRSs

2 Engineered expression system for naturally maturated mammalian proteins upon mitochondrial import

In the previous chapter we analyzed the wild type distribution of mt-aaRSs and the protein content of the sub-mitochondrial proteome. The first approach (by western blot detection) was highly informative but required a full set of antibodies, specific for each of the aaRSs. The second approach (by mass spectrometry) does not require specific reagents for each aaRSs, but still has some technical and analytical limitations. Also, the major peculiarity of proteins of mitochondrial location is they are translated in the cytosol, but their site of action is in the mitochondria. As recalled in the introduction, after synthesis, they are imported to and maturated inside the mitochondria, before reaching a final sub-mitochondrial location. The discovery of several maturation processes renders difficult the prediction of maturation sites. This has implications for the design and easy handling of the corresponding recombinant proteins (e.g. Gaudry et al., 2012). We thus decided to establish a tool that will allow the expression of any sequence for the investigation of the impact of precursor sequences on the import process, with the ultimate aim of producing large quantities of naturally processed mitochondrial proteins.

The innate ability of the Modified Vaccinia Virus Ankara (MVA-T7) to co-internalize a vector into mammalian cells was previously demonstrated in the laboratory. With this method, sufficient amounts of protein can be produced without using any further costly transfection reagents or without the production of time consuming recombinant virus (Jester et al. 2011). We used this method as dedicated tool to express mitochondrial proteins. We validated this method, in a comparative analysis between hamster (BHK21) and human (HEK293) cells, for the correct translocation, internalization and localization into mitochondria of the ectopically expressed human mitochondrial precursor proteins in BHK21 cells. We also applied this technique to resolve an important biological question by deciphering the yet unknown processing site of the mt-AspRS. Thus, we propose a powerful protein expression tool, which mimics the natural occurring processing steps followed by mitochondrial proteins and combines controlled bacterial expression (T7-Polymerase, affinity tag purification) with the processing and

post-translational maturation properties of eukaryotic cells. This study is provided in article #2. In addition to the results presented in the article, we will also discuss the impact of the N-terminal sequence of a protein on its overall pI and will present data suggesting that information outside the MTS is likely required to explain sub-mitochondrial localization of aaRSs.

3 Article #2

Expression system for naturally maturated mammalian proteins imported into mitochondria

Brian C. Jester †, Hagen Schwenzer†, Laurianne Kuhn, Philippe Hammann, Laurence Maréchal-Drouard, Catherine Florentz and Marie Sissler

Submitted

† Those authors contributed equally to the work

Expression system for naturally maturated mammalian proteins imported into mitochondria

Brian C. Jester†[#1], Hagen Schwenzer†[1], Lauriane Kuhn[1], Philippe Hammann[1], Laurence Maréchal-Drouard[2], Catherine Florentz[1] and Marie Sissler[1]*

[1] Architecture et Réactivité de l'ARN, UPR 9002, CNRS, Université de Strasbourg, IBMC, 15 rue René Descartes, F-67084 Strasbourg, France.

[2] Institut de Biologie Moléculaire des Plantes, UPR 2357, CNRS, Universtié de Strasbourg, 12 rue du Général Zimmer, F-67084 Strasbourg, France.

*to whom correspondence should be addressed: Marie Sissler, IBMC - 15 rue René Descartes, 67084 Strasbourg Cedex, France; Phone: 33 (0)3 88 41 70 62; Fax: 33 (0)3 88 60 22 18; e-mail: M.Sissler@ibmc-cnrs.unistra.fr

† these two authors contributed equally to this work

[#] Present address: Institut de Biologie Systémique et Synthétique (CNRS UPS3509); Genopole®, Genopole Campus 1 – Genavenir 6, 5 rue Henri Desbruères – 91030 EVRY Cedex, France

List of non-standard abbreviations: aaRS: aminoacyl-tRNA synthetase (specificity is indicated by the name of the amino acid abbreviated in a three-letter code, transferred to the cognate tRNA, *e.g.* AspRS stands for aspartyl-tRNA synthetase), MTS: mitochondrial targeting sequence.

Abstract

Most mitochondrial proteins are encoded by the nuclear genome, translated as precursor in the cytosol and imported into mitochondria following a process not fully deciphered in mammals. Studies of mitochondrial processes have been hindered due to the fact that mitochondrial proteins are notoriously difficult to produce. We developed an expression system to simplify the recombinant production and characterization of any mammalian mitochondrial protein. This unique tool allows for the exploration of the impact of precursor sequences on the import process and the access to naturally processed mitochondrial proteins. We demonstrate that expression of human mitochondrial proteins in a heterologous virus (MVA)-infected hamster cell-line does not alter the endogenous properties of targeting, internalization and localization. The validated MVA-mediated transfer system has a wide range of applications and here we apply it to characterize nucleus-encoded mammalian mitochondrial proteins. This powerful tool will boost the analysis of mutations in proteins causing mitochondrial pathogenesis.

Introduction

Mitochondria are the powerhouse of eukaryotic cells. They host essential metabolic pathways and harbor their own genome (mt-DNA), which codes in human for only 13 proteins. Thus, the vast majority of proteins (>99%) necessary for mitochondrial (mt) biogenesis and functions (estimated number ~1,500 [1]) is encoded by the nucleus, expressed as precursor proteins in the cytosol, and targeted to the mitochondria (reviewed in [2]). Precursor proteins are subsequently translocated through the outer mitochondrial membrane (*via* TOM complex, the mitochondria entry gate) and, depending on sorting signals, further internalized towards the matrix through the translocase of the inner membrane (TIM), or alternatively integrated into the inner or the outer membrane, or released in the inter-membrane space. More than four different pathways of importation of precursor proteins are presently reported (reviewed in [3-5]).

The sorting signals include mitochondrial targeting sequences (MTSs), either in a form of an internal signal that is non-cleavable and remains in the mature protein, or as a N-terminal pre-sequence that is proteolytically removed after import (reviewed in [6]). Among the identified peptidases, the mitochondrial processing peptidase (MPP) cleaves pre-sequences of most incoming precursor proteins [7]. Additional cleavages can subsequently occur, which remove the newly exposed N-terminal hydrophobic sorting peptides (inner membrane peptidase, IMP; mitochondrial intermediate peptidase, MIP [3,8]) or destabilizing residues (intermediate cleaving peptidases, Icp55 [9]). Although several

peptidases have been identified, deciphering where proteolytic cleavage actually occurs within the mitochondrial precursor protein still remains a challenging task. Some studies performed have decoded the cleavable motifs recognized by the MPP [7] and the constitutive proteolytic events have been profiled *in vivo* [10]. However, the absence of an absolute consensus sequence for the cleavage sites and the discovery of possible multiple proteolytic steps renders *in silico* predictions difficult. Consequently, the identification of N-terminal amino acids of naturally cleaved mitochondrial proteins by Edman degradation or mass spectrometric methods although straightforward are technically challenging. Precise determination of the N-terminal amino acid has been performed only for a few proteins of the human mitochondrial proteome. To understand the biogenesis of mitochondria *in vivo* and to design sequences compatible with efficient expression in bacterial strains, the identification of the N-terminal amino acid for a mitochondrial protein is essential [11-14]. This remains of interest for the further characterization of the functional and structural properties of mitochondrial proteins.

Some *in vivo* and *in vitro* approaches are currently available to characterize mitochondrial proteins. These techniques allow researchers to decipher several aspects of the protein-mitochondria interaction: (i) protein targeting to the mitochondria (fluorescent fusion protein [16]); (ii) protein internalization and/or the maturation of precursor proteins within the mitochondria (*in vitro* import assays [15,16]); (iii) identification of components responsible for protein import (Blue Native gel electrophoresis [17]); (iv) definition of the sub-mitochondrial localization of imported proteins (sub-mitochondrial fractionation [18]): and (v) determination of the maturation site of mitochondrial proteins (N-terminal sequencing by Edman degradation [19]). A landmark global study using GFP fusion proteins has classified ~75% of the yeast proteome into 22 distinct sub-cellular localizations [20]. The total mitochondrial proteome was further analyzed in yeast [21], and later on in mammals [1]. The determination of the mature N-termini of the majority of the known yeast mitochondrial proteins (named “N-proteome”) revealed that cleavable pre-sequence are significantly more frequent than predicted [9]. Importation/maturation pathways have been mainly studied using fungal model organisms and are far from being fully deciphered in higher eukaryotic organisms. Elucidation of these pathways for mammalian mitochondrial proteins is critical to the understanding of the mechanisms underlying mitochondrial biogenesis [3].

The goal of the present study was to establish a unique and robust tool to enable exploration of the impact of precursor sequences on the import process and ultimately to have access to naturally processed mitochondrial proteins. The innate

ability of the Modified Vaccinia Virus Ankara (MVA-T7) to co-internalize a vector into mammalian cells was previously described, and has been named MVA-mediated transfer (MVA-mtr) [22]. Used in combination with lipofectamine transfection, the MVA-mtr protocol allows for the enhanced expression of recombinant proteins [22]. We used the MVA-mtr method to express human mitochondrial proteins in hamster cells (BHK21) and validated it as an expression system for producing and characterizing naturally maturated human proteins upon mitochondrial import. Our comparative analysis demonstrates that expression in the heterologous MVA-infected hamster cell-line does not alter the targeting, internalization or localization of human mitochondrial proteins. In addition, this method was applied to solve an important biological question, namely deciphering the yet unknown processing site of the mitochondrial aspartyl-tRNA synthetase (mt-AspRS). The actual processing site was found to be different than the *in silico* prediction.

We propose a powerful protein expression tool dedicated to mitochondrial proteins, which combines the regulation capability of the widely used bacterial expression system with the processing properties of eukaryotic cells. Theoretically, any sequence can be expressed, so that the validated MVA-mtr system has a wide-range of applications for the characterization of mitochondrial proteins. It also allows for the exploration of molecular impacts of the exponentially growing number of pathology-related mutations in proteins involved in mitochondrial biogenesis (reviewed in [23-27]).

Results

Experimental workflow to express precursor sequences of mitochondrial proteins

We developed an expression system for mammalian proteins of mitochondrial localization, which enables us to express any precursor sequence and to investigate the matured protein within mitochondria.

The innate ability of the Modified Vaccinia Virus Ankara (MVA-T7) to co-internalize a vector into mammalian cells without using any further costly transfectant reagents was previously described, and named MVA-mediated transfer (MVA-mtr) [22]. This virus is species-specific restricted to hamster cells (BHK21) and thus requires only safety level 1, facilitating the co-infection of high quantities of cells. We developed a system, which combines a minimal set of expression vectors (all derived from pBCJ739.14 [22]), expressing precursor proteins fused with a combination of C-terminal epitopes, that are either affinity tags (Flag, His, Strep) or fluorescent markers (GFP, mCherry) (Fig. 1A). Sequences of interest are cloned downstream of the Bacteriophage T7 promoter, and the LacO binding site for IPTG controlled transcription. The appropriate T7-polymerase is

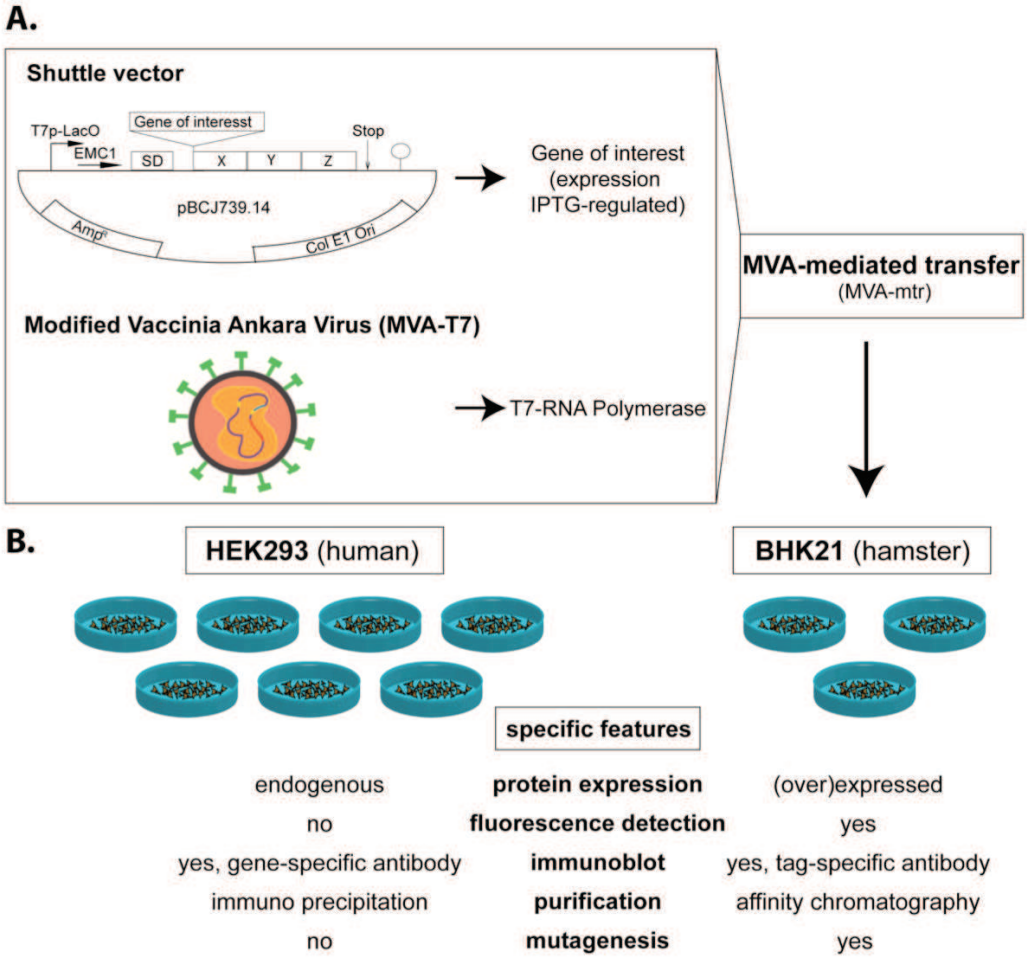


Figure 1: Schematic representation of protein overexpression in hamster cells (BHK21) and comparison with endogenous expression in human cells (HEK293). A. Genes of interest are cloned in a shuttle vector downstream of T7-LacO promoter as previously described in [22]. T7p-LacO indicates the T7 bacteriophage promoter with a downstream lacO binding site for IPTG inducible expression. EMC1 is the encephalomyocarditis virus leader region. The arrow with Stop indicates the translational stop and the stem-loop is the T7 rho-independent transcriptional terminator. The letters X, Y and Z correspond to different combinations of Flag or His-epitopes (X); GFP or mCherry sequences or nothing (Y); and to Strep II-epitope or nothing (Z). The Modified Vaccinia Virus Ankara (MVA-T7) is an attenuated strain of host specificity restricted to hamster cells and provides T7-RNA Polymerase. The shuttle vector is introduced into by MVA-mediated transfer (MVA-mtr) [22]. **B.** Features of either natural expression system (in HEK293) or engineered expression system (BHK21) are given.

additionally IPTG inducibly expressed from the viral genome [22,28]. Vectors include both bacterial and mammalian transcription/translation elements permitting the easy shuttling between bacterial and eukaryotic expression systems.

The advantages and specific features of this engineered expression system in

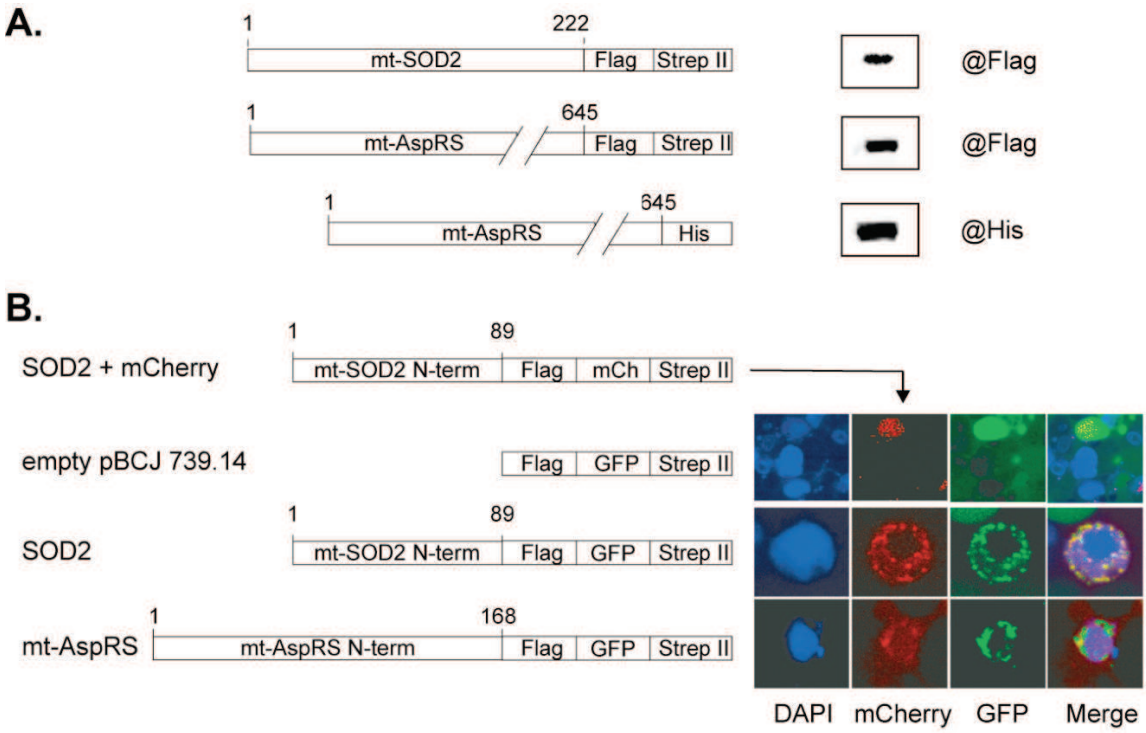


Figure 2**: Expression and location of mitochondrial proteins in hamster (BHK21) cells. A.** Detection of the expression of human mitochondrial protein sequences in BHK21 cells by western blot. Constructions used are schematized on the left. Full-length sequences of mt-SOD2 and mt-AspRS are upstream Flag/StrepII or His epitopes. Immuno detection is performed using antibodies against the Flag or His epitopes. **B.** Confocal microscopy demonstrating mitochondrial co-localization of GFP fused with different N-terminus of precursor sequences (including N-terminal mitochondrial targeting sequences). The nucleus is stained with DAPI (in blue). SOD2, a known protein of mitochondrial location, is co-expressed as a fusion with mCherry sequence (mCh) to serve as mitochondrial marker protein (in red). GFP-fused proteins are in green. The merged co-localization of mCherry and GFP expressing sequences is in yellow.

hamster cells are compared to the native expression in human cells (Fig. 1B). After transient expression of precursor proteins in the cytosol of mammalian cells, they are addressed to, imported into and processed inside mitochondria. The expressed proteins can either be detected by fluorescence microscopy or by using an affinity-tag specific antibody. In addition, the processed proteins can easily be recovered from cell lysate by affinity chromatography for downstream applications.

To validate our engineered expression system, two well-known proteins from different protein families were used: the superoxide dismutase 2 (SOD2) and the mitochondrial aspartyl-tRNA synthetase (mt-AspRS). SOD2 is a well-characterized

protein of mitochondrial matrix localization [29]. Mt-AspRS is nucleus-encoded and is a key actor of the mitochondrial translation [30].

Human mitochondrial proteins are expressed in hamster cells and targeted to mitochondria

As a first validation step, we investigated the expression of different precursor sequences (Fig. 2A). SOD2 and mt-AspRS, expressed with C-terminal fused Flag- and StrepII-tags, were detected by western blotting using an anti-Flag antibody. Due to the fact that affinity Ni^{2+}-purification technique is widely used, the expression of the mt-AspRS construct containing a C-terminal $(His)_6$-tag was additionally validated. To verify that expressed precursor proteins are indeed addressed to the mitochondria, we expressed different N-terminal sequences containing MTSs together with a C-terminal GFP fusion protein (Fig. 2B). The mt-SOD2(Nterm)-Flag-GFP-StrepII and mt-AspRS(Nterm)-Flag-GFP-StrepII constructs were introduced into BHK21 cells using MVA-mtr along with a SOD2-mCherry fusion protein (as a positive control for mitochondrial targeting) to evaluate mitochondrial co-localization *in vivo*. We observed an overlap of both the GFP and mCherry fluorescent signals, showing that the N-terminal fragments of both SOD2 and mt-AspRS were co-localized to mitochondria. As expected, the signal from empty Flag-GFP-StrepII construct remains diffuse in the cell.

Expressed proteins are internalized into mitochondria

To confirm that the recombinant fusion proteins are not only co-localized but also internalized within the mitochondria, we performed a proteinase K assay [31]. In this protocol, the fusion proteins were expressed within BHK21 cells and the mitochondria were purified. The purified mitochondria were subjected to a proteinase K treatment degrading all proteins exterior of the mitochondria that are not protected by the double membrane. As a control, recombinant GFP protein was exogenously added. This experiment was performed using the full-length SOD2 and mt-AspRS constructs with Flag and StrepII tags. The results showed that indeed both recombinant full-length products, SOD2 and mt-AspRS were resistant to 25 minutes of proteinase K treatment suggesting internalization into the mitochondria, whereas exogenous GFP was almost completely degraded at this time-point (Fig. 3).

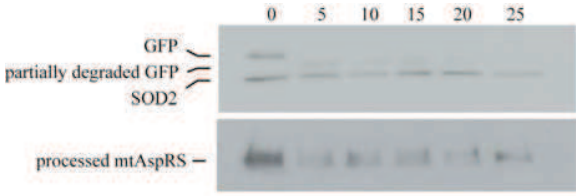


Figure 3: Western blot analysis of proteinase K protection assay. Mitochondria purified from cells expressing full-length SOD2 or mt-AspRS are mixed with recombinant GFP protein (positive control for extra-mitochondrial proteinase K-mediated

degradation) and contacted with proteinase K (20 µg/ml) on ice. Samples were removed at different time points (every five min until a final incubation time of 25 min) and the reaction was stopped by addition of PMSF and snap freezing on dry ice. Proteins are detected using anti-Flag antibody.

Intrinsic submitochondrial organization of proteins in human and hamster cells

To identity the final localization of expressed proteins inside mitochondria, we made reference to the known intrinsic localization of the SOD2 in the matrix [29] and Prohibitin in the membrane [32], and we established the one for mt-AspRS in human cells (HEK293). We thus applied mitochondrial purification and fractionation protocol [33] on both BHK21 and HEK293 cell lines. Fractionation separates mitochondria into matrix fraction and membrane fraction. The membrane fraction was further treated with sodium carbonate to separate loosely anchored proteins from deeply anchored ones [34]. Western blot detections showed that as expected SOD2 and Prohibitin are respectively localized in the matrix fraction and in the membrane fraction, in both the human and the hamster cells (Fig. 4A). This confirmed the high quality of the

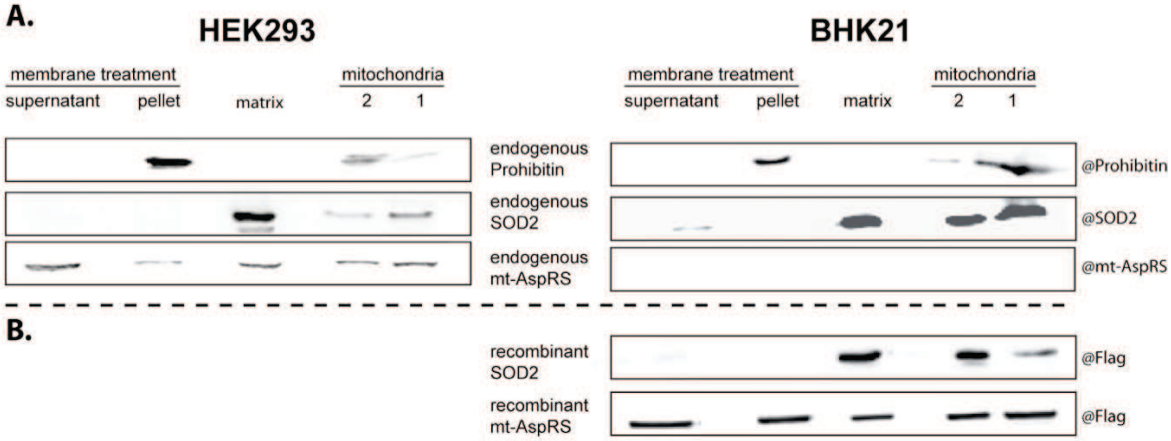


Figure 4**: Sub-mitochondrial localization of proteins in human (HEK293) and hamster (BHK21) cells. A.** Mitochondria from HEK293 or BHK21 cells were isolated and fractionated as described in Materials and Methods. *Human and hamster endogenous Prohibitin and SOD2 were detected by Western blots using human antibodies of cross-species activities. Endogenous hamster mt-AspRS could not be detected with the human-specific antibodies against mt-AspRS.* ***B.*** *Sub-mitochondrial localization of expressed proteins. Mitochondria were isolated from BHK21 cells expressing either human SOD2 or human mt-AspRS, and fractionated. Western blot detections of expressed proteins were performed using an antibody specific to the Flag epitope. In* ***A.*** *and* ***B.*** *Mitochondria 1 and 2 correspond to mitochondria before and after the density centrifugation purification step, respectively. Matrix contains proteins, which are found in the supernatant after mitochondria lysis and centrifugation. The remaining pellet (containing the membranes) was treated with Na_2CO_3 followed by a centrifugation step, which leaded to supernatant and pellet fractions*

fractionation procedure and the absence of cross-contamination from one fraction to the other one. Western blot detection revealed a multiple distribution of the mt-AspRS within all fractions of HEK293 cells (Fig. 4A). Endogenous mt-AspRS was not detected in BHK21 cells, due to the absence of antigenic cross-reactivity of antibodies directed against human mt-AspRS.

We expressed precursor SOD2 and mt-AspRS human sequences (with Flag-StrepII tags) in BHK21, purified and sub-fractionated mitochondria, and detected these proteins (using anti-Flag antibodies) in the same sub-mitochondrial fractions as the intrinsic ones (Fig. 4B). This demonstrates that localization of recombinant exogenous proteins is conserved in hamster cells as compared to human cells.

Implementation of the expression system to determine the N-terminal maturation site of mt-AspRS

We used MVA-mtr expression as a way to define the N-terminal amino acid for mature mitochondrial proteins. A peculiarity of MTSs is their non-conserved lengths, sequences and amino acid compositions [3], rendering the deciphering of N-terminal residues of every mitochondrial protein a unique issue, and sometimes challenging. The N-terminal residue of the mature SOD2 was previously deciphered in human cells [35] but only predicted for mt-AspRS [30] (Fig. 5A). As an application of our system, the precursor SOD2-Flag-StrepII sequence was expressed in BHK21, affinity-purified from isolated mitochondria and N-terminal-sequenced. The experimentally determined N-terminal peptide K_{25}HSLP corresponded to the previously reported site [35] (Fig. 5A). This corresponds to a further validation of the equivalency of hamster and human cell maturation mechanisms. Two different approaches were made to determine the N-terminal maturation site for mt-AspRS. First, the mt-AspRS was enriched by immuno-precipitation out of crude mitochondrial protein extract obtained from $120x10^8$ HEK293 cells. However, the purity and quantity of the enriched sample was not sufficient for sequencing by Edman degradation. To overcome this drawback, a multi-enzymatic cleavage approach was applied. To do so, the immuno-precipitated mt-AspRS (purified from crude mitochondrial protein extract obtained from $10x10^7$ of HEK293 cells) was in-gel-digested with chymotrypsin and trypsin. We identified by mass spectrometry the peptides Q_{35}SSQRRIPEFSS and R_{40}IPEFSSFVVR, generated respectively with chymotrypsin and trypsin (dashed boxes in Fig. 5B), as closest peptides to the N-terminus of mature mt-AspRS. To be noticed, none of the upstream theoretical tryptic-generated peptides (light gray boxes in Fig. 5B) were identified by mass spectrometry.

To further refine the N-terminal residues of mature mt-AspRS, a new strategy, namely incremental proteolytic based prediction of N-terminal peptide with mass

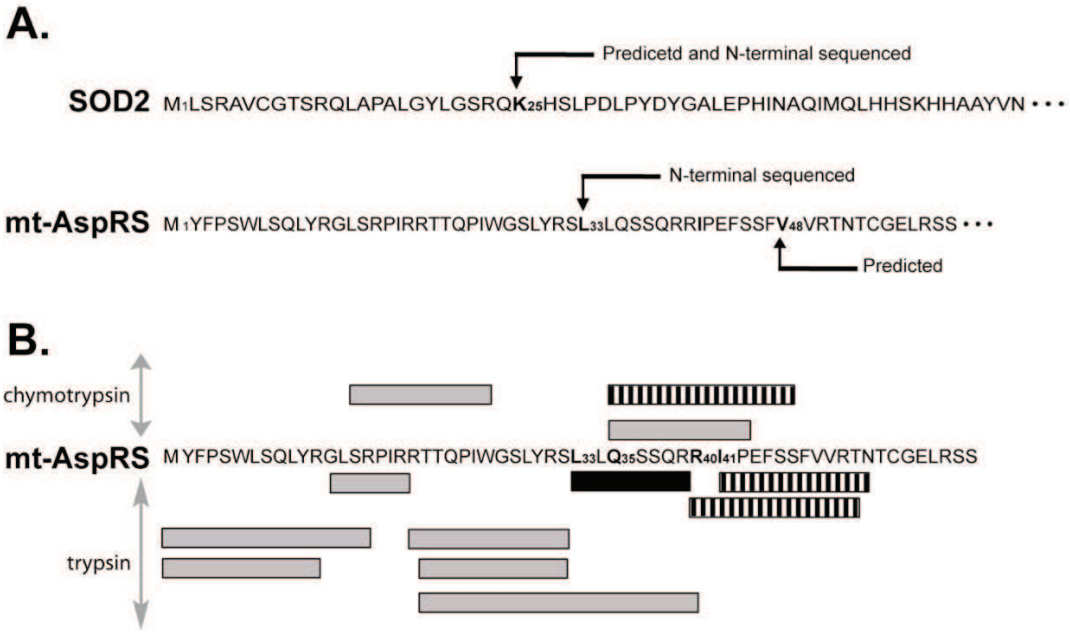


Figure 5**: Determination of N-terminal processing sites for SOD2 and mt-AspRS. A.** N-terminal sequences (from aa 1 to aa 60) of human SOD2 and mt-AspRS, where positions of predicted maturation and experimentally determined sites are displayed by arrowheads, according to references [35] and [30], respectively. **B.** Experimental determination of maturation site for mt-AspRS naturally expressed in HEK293 cells and recovered by immuno precipitation. Mt-AspRS is submitted to in-gel multi-enzyme digestion (chymotrypsin and trypsin). All possible peptides theoretically generated by chymotrypsin and trypsin digestions are indicated by boxes, those that are not experimentally found are in light gray. Peptides with dashes were those actually found by mass spectrometry. The peptide in black indicates the most probable N-terminal peptide, predicted after a theoretical sequential cleavage of N-terminal amino acids, and having a mass compatible with an existing spectrum.

spectrometry data, was developed. This strategy was to identify, within the spectra dataset, the non-predictable N-terminal tryptic-generated peptide. To do so, an *in silico* database was built up, which contains all masses theoretically calculated from tryptic-generated peptides that are sequentially digested from their N to their C parts (Supplementary Table I). Corresponding masses were searched for in the experimentally achieved dataset. Peptide identifications obtained from Mascot were validated manually by visual inspection using conventional fragmentation rules. We identified the peptide L_{33}LQSSQR as a doubly-charged peptide with a mass of 416.237 Th (m/z) and a mascot score of 38 as the most probable N-terminal peptide (the Mascot identity threshold was 31 for p-value<0.05). This revealed a maturation site at position L_{33} of mt-AspRS. The experimental mature protein thus includes 15 additional amino acids as compared to the bioinformatical prediction designing

maturation at position V_{48} [30]. The possibility that identified peptides downstream of position 33 and upstream of position 48 could correspond to second cleavage sites needs to be further explored.

To confirm our finding and to demonstrate the utility of the newly developed expression system, the precursor mt-AspRS-Flag-StrepII sequence was expressed in BHK21 cells using MVA-mtr. The matured mt-AspRS was purified from isolated mitochondria using affinity chromatography and resolved on SDS-PAGE. Gel bands containing mt-AspRS were sliced out and subjected to N-terminal Edman sequencing. The purity and quantity of purified mt-AspRS obtained from only ten 150 cm culture plates was sufficient for a successful N-terminal sequencing (Fig. 5A). This confirmed a maturation site at position L_{33} for mt-AspRS.

These results emphasize that overexpressed precursor proteins undergo the same mitochondrial processing pathways when expressed in a hamster cell-line as when intrinsically expressed in a human cell-line. They demonstrate that our expression system can be useful for defining N-terminal sequences for a wide range of mitochondrial processed proteins.

Discussion

Efficient expression of mammalian mitochondrial proteins (deprived of targeting sequence) in bacteria is difficult due to poor solubility linked to incorrect N-terminal definition [13], to the absence of adequate post-translational processing pathways (*e.g.* post-translational modifications essential for structure and function [36]), or to a prokaryotic protein synthesis that does not retrace the route followed by eukaryotic mitochondrial precursors proteins (*i.e.* synthesis in one compartement; processing, folding and maturation in another one). A comparative evaluation of the advantages and limitations of the frequently used eukaryotic expression systems [37] indicated that (i) although combining easy handling, rapid expression and high production, the yeast expression systems do not fully achieve mammalian post-translational modifications [38]; (ii) the baculovirus-insect expression system possesses eukaryotic processing capabilities and allows the production of high quantities of protein, but it is limited by the time-consuming preparation of recombinant baculovirus, which renders this method inappropriate for simple expression screening experiments [37,39]. Therefore, mammalian expression systems remain the system of choice. Transient transfection of expression vectors allows for fast screening of different constructions and for sufficient expression yield.

The method described here for investigating nucleus-encoded proteins of human (mammalian) mitochondrial location is remarkable in its simplicity. We used the MVA-mtr method [22] to express human mitochondrial proteins within hamster cells (BHK21) and validated it by characterizing

naturally maturated human proteins after mitochondrial import. We demonstrated by a comparative analysis that the expression in the heterologous MVA-infected hamster cell-line does not alter endogenous properties of targeting, internalization and localization of human mitochondrial proteins. It further does not alter their N-terminal maturation site. This system combines the regulation pattern of the bacterial expression system (T7-Polymerase, IPTG-regulated, fusion-tag purification) with the processing and maturation properties of eukaryotic cells. It also combines the manageability of the transient recombinant protein expression (easy cloning and fast screening) with the productivity of stable cell lines or recombinant baculovirus techniques (yield).

The expression procedure presented here can be used for a wide range of applications to investigate nucleus-encoded mammalian proteins. It has been utilized in the present study to establish the sub-mitochondrial distribution of selected proteins and to decipher the N-terminal amino acid of the mitochondrial aspartyl-tRNA synthetase. Additionally, the approach developed to refine N-terminal residues by generating theoretical proteolytic N-terminal peptide database(s) turned out to be of great value. The combination of MVA-mtr expression along with our novel incremental proteolytic based prediction of N-terminal peptide with mass spectrometry data analysis will undoubtedly be of great utility towards the systematic deciphering of N-terminal residues for mammalian mitochondrial proteins. The approach is extremely rapid and inexpensive when compared to standard techniques. Our expression system is potentially applicable to any sequence. It allows for mutagenesis and for the exploration of molecular basis for addressing, internalization, maturation, and localization. It is a powerful tool to answer emerging questions not only under wild-type situations/conditions but also to explore the impact(s) of a growing number of pathology-related mutations within nucleus-encoded genes coding for proteins involved in mitochondrial biogenesis (reviewed in [23-27]). The mt-AspRS belongs to a family of enzymes where 65 mutations were recently reported to cause human disorders with frequently severe consequences (reviewed in [40-42]). Several of the 28 mutations found in the mt-AspRS gene can be now analyzed in a systematic way, which should contribute to our understanding of the diseases and disorders related to these complex biological systems.

Online Methods

Cells, virus and growth conditions

For routine cloning, *Escherichia coli* DH5 alpha strain was grown in LB media supplemented with Ampicillin (100 µg/ml), when appropriate. For expression studies, Modified Vaccinia Ankara strain (MVA-EM24) was used as described previously [22,28]. Baby Hamster Kidney cells strain 21 (BHK21) (ATCC # CRL-12072) were cultivated in the Glasgow modification of Minimum Essential Medium (GMEM) supplemented with 10% fetal calf serum (FCS) and 1.5 g/ml bacto-tryptose phosphate. Standard cell cultivation conditions were used for BHK21 growth and viral infections (37º C, 5% CO_2). Viral stocks were prepared as described previously [28] except for final resuspension of viral stocks in PBS. Routine vaccinia manipulations were performed as described previously [43]. 1 mM IPTG was added to the growth media to induce expression of constructs. Human embryonic kidney cells (HEK) 293F (Invitrogen, Carlsbad, CA. USA) were grown in suspension under constant rotation (10 rpm) in a miniPERM bioreactor (Greiner Bio-one) using 293 SFM II medium (Gibeco®) with 1% Penicillin and Streptomycin at 37°C and 5% CO_2. Cells were harvested at a density of 10^7 cells/ml.

Plasmids

The MVA-*E. coli* shuttle vector that places the Flag-Strep II affinity tag at the C-terminal end of expressed proteins (pBCJ735.20) was previously constructed [22]. To construct the full-length mt-AspRS expression clone the *DARS2* gene was eluted as a Nsi I fragment from pFL and cloned into pBCJ735.20 resulting in pBCJ749.77. The SOD2 gene was amplified from cDNA that was constructed from RNA extracted from HEK 293 using Superscript II RT (Invitrogen, Carlsbad CA, USA) per manufacturer's protocol. Cloning of the SOD2 gene into the multi-cloning site of pBCJ735.20 resulted in the full-length affinity tagged constructs (pBCJ762.23). For fluorescent imaging studies the N-terminal peptides for the different genes were fused to GFP. The different gene fragments were amplified from cDNA and ligated into the multi-cloning site of pBCJ739.14. The N-terminal fragments cloned were: 168 aa of mt-AspRS (pBCJ753.2), and 89 aa of SOD2 (pBCJ747.2). To make a construct that expressed a known mitochondrial targeted protein we replaced the GFP in pBCJ747.2 with mCherry amplified from pRSET-mCherry (Invitrogen, Carlsbad CA, USA) resulting in pBCJ722.1.

Confocal imaging

For confocal imaging, $2x10^5$ BHK21 cells were seeded onto coverslips, previously treated with HistoGrip™ (Invitrogen Carlsbad, CA. USA), in a 6-well plate. Cells were grown using standard culture conditions. After one day of culture, 2 µg of GFP expression plasmid mixed with 2 µg control

plasmid expressing mCherry (pBCJ722.1) were introduced into BHK21 cells using MVA-mtr as described previously [22]. IPTG was added to the growth medium and cells were grown 24 hours. Coverslips were washed with PBS and cells were fixed using Vectashield® mounting medium with DAPI (Vector Laboratories, Burlingame, CA, USA), following the manufacturer instructions. Confocal imaging was conducted at "Plateforme RIO d'imagerie cellulaire Strasbourg Esplanade" (IBMP, Strasbourg).

Mitochondria purification and fractionation

The protocol for mitochondrial isolation and fractionation was adapted from a previously described method [33]. Briefly, mitochondria were isolated from cells growing in fifteen T175 culture flasks. Cells were washed with PBS, resuspended in 10 ml Lysis buffer (220 mM Mannitol, 70 mM Sucrose, 10 mM Hepes-KOH pH 7.5, 1 mM $MgCl_2$, 1 mM EDTA and 1 % BSA), and cracked with a warring blender or a dounce homogenizer. Intact cells were removed by brief centrifugation at 800*g*. Mitochondria were pelleted at 8 600*g*, resuspended in Lysis buffer and centrifuged again at 800*g* before being pelleted at 8 600*g* and loaded on a Percoll-Sucrose gradient, that was subsequently centrifuged at 100 000*g* for 45 min. Purified mitochondria, recovered from the gradient, were washed three times with Wash buffer (300 mM Mannitol, 10 mM K_2HPO_4/KH_2PO_4 pH 7.5, 1 mM EDTA). Wet mitochondrial pellet was weighed and resuspended in 10 µl/mg Wash buffer. Mitochondria were lysed by multiple freeze and thaw cycles and brief sonication. Intact mitochondria were removed by centrifugation at 8 600*g*. Matrix proteins were separated from insoluble membrane proteins by centrifugation at 100 000*g*. The pellet, containing the mitochondrial membranes, was treated 30 min with 0.1 M Na_2CO_3 pH 11.5 at room temperature [34]. Na_2CO_3 sensitive proteins were recovered in the supernatant after an additional centrifugation step at 100 000*g*. The remaining pellet contained deeply anchored membranes proteins.

Proteinase K protection assay

Affinity-tagged proteins were expressed using five to ten 10 cm cell culture dishes of BHK21 cells, employing the MVA-mtr protocol as described previously [22]. After mitochondria purification (see above), protein quantification was performed using Bio-Rad protein assay according to the manufacturer's instructions (BioRad, Hercules, CA. USA). Proteinase K protection assays [31] were set up on ice as follows where the final reaction concentrations were 6.25 mg/ml purified mitochondria, 1 mg/ml BSA, 20 µg/ml proteinase K, 8.75 µg/ml recombinant GFP in Assay buffer (10 mM Hepes-KOH pH7.5, 25 mM sucrose, 75 mM sorbitol, 100 mM KCl, 10 mM KH_2PO_4/K_2HPO_4 pH 7.5, 5 mM $MgCl_2$). At various time-points, 40 µl aliquots were

removed and the reaction was stopped by adding PMSF (final concentration 5 mM) and snap freezing on dry ice. Samples were separated on 10 or 12% SDS-PAGE gels and blotted to a 0.2 µm PVDF membrane. Immunoblots were performed using standard methods. Recombinant proteins were detected with an antibody specific for the Flag peptide (Sigma, St. Louis, MO, USA). Chemiluminescent detection was carried out using Pierce Detection Kit (Rockford, IL, USA) according to the manufacturer's instructions.

Expression, purification and N-terminal sequencing

For sequencing by Edman degradation: To obtain *in vivo* processed protein, affinity-tagged protein was over-expressed in ten 150 cm cell culture dishes of BHK21 cells, employing the MVA-mtr protocol as described previously [22]. After 48 hours of expression (IPTG-induced), cells were harvested by scraping with a rubber spatula and centrifugation. The cell pellet was washed once with PBS. Mitochondria were isolated as described above. Recombinant protein was isolated by streptavidin affinity chromatography using the Strep-Tactin system following the manufacturer's protocol (IBA, Göttingen, Germany). Elutions from the column were resolved on a 10% SDS gel, stained with coomassie blue and the bands containing the protein of interest were identified by MALDI-TOF mass spectrometry (IBMC proteomic platform). The identified bands were excised from the gel and N-terminal sequenced using Edman degradation by Alta Biosciences (Birmingham, UK).

For sequencing by multiple enzyme digestion: HEK293F cells were harvested when at a confluence of 80%. The cell pellet was washed in buffer A (20 mM Hepes-KOH pH 7.5, 200 mM KAc, 0.2% Triton X100) and cracked by sonication (5 strokes at 6 watts for 10 sec). The lysate was diluted two times in buffer B (20 mM Hepes-KOH pH 7.5, 200 mM KAc) in order to reduce the Triton concentration <0.05%, which is suitable for further immuno-precipitation experiment. Contaminating membranes were removed by centrifugation at 100 000*g* for 45 min at 4°C. The crude cellular extract was incubated for 1 hour at 4°C with first protein A agarose beads, and second with non-specific IgG covalently coupled with protein A agarose beads. The pre-cleared lysate was then incubated for 2h at 4°C with human mt-AspRS-specific IgG covalently coupled with protein A agarose beads. Afterward, the beads were incubated with 10 volumes of PBS for 30 min at 4°C, loaded on a gravity-flow column, and further washed with 5 volumes of PBS. Mt-AspRS was recovered from the beads by incubation 5 min at 50°C in 250 mM Tris-HCl pH 7.5 and 2% SDS (2 times), concentrated using Centricon centrifugal filter device (10kDa, Millipore), and loaded on 10% SDS-PAGE.

Mass spectrometry analysis:

Sample preparation

Protein bands excised from the SDS-PAGE electrophoresis gel were transferred into 96-well microtitration plates. Gel slices were washed with three cycles of incubations in 25 mM ammonium bicarbonate for 15 min and then in 25 mM ammonium bicarbonate containing 50% (v/v) acetonitrile for 15 min. Samples were dehydrated with 100% (v/v) acetonitrile and then reduced with 10 mM DTT for 1 hour at 56°C. Proteins were then alkylated with 55 mM iodoacetamide for 1 hour in the dark at room temperature. Gel pieces were washed again with the destaining solutions described above. 20 ng of sequencing-grade enzyme (5ng/µL, see below) in 25 mM ammonium bicarbonate was added on each dehydrated gel piece. After 30 min of rehydratation at room temperature, 30 µL of 25 mM ammonium bicarbonate was added on gel pieces before digestion, overnight at room temperature for an optimal and efficient digestion. The resulting peptides were then extracted from the gel pieces in 30 µL of 60% (v/v) acetonitrile and 5% (v/v) formic acid. The initial digestion and extraction supernatants were pooled and vacuum-dried in a SpeedVac.

Multi-enzymatic digestions

Protein were submitted to digestion with the following enzymes: (i) trypsin (V5111, Promega, Madison, WI); (ii) chymotrypsin (V106A, Promega); (iii) Asp-N (V1621, Promega); (iv) Arg-C (Protease Profiler Kit PP0500, Sigma, St. Louis, MO, USA); (v) Glu-C (PP0500, Sigma). A first overnight enzymatic digestion was performed with each of the 5 previously described enzymes. Two double-digestions were also carried out, using a combination of (vi) trypsin and Asp-N or (vii) Asp-N and Arg-C. In the case of a double-digestion, 20 ng of the second enzyme were added directly on the peptides mixture for a second digestion for 5 hours at 37°C.

Identification of the N-terminal peptide by nanoLC-MS/MS

The dried peptides were re-suspended in 15 µL of water containing 0.1% FA (solvent A). Peptide mixtures were analyzed using a NanoLC-2DPlus system (with nanoFlex ChiP module; Eksigent, ABSciex, Concord, Ontario, Canada) coupled to a TripleTOF 5600 mass spectrometer (ABSciex) operating in positive mode. After 10 min of desalting and concentration, the precolumn was switched online with the analytical ChIP C-18 analytical column (75µm ID x 15 cm ChromXP; Eksigent). Peptides were eluted by using a 5%-40% gradient of solvent B (0.1% formic acid in ACN) for 60 min at a flow rate of 300 nL/min. The TripleTOF 5600 was operated in data-dependant acquisition. Up to 20 of the most intense multiply charged ions were selected for CID fragmentation: this “Top20” method had a constant cycle time of 3.3 s and was set to a high-sensitivity mode.

Theoretical proteolytic N-terminal peptide database

The full-length protein sequence of mt-AspRS was *in silico* digested: the N-terminal amino acid was sequentially removed and the mass of the new generated peptide (x-1) was calculated. This procedure was performed till the first N-terminal experimental determined tryptic peptide (R_{40} IPEFSSFVVR). Mass spectrometer data were searched against this theoretical database using the Mascot algorithm (version 2.2, Matrix Science, London, UK). Peptide modifications allowed during the search were: N-acetyl (protein), carbamidomethylation(C) and oxidation (M). Mass tolerances in MS and MS/MS were set to 20 ppm and 0.5 Da, respectively; the instrument setting was specified as ESI-QUAD-TOF and 2 miscleavages were allowed. Trypsin was specified for enzyme specificity. Peptide identifications obtained from Mascot were validated manually by visual inspection using the conventional fragmentation rules. Mascot algorithm returned an individual identity score of 31 above which p-value was below 0.05.

Acknowledgments

We thank J. Mutterer (Microscopy and Imaging Platform, IBMP, Strasbourg) and E. Julien for their contributions. This work was supported by Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg (UdS), Association Française contre les Myopathies (AFM), ANR MITOMOT (ANR-09-BLAN-0091-01/03) and by the French National Program "Investissement d'Avenir" (Labex MitCross), administered by the "Agence National de la Recherche", and referenced ANR-10-IDEX-002-02. B.C.J. was supported by CNRS postdoctoral funding, and H.S was supported by Région Alsace, Université de Strasbourg, AFM, and Fondation des Treilles.

Authors Contributions

BCJ, HS, LMD and MS designed experiments. BCJ, HS, LK conducted experiments. LK and PH performed mass spectrometry analysis. BCJ, HS, LK, LMD, CF and MS analyzed the data. HS and MS wrote the paper. BCJ, LK, LMD and CF contributed to the critical revision of the paper.

References

1. D.J. Pagliarini, S.E. Calvo, B. Chang et al., *Cell* **134**, 112 (2008).
2. W. Neupert and J.M. Herrmann, *Annu. Rev. Biochem.* **76**, 723 (2007).
3. A. Chacinska, C.M. Koehler, D. Milenkovic et al., *Cell* **1387**, 628 (2009).
4. O. Schmidt, N. Pfanner, and C. Meisinger, *Nat Rev Mol Cell Biol.* **11**, 655 (2010).
5. T. Becker, L. Böttinger, and N. Pfanner, *Trends Biochem Sci.* **37**, 85 (2012).
6. D. Mossmann, C. Meisinger, and F.N. Vögtle, *Biochim. Biophys. Acta* **1819**, 1098 (2012).
7. O. Gakh, P. Cavadini, and G. Isaya, *Biochim. Biophys. Acta.* **1592**, 63 (2002).
8. G. Isaya, F. Kalousek, W.A. Fenton et al., *J. Cell Biol.* **113**, 65 (1991).

9. F.N. Vögtle, S. Wortelkamp, R.P. Zahedi et al., *Cell* **139**, 428 (2009).
10. J.C. Timmer, M. Enoksson, E. Wildfang et al., *Biochem. J.* **407**, 41 (2007).
11. J.M. Bullard, Y.-C. Cai, and L.L. Spremulli, *Biochem. Biophys. Acta* **1490**, 245 (2000).
12. Y.N. Yao, L. Wang, X.F. Wu et al., *FEBS Lett.* **534** (1-3), 139 (2003).
13. A. Gaudry, B. Lorber, M. Messmer et al., *Protein Engineering, Design and Selection* **25**, 473 (2012).
14. A. Neuenfeldt, B. Lorber, E. Ennifar et al., *Nucleic Acids Res.* **41**, 2698 (2013).
15. D. Stojanovski, N. Pfanner, and N. Wiedemann, *Methods Cell Biol.* **80**, 783 (2007).
16. C. Stadler, E. Rexhepaj, V.R. Singan et al., *Nat Methods.* **10**, 315 (2013).
17. H. Schägger and G. von Jagow, *Anal Biochem.* **199**, 223 (1991).
18. F. Pallotti and G. Lenaz, *Methods Cell Biol.* **80**, 3 (2007).
19. P. Edman, *Arch Biochem.* **22**, 475 (1949).
20. W.K. Huh, J.V. Falvo, L.C. Gerke et al., *Nature* **425**, 686 (2003).
21. J. Reinders, R.P. Zahedi, N. Pfanner et al., *J. Proteome Res.* **5**, 1543 (2006).
22. B.C. Jester, R. Drillien, M. Ruff et al., *J. Biotechnol.* **156**, 211 (2011).
23. H.T. Jacobs and D.M. Turnbull, *Trends Genet.* **21**, 312 (2005).
24. G.C. Scheper, M.S. van der Knaap, and C.G. Proud, *Nat. Rev. Genet.* **8**, 711 (2007).
25. S.R. Pieczenik and J. Neustadt, *Exp Mol Pathol.* **83**, 84 (2007).
26. A. Rötig, *Biochim Biophys Acta.* **1807**, 1198 (2011).
27. J. Nunnari and A. Suomalainen, *Cell* **148**, 1145 (2012).
28. M. Hebben, J. Brants, C. Birck et al., *Protein Expr Purif.* **56**, 269 (2007).
29. J.A. Melendez, R.P. Melathe, A.M. Rodriguez et al., *Cell Growth Differ.* **10**, 655 (1999).
30. L. Bonnefond, A. Fender, J. Rudinger-Thirion et al., *Biochemistry* **44**, 4805 (2005).
31. V. Shalak, M. Kaminska, and Mirande M., *Biochemistry* **48**, 9956 (2009).
32. M. Artal-Sanz and N. Tavernarakis, *Trends Endocrinol Metab.* **20**, 394 (2009).
33. H. Suzuki, T. Ueda, H. Taguchi et al., *J. Biol. Chem.* **282** (6), 4076 (2007).
34. L. Delage, P. Giegé, M. Sakamoto et al., *Biochimie* **89**, 658 (2007).
35. D. Barra, M.E. Schinina, M. Simmaco et al., *J Biol Chem.* **259**, 12595 (1984).
36. S. Sahdev, S.K. Khattar, and K.S. Saini, *Mol. Cell. Biochem.* **307**, 249 (2008).
37. J. Yin, G. Li, X. Ren et al., *J Biotechnol.* **127**, 335 (2007).
38. M. Schuster, A. Einhauer, E. Wasserbauer et al., *J Biotechnol.* **84**, 237 (2000).
39. M.M. van Oers, *J Invertebr Pathol.* **107**, S3 (2011).
40. T. Suzuki, A. Nagao, and T. Suzuki, *Annu. Rev. Genet.* **45**, 299 (2011).

41. S. Konovalova and H. Tyynismaa, *Mol Genet Metab.* **108**, 206 (2013).
42. H. Schwenzer, J. Zoll, C. Florentz et al., in *Topics in Current Chemistry-Aminoacyl-tRNA* (Springer, 2013).
43. P.L. Earl and B. Moss, *Curr Protoc Protein Sci.* **Chapter 5**, Unit5.14 (2001).

SUPPLEMENTARY DATA

.

Theoretical peptide sequence	Sequence name in the dedicated database	Sequence experimentally observed	m/z observed	Mascot Score
MYFPSWLSQL YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA1	SLLQSSQRR	1073.5680 (537.7913=2+)	3
		RIPEFSSFVVR	1335.7273 (446.2497=2+)	39
		IPEFSSFVVR	1179.6222 (590.8184=2+)	53
YFPSWLSQL YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA2	n.d.	n.d.	n.d.
FPSWLSQL YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA3	n.d.	n.d.	n.d.
PSWLSQL YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA4	n.d.	n.d.	n.d.
SWLSQL YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA5	n.d.	n.d.	n.d.
WLSQL YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA6	WLSQLYR	964.4956 (483.2551=2+)	2
LSQL YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA7	n.d.	n.d.	n.d.
SQL YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA8	SQLYRGLSR	1120.5786 (561.2966=2+)	8
QL YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA9	n.d.	n.d.	n.d.
L YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA10	n.d.	n.d.	n.d.
YRGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA11	n.d.	n.d.	n.d.
RGLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA12	RGLSRPIR	953.6158 (477.8152=2+)	1
GLSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA13	GLSRPIR	839.5205 (420.7675=2+)	7
LSRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA14	n.d.	n.d.	n.d.
SRPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA15	n.d.	n.d.	n.d.
RPIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA16	n.d.	n.d.	n.d.
PIRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA17	n.d.	n.d.	n.d.
IRR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA18	n.d.	n.d.	n.d.
RR TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA19	n.d.	n.d.	n.d.
R TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA20	n.d.	n.d.	n.d.
TTQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA21	SWLSQLYR	1093.5482 (547.7814=2+)	3
TQPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA22	n.d.	n.d.	n.d.
QPIWGSLY RSLLQSSQRR IPEFSSFVVR	AA23	n.d.	n.d.	n.d.
PIWGSLY RSLLQSSQRR IPEFSSFVVR	AA24	n.d.	n.d.	n.d.
IWGSLY RSLLQSSQRR IPEFSSFVVR	AA25	n.d.	n.d.	n.d.
WGSLY RSLLQSSQRR IPEFSSFVVR	AA26	WGSLYR	822.4238 (412.2192=2+)	6
GSLY RSLLQSSQRR IPEFSSFVVR	AA27	n.d.	n.d.	n.d.
SLY RSLLQSSQRR IPEFSSFVVR	AA28	n.d.	n.d.	n.d.
LY RSLLQSSQRR IPEFSSFVVR	AA29	n.d.	n.d.	n.d.

Y RSLLQSSQRR IPEFSSFVVR	AA30	n.d.	n.d.	n.d.
RSLLQSSQRR IPEFSSFVVR	AA31	n.d.	n.d.	n.d.
SLLQSSQRR IPEFSSFVVR	AA32	n.d.	n.d.	n.d.
LLQSSQRR IPEFSSFVVR	AA33	LLQSSQR	830.4587 (416.2366=2+)	**38**
LQSSQRR IPEFSSFVVR	AA34	n.d.	n.d.	n.d.
QSSQRR IPEFSSFVVR	AA35	n.d.	n.d.	n.d.
SSQRR IPEFSSFVVR	AA36	n.d.	n.d.	n.d.
SQRR IPEFSSFVVR	AA37	n.d.	n.d.	n.d.
QRR IPEFSSFVVR	AA38	n.d.	n.d.	n.d.
RR IPEFSSFVVR	AA39	n.d.	n.d.	n.d.
R IPEFSSFVVR	AA40	n.d.	n.d.	n.d.

Supplementary Table I: Identification by mass spectrometry of the different N-terminal peptides arising from the theoretical database. The 50 first amino acids from mt-AspRS were used to create a theoretical fasta database, in which each sequence was designed as a sub-sequence from these 50 amino acids (1-50, 2-50, 3-50, …, 39-50, 40-50). The Mascot algorithm was used to evaluate the most probable N-terminal peptide: the threshold to find a good hit was set to 31, corresponding to the individual identity score calculated by Mascot for p-value<0.05. The table lists the 40 theoretical N-terminal peptides: in case of detection, the peptide sequence, the m/z value, the charge state and the Mascot score is indicated (n.d. = not detected).

4 Variation of the theoretical pI with the length of N-terminal sequence of mt-aaRSs

It is known that the MTS contains basic residues and contain amphiphilic α-helices (reviewed in e.g. Chacinska et al., 2009). It would thus be of interest to see if those residues are mostly concentrated within the MTS, and if the removal of the MTS change the isoelectri point (pI) of the protein. We determined the theoretical pI using ProtParam prediction tool (Gasteiger et al., 2003) of full-length mt-aaRSs and of partially deleted sequences (by step of 5 aa until having remove 100 aa) (**Figure 31**).

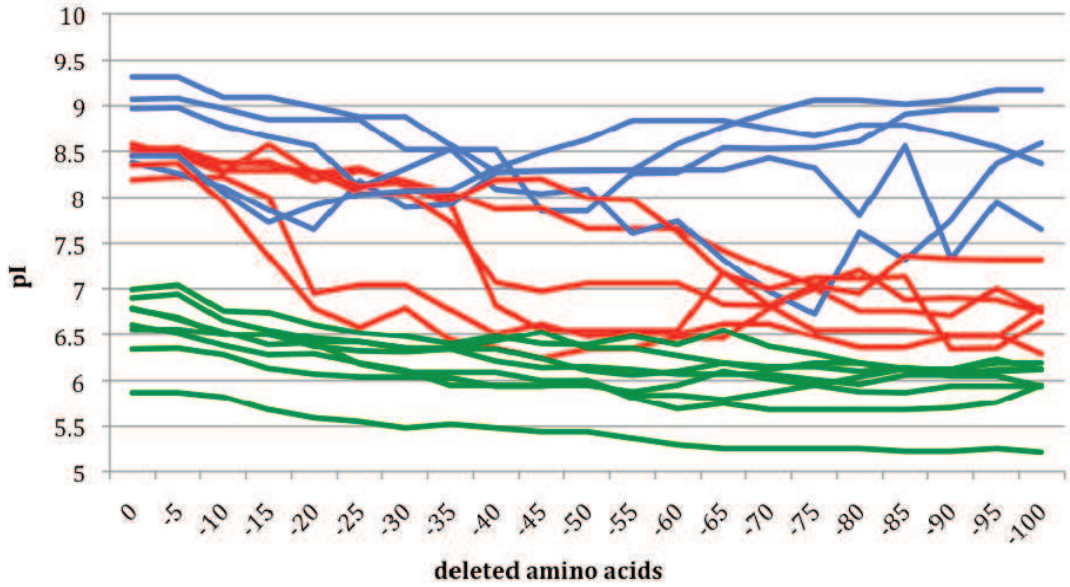


Figure 31: Predicted pIs as a function of the sequence size of mt-aaRSs. The full-length protein sequences were cleaved from the N-terminal site of the protein. The pI's of the theoretical proteins were calculated using ProtParam. The green (mt-PheRS, mt-ThrRS, mt-IleRS, mt-GlyRS, mt-LysRS, mt-AlaRS, mt-AsnRS and mt-ValRS), red (mt-CysRS, mt-LeuRS, mt-AspRS, mt-SerRS, mt-HisRS and mt-ProRS) and blue(mt-Glu, mt-TrpRS, mt-TyrRS, mt-MetRS and mt-ArgRS) color display different groups sequences in regard to the correlation of the pI with the protein sequence.

When considering solely the full-length sequences, two clusters of mitochondrial precursor aaRSs are visible (**Figure 31**). The first cluster (blue+red curves) consists of aaRSs, which have a pI between 8 and 9.3 (mt-TrpRS, mt-TyrRS, mt-GluRS, mt-CysRS, mt-LeuRS, mt-ArgRS, mt-MetRS, mt-AspRS, mt-SerRS, mt-HisRS and mt-ProRS) and the second (green curves) consists of aaRSs having a pI between 5,8 and 7

(mt-PheRS, mt-ThrRS, mt-IleRS, mt-GlyRS, mt-LysRS, mt-AlaRS, mt-AsnRS and mt-ValRS). Similar, pI distributions are reported earlier (Sissler et al., 2008). When, the N-terminal amino acids are sequentially deleted, prediction of pI values lead to three main clusters. Some sequences kept predicted pIs despite deletion of N-terminal residues [(mt-Glu, mt-TrpRS, mt-TyrRS, mt-MetRS and mt-ArgRS for pI>7.5, blue curves) and (mt-PheRS, mt-ThrRS, mt-IleRS, mt-GlyRS, mt-LysRS, mt-AlaRS, mt-AsnRS and mt-ValRS for p<7, green curves)]. Some sequences underwent a significant pI change from basic to acidic when the N-terminal amino acids are removed (mt-CysRS, mt-LeuRS, mt-AspRS, mt-SerRS, mt-HisRS and mt-ProRS, red curves).

To further characterize this observation, we plotted the pI as a function of the molecular weight (**Figure 32**). For clarity of representation, the three clusters are displayed separately. In this illustration, it appears that precursor proteins with basic pI (blue and red clusters) have also molecular weight < 70kDa (except mt-LeuRS), while precursor aaRSs with a low pI<7 (green cluster) have mainly molecular weights >70kD (exception mt-PheRS and mt-AsnRS). Of note, the typical MTS has a size of 30-40 amino acids and its cleavage site can be predicted. Those predictions are indicated in the figure. It is interesting to observe that the aaRSs in the red cluster (**Figure 32**) may undergo a drastic change in their overall-net charge by the removal of some N-terminal amino acids. Thus, if predictions turn out to be true, mature aaRSs from this cluster have pI significantly lower than the one from the corresponding precursor aaRSs. As an example, the precursor of mt-AspRS has a pI of 8,2 and the first cleavage site is predicted to generate a protein with a pI of 7. More drastic changes are observed for the mt-ProRS. In this case, the precursor protein has a pI of 8,5, a first cleavage generates a protein with a pI of 7,9 and a second predicted cleavage leads to mature protein with a pI of 6,8.

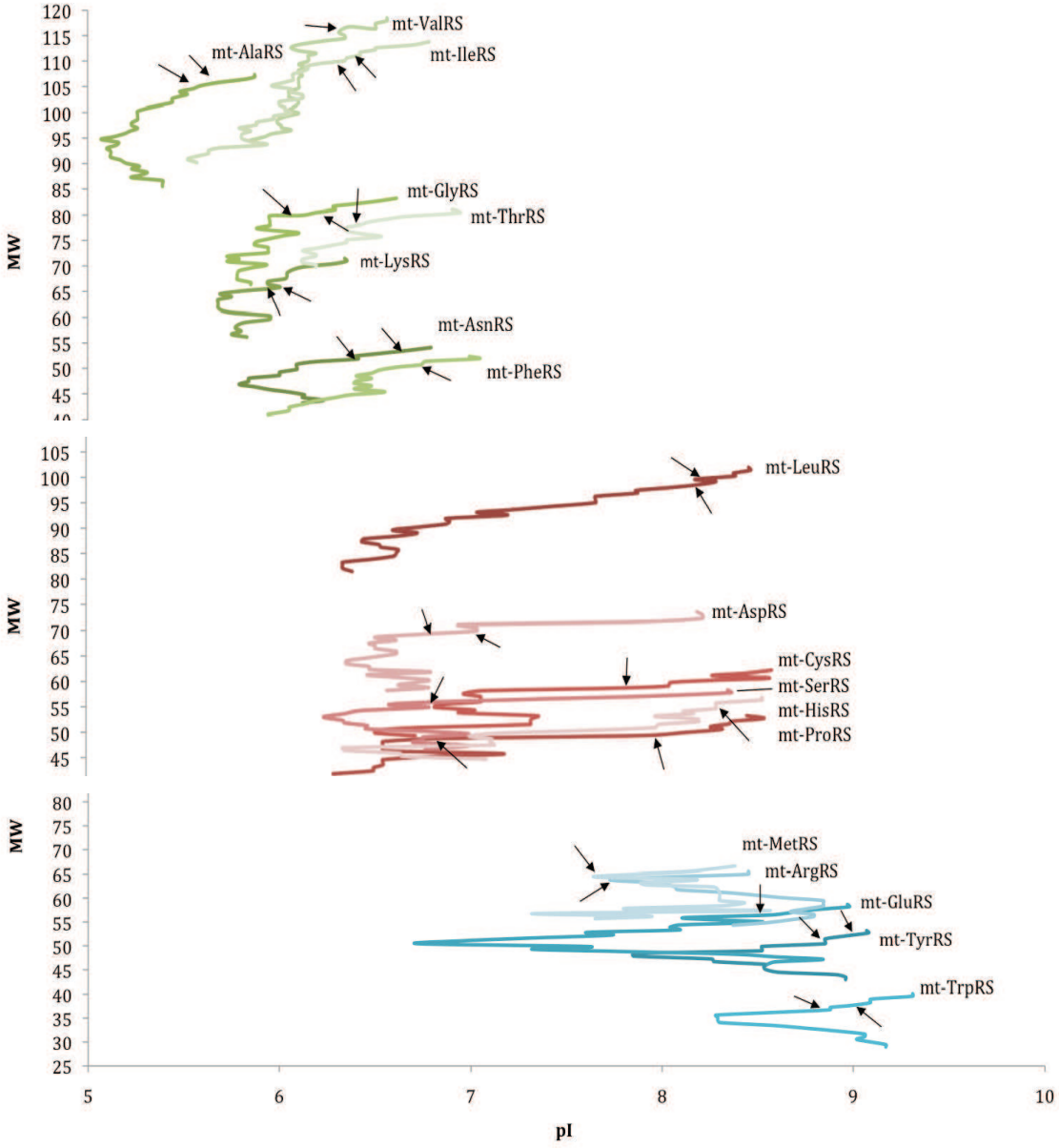


Figure 32: Relationship of pI and the molecular weight (MW) of the mt-aaRSs. The full-length protein sequences were cleaved from the N-terminal side of the protein. The pI's of the theoretical proteins were calculated using ProtParam. Colors green, red, blue correspond to the clusters in **Figure 31**. Predicted cleavage sites are indicated with arrows.

In summary, this theoretical analysis showed that the precursor and mature aaRSs might significantly differ regarding their physical-chemical properties (**Figure 33**). Considering the precursor aaRSs, some are basic (blue and red clusters) and some are acidic (green clusters). This study showed, that some precursor aaRSs undergo drastic changes in pI once their MTS is cleaved (e.g. mt-AspRS or mt-ProRS).

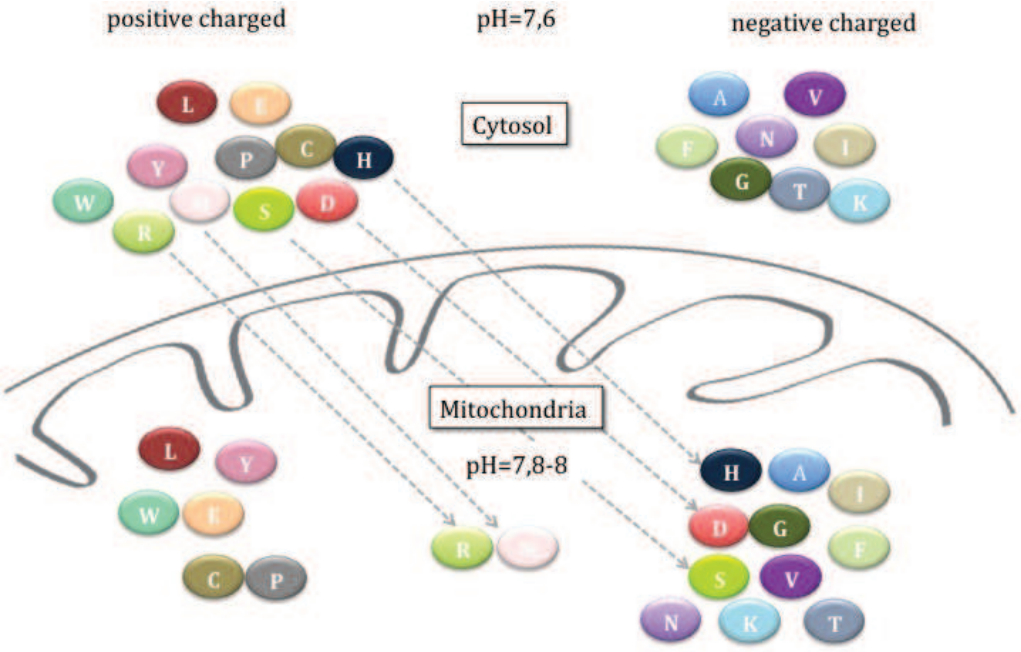


Figure 33: Schematized classification of aaRSs according to their predicted net charge. The cytosolic and mitochondrial location are illustrated. Proteins with a pI>pH have a positive net charge and proteins with a pI<pH have a negative net charge. The approximate pH of the cytosol and the matrix is given. Of note, precursor mt-aaRSs a likely unfolded upon the import inside mitochondria and thus do not have the same shape like the folded matured proteins.

It could be then hypothesized that that precursor proteins in cytosol need different conditions (e.g. partner proteins) than the mature proteins inside mitochondria. Modification of the net charge may have consequences on the folding, possible interaction(s) with cellular partners or may impact the sub-mitochondrial location. Of note, the pI of proteins is defined as the overall net charge of all amino acids. If the local pH is equal to the pI of a protein, than the protein have no net charge (protein has the same amount of positive and negative charged amino acids). If the pH < pI, than proteins have a positive net charge and at pH>pI they have a negative overall net charge.

Box 2: Relationship of overall net charge of protein to the local pH

$$\boxed{\text{pH}} \;<\; \boxed{\text{pI=pH}} \;<\; \boxed{\text{pH}}$$

$$\oplus \qquad\qquad 0 \qquad\qquad \ominus$$

Overall net charge of protein

In addition, it is estimated that the pH inside mitochondria is about 8 and in the cytosol about 7,6 (Alberts et al., 2002). We used these values as an approximation, despite the fact that they remain controversial. The different pH of the two cellular compartments may obviously impact the belonging of the proteins to the positively charged set of proteins or to the negatively charged set of proteins (**Figure 33**). This is indeed the case for the mt-AspRs, mt-HisRS and mt-SerRS, which may have change their net charge after the mitochondrial import and N-terminal maturation.

The pI of the mt-ArgRS and mt-MetRS, is predicted to be similar to the pH inside mitochondria, thus their net charge is close to zero. Of note, a posttranslational modification (e.g. phosphorylation, N-terminal acetylation) may further alter the net charge of the protein. This aspect was not considered at this point.

What is the implication of these results? First we know that the membrane is negatively charged because of the phospholipids. So positively charged proteins can interact with the membrane. We saw indeed that the most basic aaRS (mt-TrpRS) were found only in the membrane fraction, as determined by western blot analysis of the mitochondrial sub fractions. We also saw that the predicted negatively charged mt-GlyRS and mt-IleRS were only found in the matrix. Of note, the mt-CysRS was found only in the matrix despite the fact that it is predicted to be a positively charged protein. In this case, the situation is not so clear because the pI of the predicted matured protein is close the pH of the mitochondria matrix suggesting that the net charge is neutral or slightly positive. This hypothesis is however not sufficient to explain the localization of all aaRSs. Mt-AspRS and mt-SerRS are both predicted as negatively charged and found both in the matrix but also in the membrane. The membrane location of mt-AspRS was more extensively investigated. We showed that it is anchored more strongly than via solely electrostatic interactions. We also cannot exclude the possibility that post-translational modification(s) may play a role. The role of the switch from positively to negatively charged proteins after N-terminal maturation needs to be figured out in regard the function and localization of the protein.

5 Experimental determination of the N-terminal processing site of mt-AspRS

In the article #2, the N-terminal sequence of the mt-AspRS was deciphered at a position distant from those predicted using bioinformatic tools (**Figure 34**). As mentioned at several instances, differences at the N-termini may cause differences in the pI. And indeed, the recombinant protein starting at isoleucine 41 as N-terminal amino acid had a predicted pI of 6.53 and was shown to have an improved solubility and activity that facilitated its crystallization (Gaudry et al., 2012) compared to the other variants V48 (pI of 6,65) or L33 (pI of 6,9). Interestingly, the residue L33 was experimentally determined to be the N-terminal amino acid.

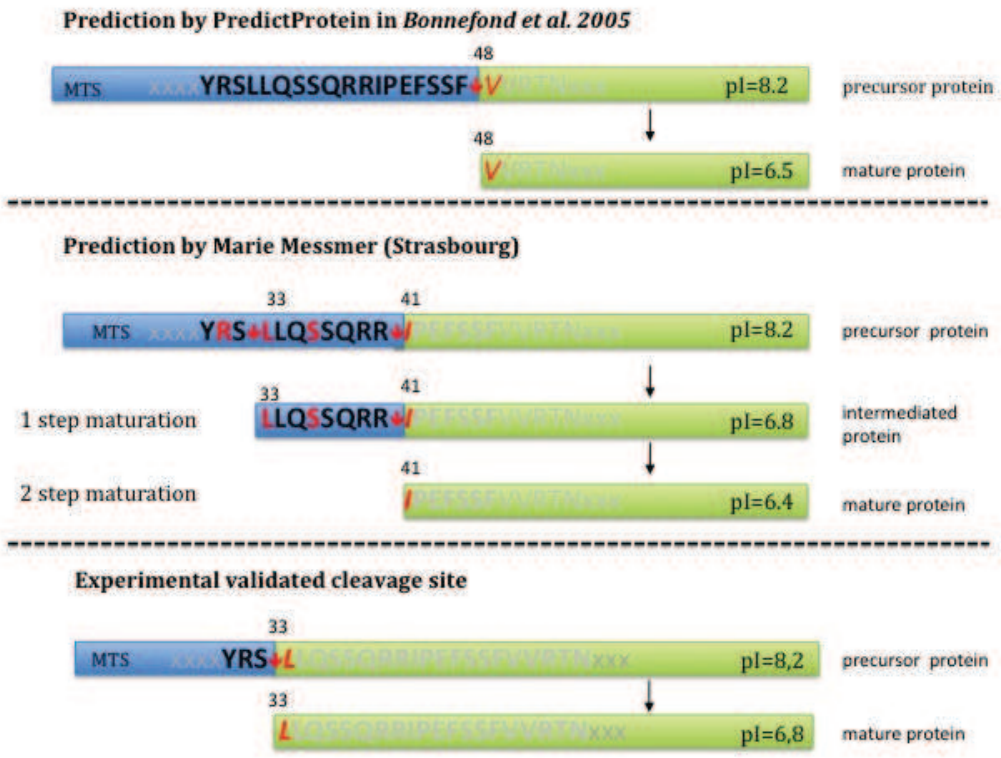


Figure 34: Schematized N-terminal cleavage sites of the mt-AspRS upon import into mitochondria. The corresponding amino acid sequence of the junction of the MTS (blue) and the "matured" protein (green) is illustrated and the corresponding pI is given. Red letters highlight important amino acids for cleavage. First amino acid of the cleaved matured protein are indicated by red Latin letters. Arrows indicate cleavage sites.

The difference in the predicted and experimental cleavage site suggest that different conditions exist *in vivo* stabilizing the protein, enhance the solubility and/or regulate the activity. Such conditions can be protein partners, cofactors and/or certain physiological conditions, which are not yet deciphered and thus cannot be mimicked *in vitro*. Another hypothesis is connected with the N-end rule of proteins describing the stabilization and destabilization of N-terminal amino acids (N-end rule; Bachmair et al., 1986). Of note, the experimentally validated form of mt-AspRS, starting at position L33, does not exclude that a minor population, starting at I41 exists *in vivo*. The observed discrepancies in their handling *in vitro* may correlate with their theoretical half-life [5,5h for mt-AspRS (L33) and 20h for mt-AspRS(I41)(Gonda et al., 1989)]. For example, one protease (Icp55) was identified in mitochondria to cleave pre-protein intermediates that carry a destabilizing N-terminal amino acid and to generate a stable mitochondrial protein (Vögtle et al., 2009). In addition, it has been hypothesized, that the N-end rule is a function of the cell's physiological state and could provide a mechanism for selective destruction of pre-existing, otherwise long-lived and undamaged proteins during cell differentiation, cell cycle progression, and other changes in the cell's state (reviewed in e.g. Mogk et al., 2007). Thus, we speculate that, in a normal state, cells keep the short life proteins as tool of regulation and in case of a physiological need such as high translation activity or starvation, the protein is further cleaved into a more stable protein with a higher half-life. Of note, these hypotheses are speculative and have to be investigated since the N-end rule pathways are not fully understood in mitochondria.

6 Mutational analysis of MTS of mt-AspRS and SOD2

We showed, by several experiments in HEK293 or BHK21 cells that the mt-AspRS is localized in both matrix/membrane and that SOD2 is only present in the matrix. In order to investigate the contribution of the MTS sequence to the final sub-mitochondrial localization, we exchanged MTSs from both proteins. At the time of this experiment, the cleavage site of SOD2 was known (Barra et al., 1984), and the one of mt-AspRS was solely predicted. We thus removed from the latter the first 49 amino

acids, compatible with the integrity of secondary structures. The two new constructions were named MTS^{SOD2}-AspRS and MTS^{Asp}-SOD2 (**Figure 35**). The corresponding constructs were cloned inside a vector suitable for the MVA-mediated internalization and expression in BHK21 cells. The expressed proteins were detected using antibodies against the co-expressed flag-epitope (**Figure 35**).

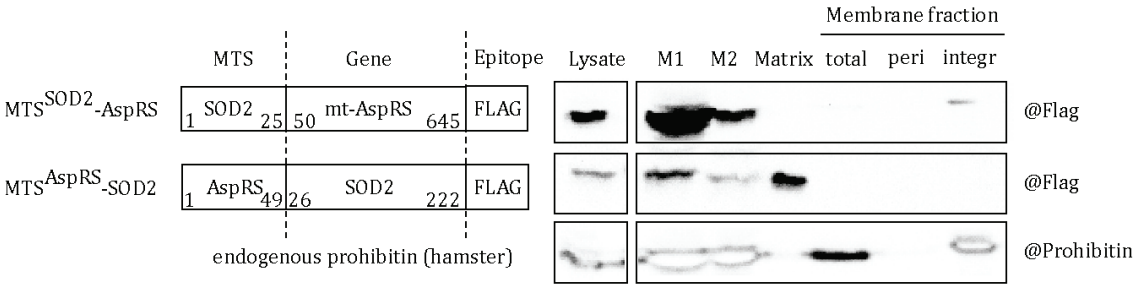


Figure 35: Expression of proteins with mutated MTS and detection in sub-mitochondrial fraction. Mutant constructs MTS^{SOD2}-AspRS and MTS^{AspRS}-SOD2 were expressed in BHK21 using vaccinia virus mediated transfer. Cells were lysed, mitochondria purified and fractionated. Samples were loaded on 10% SDS-PAGE and separated proteins were transferred to PVDF membrane. Recombinant proteins were detected using specific antibody against the co expressed Flag epitope. In addition the quality of fractionation was tested using an antibody against the endogenous hamster proteins prohibitin. Numbers indicate aa position corresponding protein. Peri stands for peripheral proteins and integr stands for integral proteins.

The expression and sub-mitochondrial distribution of the wild-type sequences for SOD2 and mt-AspRS were discussed in article #2. We deciphered their location in HEK cells and demonstrated that the modulated expression of recombinant proteins did not hamper their canonical distribution in BHK cells. As shown in **Figure 35**, exchanging the MTSs did not lead to the expected results. While MTS^{Asp}-SOD2 is internalized inside the mitochondria, it remained in the matrix. This indicated that the MTS^{Asp} sequence is not sufficient to lead to an additional membrane location. In addition, MTS^{SOD2}-AspRS is neither in the matrix nor in the membrane. Among all the possible explanations, it can be proposed that the chimeric MTS^{SOD2}-AspRS:

i) was not translocated as already observed for a pathology-related mutant of mt-AspRS (mt-AspRS S45G, see chapter V and article #3) or

ii) was not correctly designed so that the corresponding protein is not properly folded and aggregates

A possible further speculation is that both proteins are imported by two different mechanisms. And indeed, it exist two proposed mechanism of proteins import through TIM23. For loosely folded preproteins, a mechanism based on Brownian movement of the unfolded protein trough the translocase might be sufficient. Once the protein enters the matrix, it will be then trapped by mt-Hsp70 to prevent a backsliding of the preprotein (Pfanner and Geissler, 2001). This mechanism could be a mechanism for the small SOD2. In contrast, proteins with tightly folded domains require an additional pulling mechanism inside mitochondria. Of note, mt-AspRS is a highly structurally organized protein. In addition, it was stated that preproteins have to be recognized at several levels of import and therefore different signal are required at each level (Pfanner and Geissler, 2001, Becker et al., 2012). Thus, we hypothesize that both MTSs have more information than solely for the localization to mitochondria but also additional information for the translocation through the translocases. Consequently, it can be speculated that the MTS^{SOD2} is not sufficient to pull the bigger mt-AspRS through the inner membrane. In addition, the electric potential of the membrane is required to force the location of the protein trough the translocase towards the matrix. It is believed that the positively charged amino acid on the N-terminus form a "gradient" that is transmitted through the whole polypeptide chain (Huang et al., 2002, Prakash and Matouschek, 2004). This physiological property facilitates the unfolding and translocation through the membrane. This means, that the positive MTS is orientated towards the negative charge matrix side of the inner membrane. Of note, we predicted following pI changes for the mutants: MTS^{SOD2}-SOD2 (pI 8.4) to MTS^{Asp}-SOD2 (9.1); MTS^{Asp}-AspRS (pI 8.2) to MTS^{SOD2}-AspRS (pI 7,2). This hypothesis is speculative and need to further investigations.

In summary, we theoretically and experimentally investigated the N-terminal end of the mt-aaRSs with a particular focus on mt-AspRS. We showed that mt-aaRSs could be clustered into groups with high or low pI. We hypothesize that the pI could influence the cytosolic and mitochondrial properties of these proteins due to different overall net charges. A summary of previously published data and properties observed in this thesis are given in **Table 4**. It is obvious that only for the blue group, a correlation

with other properties can be drawn. No further correlation(s) between individual properties of the mt-aaRSs and the belonging to one of the group could be established.

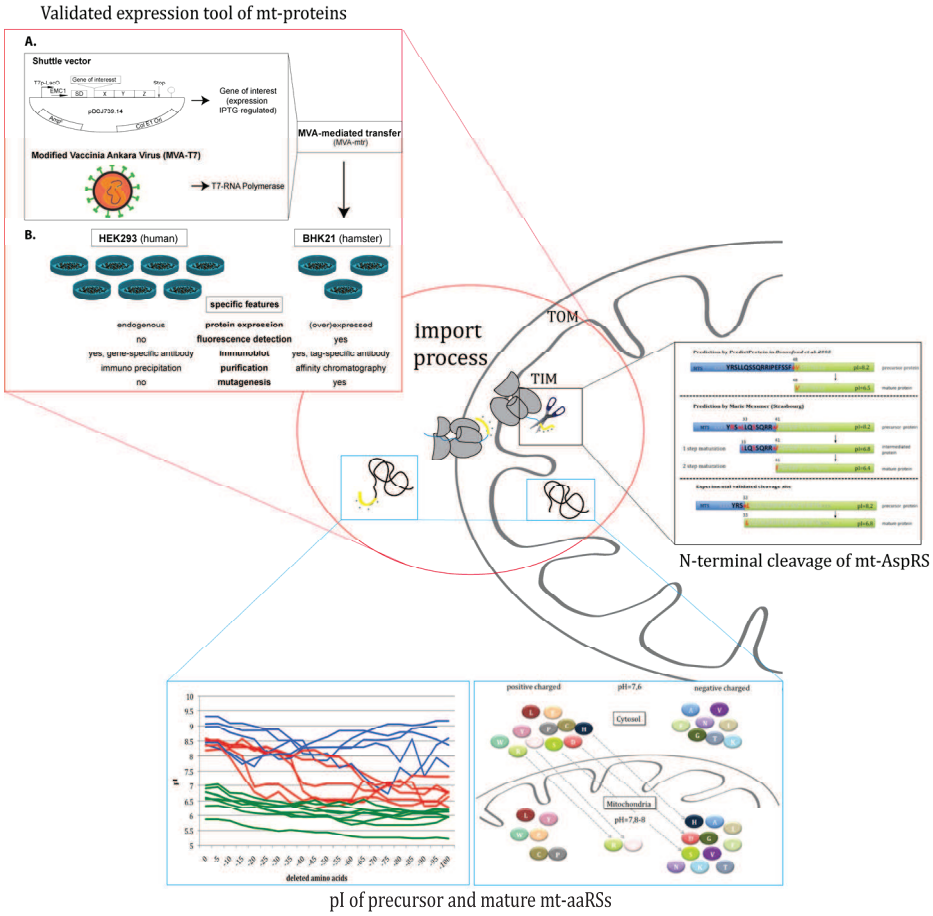
pI	Classical classification	motifs	origin	ΔpI bacterial	ΔpI cytosolic	MW	localization	disease
AlaRS	II	R-10	E(bacterial)	0	0	a	ns	y
AsnRS	II	R-10	P	1.5	0.5	b	mm	y*
GlyRS	II	R-10	E	0.5	0	a	0m	n
IleRS	I	R-10	P	0	0	a	m0	n
LysRS	II	R-10	E	1	0	a	mm	n
PheRS	II	R-3	P	1.5	-0.5	b	mm	y
ThrRS	II	R-3	E	0.5	0	a	mm	n
ValRS	I	R-2	E(bacterial)	1	-0.5	a	mns	n
AspRS	II	R-10	P	1.5	1.5	a	mm	y
CysRS	I	R-2	E	1.5	0	b	m0	n
HisRS	II	R-3	E	2	2	b	mm	y
LeuRS	I	R-10	P	2	0.5	a	mm	y**
ProRS	II	R-10	P	1.5	0.5	b	mm	n
SerRS	II	no	P	1.5	0.5	b	mm	y
ArgRS	I	no	E(bacterial)	2.5	1.5	b	mns	y
GluRS	I	no	P	2.5	0	b	mm	y
MetRS	I	R-10	P	2	2	b	mns	y
TrpRS	I	R-10	P	2.5	2	b	0m	n
TyrRS	Ib	R-10	P	3.5	2	b	mm	y

Table 4: Clustering of all known data in regard to the classification after the pI of the mt-aaRS. The three clusters of aaRSs are displayed in green, red and blue (corresponding to results in chapter 3). Properties, which correlate with the pI classification, are colored in the respective group color. Non-correlation are colored in brown. Classical classification and molecular weight are published in Bonnefond et al., 2005. Prediction corresponds to those proposed by Marie Messmer, origin of mt-aaRSs was provided from Marie Sissler (personal communication). Isoelectric point for bacterial and cytosolic homolog was extracted from Sissler et al. Correlation to disease was extracted from Schwenzer et al. 2013. * Disease correlation extracted from unpublished data (Chapter 5) and ** Pierce et al., 2013. Origine P and E correspond to prokaryotic and eukaryotic origin respectively. MW correspond to molecular weight and indicate if molecular weight is above (a) or below (b) the average (~60 kDa). Localization indicated by mm (matrix/membrane), m0(only matrix), 0m (only membrane) are indicated. Ns indicate no specific result available. Y and N correspond to yes and no, respectively.

7 Achievements at a glance

- We validated a protein expression tool, which mimics the natural occurring processing steps in mitochondria
- We experimentally deciphered the N-terminal cleavage site of mt-AspRS at distance from the predicted one
- We performed a theoretical investigation and revealed some peculiarities concerning pI of precursor protein and the matured proteins
- We found evidences indicating that the MTS is not solely responsible for the sub-mitochondrial distribution

8 Graphical Summary



Chapter IV: Biological partners of the human mt-AspRS

1 Introduction

In the previous chapter, our aim was to decipher the organization of the mt-aaRSs by i) establishing their sub-mitochondrial distribution (most of them were found in both the matrix and in the membrane fractions); and ii) trying to integrate this distribution into a functional network (with preliminary experiments aimed at correlating mt-aaRSs distribution with mt-tRNAs distribution). The purpose of this study was promoted by the broad knowledge acquired over the past decades on cytosolic aaRSs. With the discovery of alternative functions for cytosolic aaRSs, we know that organization of aaRSs can be a dynamic process depending on different localization and different protein interactions. As recalled in the introduction, the integration of nine of those cytosolic aaRSs into a macromolecular complex named MARS is well established (Mirande et al., 1982), as well as the involvement in this complex of three regulatory non-aaRS proteins (p18, p38 and p43). Also, the MARS complex is nowadays considered as a molecular reservoir ("depot hypothesis") where anchored proteins display canonical functions and undergo a functional switch towards non-translational function after release (Han et al., 2006; Ray et al., 2007).

Thus, while the organization of human cytosolic aaRSs is well explored and their involvement into alternate functions clearly established for some of them, the properties of the human mt-aaRSs remain scarcely known. An additional straightforward way to integrate aaRSs into novel functional networks and cellular activities is to nail down the involved cellular partners and potential interacting components. Therefore, the aim of the present chapter was to further explore the sub mitochondrial organization of the aaRSs. To do so, we searched for partner proteins and higher molecular complexes containing aaRSs,

following three major approaches:

I) To integrate the aaRSs into putative functional networks by data mining.
II) To investigate the involvement of aaRSs in macromolecular complexes by *in vivo* cross-linking and native gel electrophoresis.
III) To identify partner proteins by co-immuno precipitation.

We used the mt-AspRS as a "model" aaRS for three main reasons: (i) the established double distribution of this protein in the matrix and membrane could indicate different functions at different locations (ii) a missing link has been hypothesized concerning pathology related mutations in mt-AspRS and biological function (Scheper et al., 2007; Schwenzer et al., 2013) (iii) the human mitochondrial aspartylation has been studied for a long time in the laboratory so that all the requested materials (recombinant protein, specific antibodies ...), but also knowledge on the system, were available at the time of project initiation. The results presented here will serve as a foundation for further investigations with the remaining mt-aaRSs.

2 Data mining

2.1 STRING: Functional protein association network

Prior to experimental investigations, we used an online database resource Search Tool for the Retrieval of Interacting Genes (STRING 9.2 database, Franceschini et al., 2013) to predict functional association network. This database combines known and predicted interactions from direct and indirect proven datasets and provides an easy way to create association networks. The datasets are derived from genomic context, from high throughput experiments, from co-expression data and from published knowledge. The database connects all these data and predicts possible functional interaction networks. We therefore used this tool to predict the functional interaction networks of the mt-AspRS (DARS2) **Figure 36**. To do so, we performed a prediction using all available prediction parameters (neighborhood, gene fusion, co-occurrence, co-expression, experiments, databases, text-mining) and calculated the 50 most prevalent interaction partners with a confidence total score threshold of 0,700 (**Figure 36**A). Of note, the confidence score threshold gives a rough estimation of the confidence of the correlation between the two proteins, based on all of the available information predicting an association. We observed an accumulation of cytosolic and mitochondrial aaRSs in the close neighborhood of the mt-AspRS. Other proteins of the mitochondrial translation system (e.g. ribosomal proteins, initiation and elongation factors) were also predicted to be involved in this functional network. Since the prediction was performed with medium stringency (count first 50 protein hits with required interaction confidence score threshold 0,400) and the accuracy of text-mining tool appears doubtful, we increased the stringency of our prediction parameters.

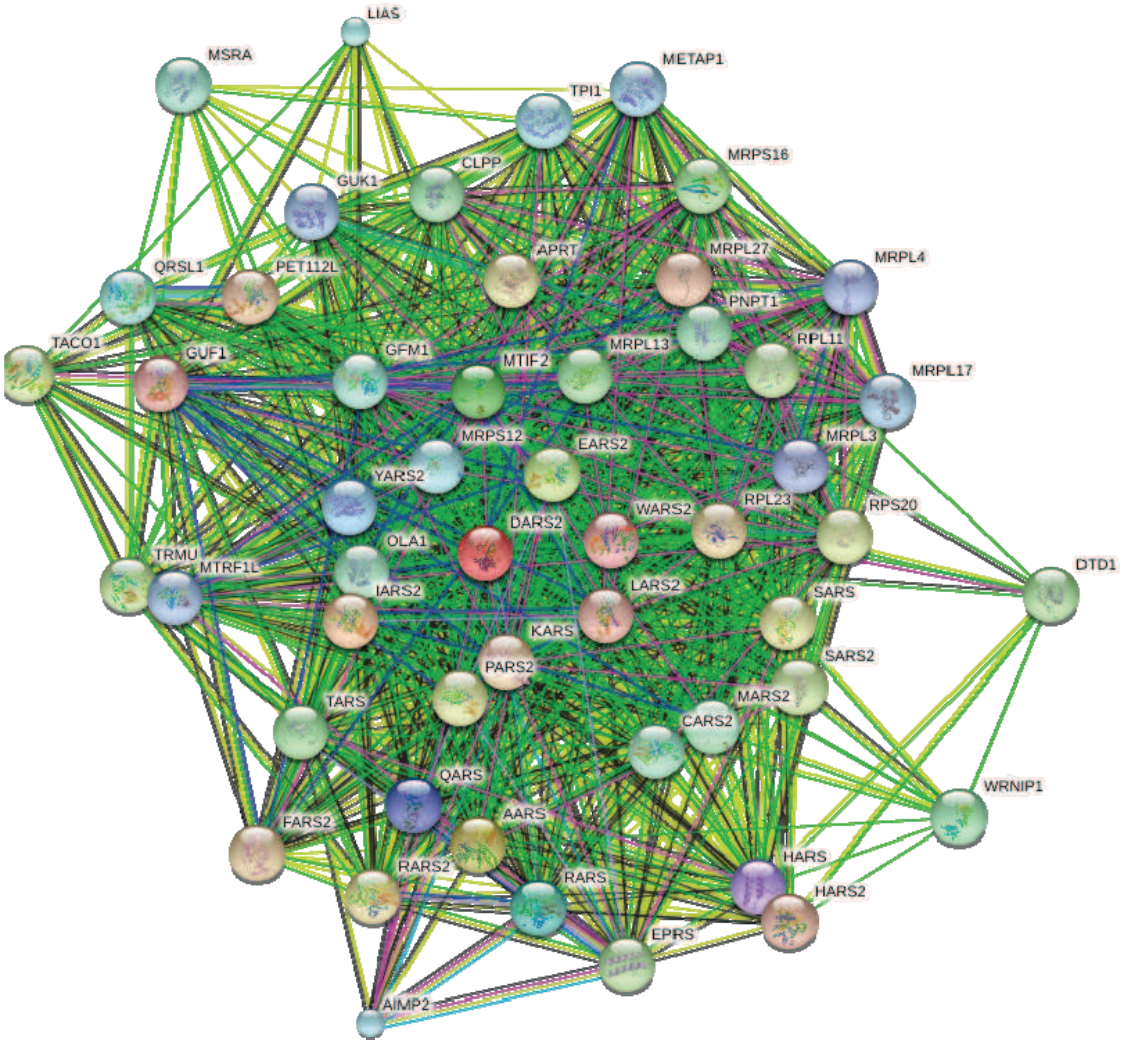


Figure 36A: STRING-Predicted functional interaction continued on the next page

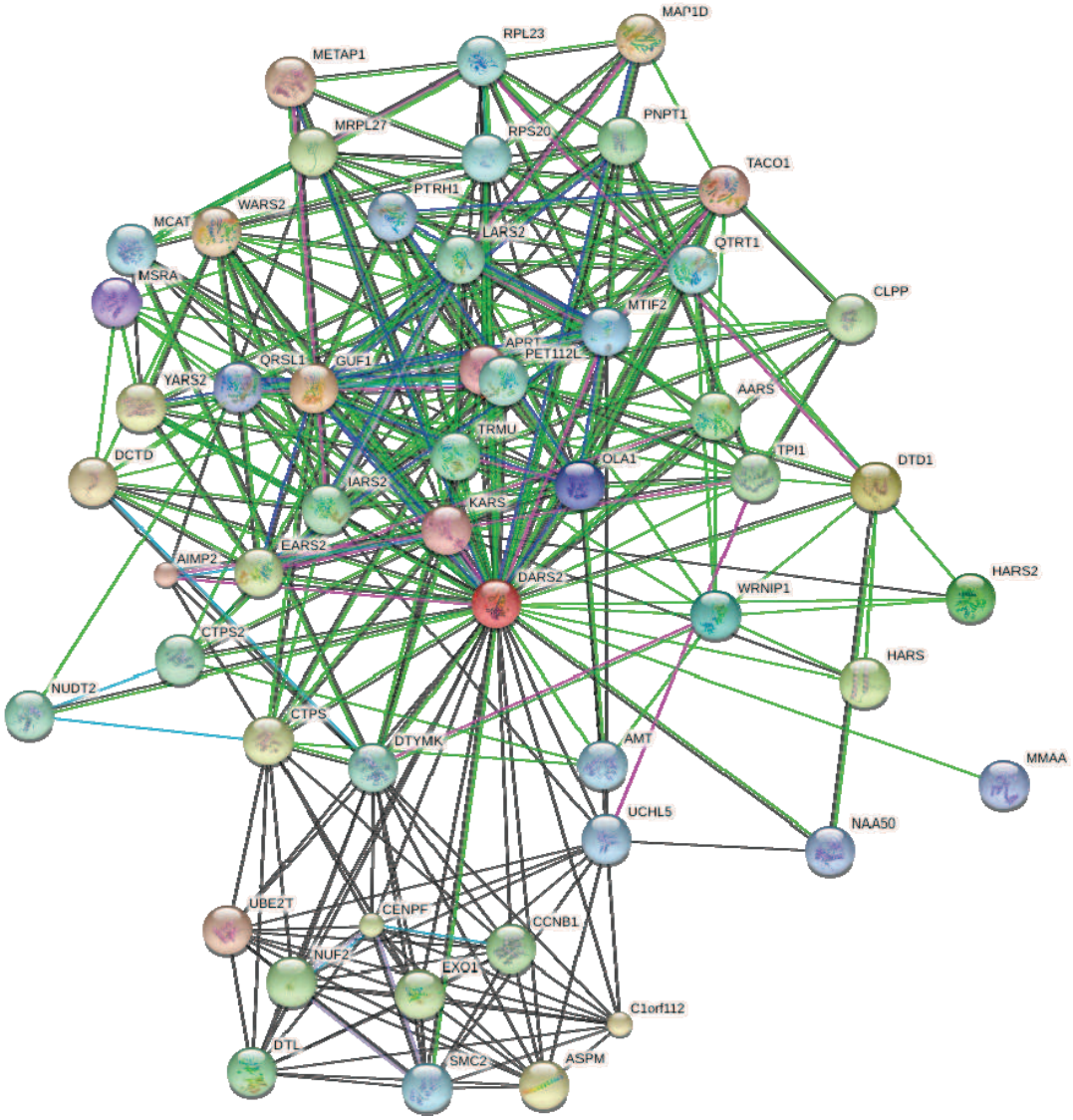


Figure 36B: STRING-Predicted functional interaction continued on the next page

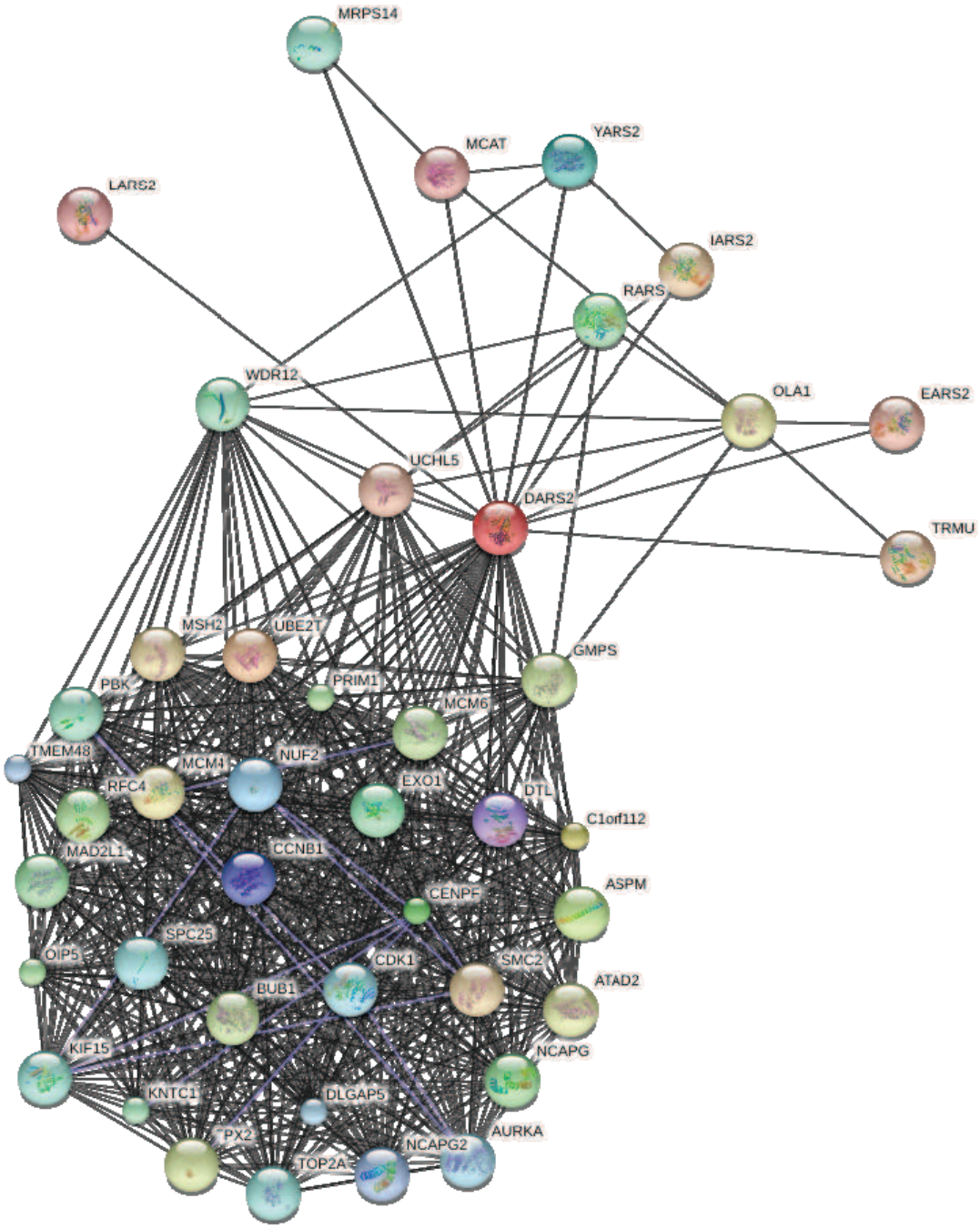


Figure 36C: STRING-Predicted functional interactions. A) Combined prediction methods (data from neighborhood, gene fusion, co-occurrence, co-expression, experimental, databases, text mining). B) Restricted prediction methods (without text mining) C) Co-expression based prediction

Concerning the text-mining tool, it appears that this method calculates the functional interaction network based on the co-appearance of two proteins in scientific text. Unsurprisingly, this approach frequently leads to false positives. For example, texts as following are taken for the calculation of the functional network: "[…] For mutations in two aminoacyl-tRNA synthetase genes, DARS2 and RARS2, normal or mild OXPHOS deficiencies in fibroblasts, respectively, have been reported, while [...] aciduria, Leigh-like encephalomyopathy, dystonia and deafness Mutant mitochondrial elongation factor G1 and combined oxidative phosphorylation deficiency Inherited mitochondrial [...] to function Defective mitochondrial translation caused by a ribosomal protein (MRPS16) mutation Impaired mitochondrial glutamate transport in autosomal recessive neonatal […]" (Rodenburg, 2011). Consequently, text-mining of this passage erroneously links DARS2 in a functional network with RARS2, EF-G1 and MRPS16. This is true concerning mitochondrial diseases but leads to wrong interpretation in regard to physical interaction networks.

To exclude these falses positifs, a second prediction excluding the text mining method was performed (**Figure 36**B). The observed interaction network is significantly reduced. In particular, the cytosolic aaRSs disappeared and a new group of proteins emerged as an individualized cluster of interactions (UBE2T, EXO1, NUF2, DT1, CENPF, SMC2, CCNB1, C1orf112 and ASPM). Of note, some of these proteins have a cell cycle activity. The possible relevance of their network interaction with the mt-AspRS is unclear.

In a next step, the STRING database allows one to solely display the co-expression data obtained from experimental data. Co-expression behaviors may shed light on functional correlations as it can be assumed that proteins forming a complex need to be express simultaneously to ensure the correct constitution of the putative complex. Thus, the STRING prediction parameters were set up to include only co-expression data (**Figure 34**C). It was observed that LARS2, YARS2, IARS2, EARS2 and RARS are co-expressed with DARS2. In addition, the mitochondrial ribosomal protein 14 and the mitochondrial tRNA modification protein TRMU are co-expressed. The previously observed cluster containing (UBE2T, EXO1 and ASPM) is further enriched in this analysis. A brief

functional clustering (DAVID ABCC 6.7 Huang et al., 2009a, 2009b) revealed that the majority of these genes are involved in cell cycle and chromosome maintenance.

In summary, three analyses were performed with an increasing stringency of parameters. The first analysis include all available prediction parameters (neighborhood, gene fusion, co-occurrence, co-expression, experiments, databases, text-mining), the second analysis exclude the text mining method and the third analysis only consider experimentally obtained co-expression data. Altogether, it appears that the co-expression data is the most informative analysis at this stage of investigation. Actually, we observed that three groups of proteins are co-expressed with mt-AspRS. The first group concerns other mitochondrial synthetase, the second group includes protein involved in mitochondrial translation (e.g. ribosomal proteins and tRNA modification proteins) and the last group contains proteins being involved in cell cycle or chromosome maintenance. We are not able to draw any conclusion from this data alone, but the involvement of the first and second group of proteins in the network of mt-AspRS is not surprising, since they all play an important role in the well-defined process of mitochondrial translation. The role of the third groups and their functional connection to mt-AspRS remains unclear.

2.2 Data mining of existing high-throughput data

In addition to bioinformatics prediction of a functional interaction network, experimentally derived data were extracted from BioGrid3 database or from IntAct (**Table 5**). BioGrid3 database reported seven potential interaction partners of mt-AspRS obtained by Two-Hybrid and affinity capture methods whereas IntAct reported three partner proteins obtained by Two-Hybrid methods or by Tandem Affinity purification.

All of these proteins are of cytosolic or nuclear localization. P29322, Q9UBS5 are receptor proteins located in the plasma membrane. P0CG48 is involved in protein degradation and is localized in both cytoplasm and nucleus. Q9BPU6 is connected with neuronal differentiation and localized in cytosol. The Q15554 is a protein involved in

telomere maintenance and responsible for recruiting of protein factors. Q9BPU6 is connected with neuronal differentiation and localized in cytosol.

UniProt	Name	Method	Reference
BioGrid3.2			
Q13155	AIMP2	Two-hybrid	Quevillon et al., 1999
Q9BPU6	DPYSL5	Two-hybrid	Wang et al., 2011
P29322	Ephrin type-A receptor 8	Two-hybrid	Wang et al., 2011
Q15046	Lysyl-tRNA synthetase	Two-hybrid	Quevillon et al., 1999
Q15554	telomeric repeat binding protein 2	Affinity Capture-MS	Giannone et al., 2010
P0CG48	Polyubiquitin-C	Affinity Capture-MS	Danielsen et al., 2011
P0CG48	Polyubiquitin-C	Affinity Capture-MS	Wagner et al., 2011
P0CG48	Polyubiquitin-C	Affinity Capture-MS	Povlsen et al., 2012
Q9UBS5	Gamma-aminobutyric acid type B receptor subunit 1	Affinity Capture-MS	Aichem et al., 2012
IntAct			
Q9BPU6	Dihydropyrimidinase-related protein 5	Two Hybrid array	Wang et al., 2011
P29322	Ephrin type-A receptor 8	Two Hybrid array	Wang et al., 2011
Q8JPQ9	Non-structural protein NS-S	Tandem Affinity purification	Pichlmair et al., 2012

Table 5: Interactome of mt-AspRS. Summary of reported interaction proteins reported in BioGrid3.2 and in InterAct.

So far, the lysyl-tRNA synthetase is the only putative mt-AspRS interacting partner that can be of mitochondrial location. However, these data date from 1999 (Quevillon et al.,1999) and the mt-AspRS was first annotated in 2005 (Bonnefond et al., 2005). Thus, these data need to be treated with caution, as high-throughput analyses are not always manually corrected.

So far, no clear connection between the data obtained by predicted functional association analysis or by analysis of high-throughput data can be drawn. As described before, the predicted interaction networks are only one brick towards the discovery of new partners of the mt-AspRS and need to be validated by experimental data.

3 Identification of putative partner proteins of mt-AspRS

3.1 Involvement of mt-AspRS in macromolecular complexes

A part of the cytosolic aaRSs is organized into a macromolecular complex consisting of nine aaRSs and three non-aaRSs components with an approximate size of 1.5MDa (Mirande et al., 1982). In addition, data obtained by size exclusion chromatography suggest the involvement of the mt-TyrRS in a macromolecular complex of about 1MDa (Van Berge et al., 2012a; Sasarman et al., 2012; Schägger et al., 1994). Those evidences prompted us to investigate the presumed involvement of mt-AspRS in such a macromolecular complex. To do so, we used Blue Native Polyacrylamide Gel electrophoresis (BN-PAGE), which was established in 1991 to purify membrane protein complexes in a native form (Schägger and Von Jagow, 1991). Since we discovered that a fraction of the mt-aaRSs is organized on the mitochondrial inner membrane (Chapter II), we decided to apply this method to reveal possible-membrane bound complexes.

In a preliminary experiment, we tested the practicability of this method under classical experimental conditions. The dimeric state of the recombinant mt-AspRS was verified previously (Bonnefond et al., 2005, van Berge et al., 2012) but never confirmed using native protein extracted from human cells. The samples were prepared as described in Materials and Methods. Different amounts (1.8ng-114ng) of recombinant mt-AspRS and protein samples extracted from HEK293 cells were separated under native conditions on the BN-PAGE (4-16%). As a control, one sample of the recombinant mt-AspRS (114ng) was prepared under denaturizing condition (with SDS and ß-Mercaptoethanol) and separated. After the separation, the proteins were transferred to a PVDF membrane and the mt-AspRS was detected with specific antibodies. The result of this experiment is shown in **Figure 37**. The major fraction of recombinant mt-AspRSs (left panel) was found at a size of approximately 150kDa. A second major band (66kDa) was observed, for amounts higher than 14ng of loaded recombinant mt-AspRS. Both sizes correspond to the dimeric and monomeric form of the mt-AspRS, respectively. Aggregations (>242kDa) were visible at higher concentrations. In the right panel, we tested the oligomeric state of the mt-AspRS

separated from native cellular extract. We observed a major band of 150kDa corresponding to the dimer of mt-AspRS. This band was shown in several experiments to be the dominant form in native cellular extract.

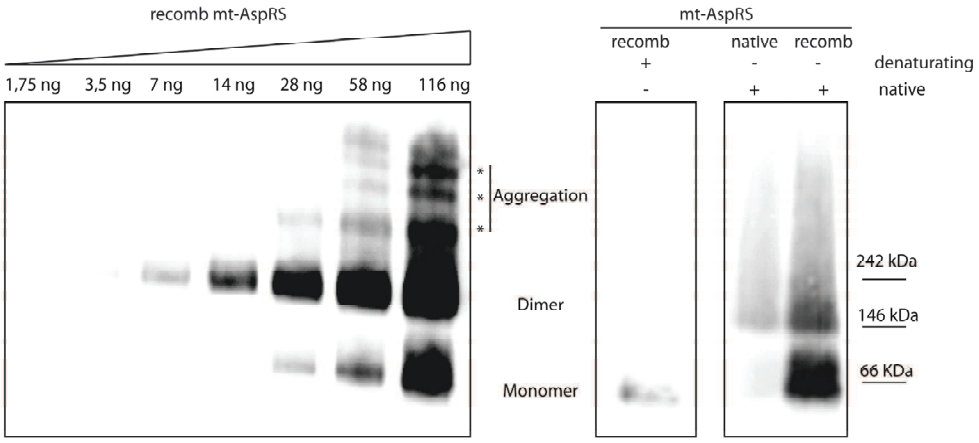


Figure 37: Identification of the oligomeric status of the native and recombinant mt-AspRS by BN-PAGE. Recombinant protein at different concentrations or cellular lysate was separated under native conditions on a BN-PAGE and proteins were transferred to a PVDF membrane. In addition, one sample of recombinant protein prepared under denaturizing condition was separated and transferred. The mt-AspRS was detected using a specific antibody. Left panel shows different concentration of recombinant protein. Mono- and dimeric forms are indicated. Asterisks indicate aggregations. Right panel show the denaturated recombinant protein and the separation of the native and recombinant protein under native conditions. Molecular sizes are indicated.

In the native sample we observed also a band above 1000kDa (not shown in **Figure 37**) indicating the possible presence of mt-AspRS in a macromolecular complex. To further investigate this observation, and to stabilize the possible complex, we performed formaldehyde cross-linking experiments. To do so, we treated cells with different concentrations of formaldehyde (0,2%-1%) and with different times of incubation (0h-1h). The results are displayed in **Figure 38**. A slight additional band around 480kDa (**Figure 38**, left) appeared with an increased amount of Formaldehyde. In addition, the intensity of the previously observed band above 1000kDa seems to be independent of formaldehyde concentration. Instead, we observed that this band was sensitive to the incubation time with formaldehyde (**Figure 38**, right).

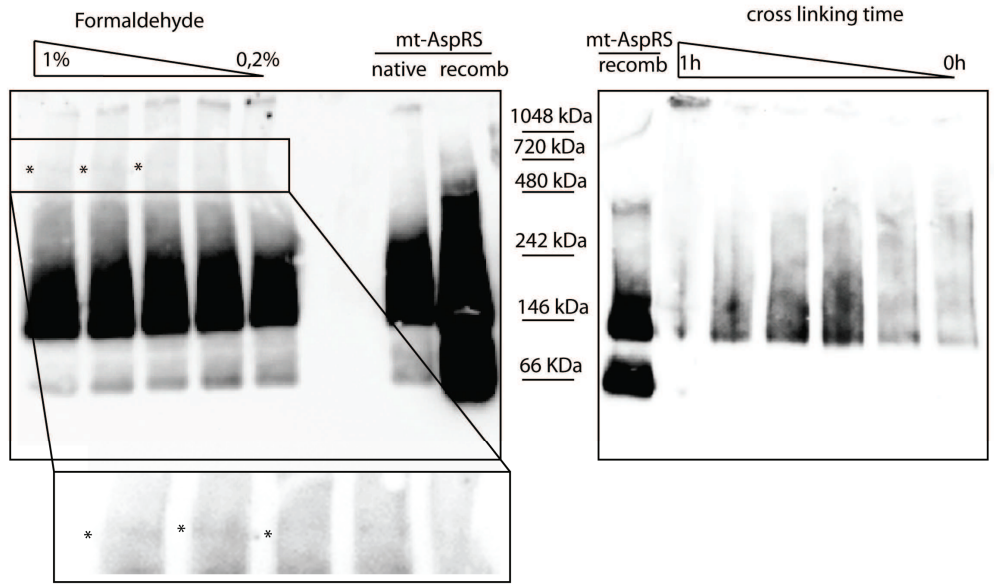


Figure 38: Cross-linking of mitochondrial proteins and subsequent separation on BN-PAGE. Mitochondrial proteins were cross-linked using formaldehyde, separated on BN-PAGE and transferred to a PVDF membrane. Two cross-linking conditions are indicated. Left panel shows the cross-linking efficiency depending on the formaldehyde concentration and right panel shows the cross-linking efficiency depending on the time using 1% Formaldehyde. A zoom inset highlights observed higher molecular bands of an approximately size of 480kDa. Asterisks indicate the location of these bands. Molecular sizes are indicated.

In summary, we found evidence for two higher molecular weight mt-AspRS-containing complexes of approximately 480kDa and >1000kDa size. To further characterize these complexes, we investigated their existence in the different sub mitochondrial fractions. To do so, we purified mitochondria and incubated them with or without formaldehyde. Subsequently, we prepared sub mitochondrial fractions containing membrane and matrix proteins (as defined in chapter II). These fractions were separated on BN-PAGE and transferred to PVDF membrane as described before (**Figure 39**, top). As control, samples were prepared under denaturizing conditions and separated on a SDS-PAGE and transferred to PVDF membrane (**Figure 39**, bottom). By immuno detection, bands of 146kDa, corresponding to the mt-AspRS dimer, were detected in lysate, mitochondria before (Mito1) or after (Mito2) purification on a percol gradient, cross-linked membrane and non cross-linked matrix fraction. In addition, in the non-cross linked membrane fraction and in the cross-linked matrix fraction no clear bands were visible at

~146kDa. Furthermore, a strong smear in the cross-linked matrix fraction likely implied a higher molecular complex. In the cross-linked membrane fraction two light additional bands were observed at 242kDa and 480kDa. The latter one is of similar size to the previously observed band (**Figure 38**, left). No macromolecular complex of over 1000kDa was observed. The unexpected migration pattern of the non-cross linked membrane fraction cannot be interpreted. Unfortunately, this particular experiment could not be reproduced in a satisfying way to confirm the obtained results. Nevertheless, these experiments provide some evidence that mt-AspRS could form a complex with other proteins. To identify these proteins, the mt-AspRS was immuno precipitated, and co-purified putative partner proteins were identified by mass spectrometry.

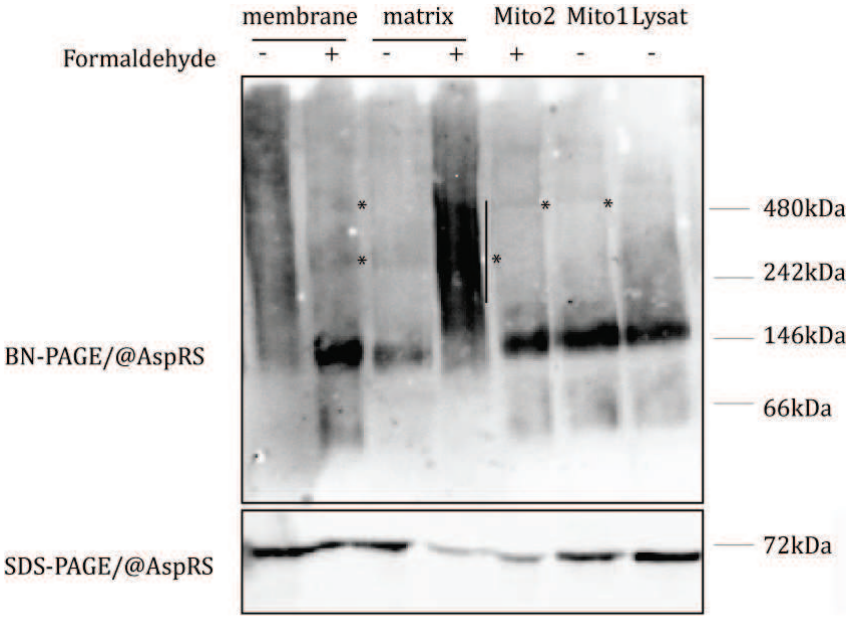


Figure 39: Cross-linking of mitochondrial proteins and subsequent fractionation of matrix and membrane fractions. Purified mitochondria were cross-linked (using 1% Formaldehyde) or not and subsequently fractionated. The obtained fractions were separated on BN-PAGE (top panel) and proteins were transferred PVDF membrane. Mitochondrial AspRS were detected using a specific antibody. In lower panel, one sample of each fraction were prepared under denaturizing conditions and separated on 10% SDS-PAGE and transferred to PVDF membrane. Asterisks indicate the location of putative higher molecular complexes containing mt-AspRS. Molecular sizes are indicated.

3.2 Identification of putative protein partners

To identify partner proteins of mt-AspRS, we performed co-immuno precipitation experiments (co-IP) with specific antibodies against the mt-AspRS. The principle of immuno precipitation is described in materials and methods and briefly recalled here and schematized in **Figure 40**. We coupled antibodies specifically targeted against mt-AspRS with agarose beads containing protein A (agarose-A beads). These beads were incubated with total or mitochondrial lysate extracted from HEK293 cells under native conditions. After several washing steps, the proteins that co-eluted with mt-AspRS were analyzed by LC-MS/MS.

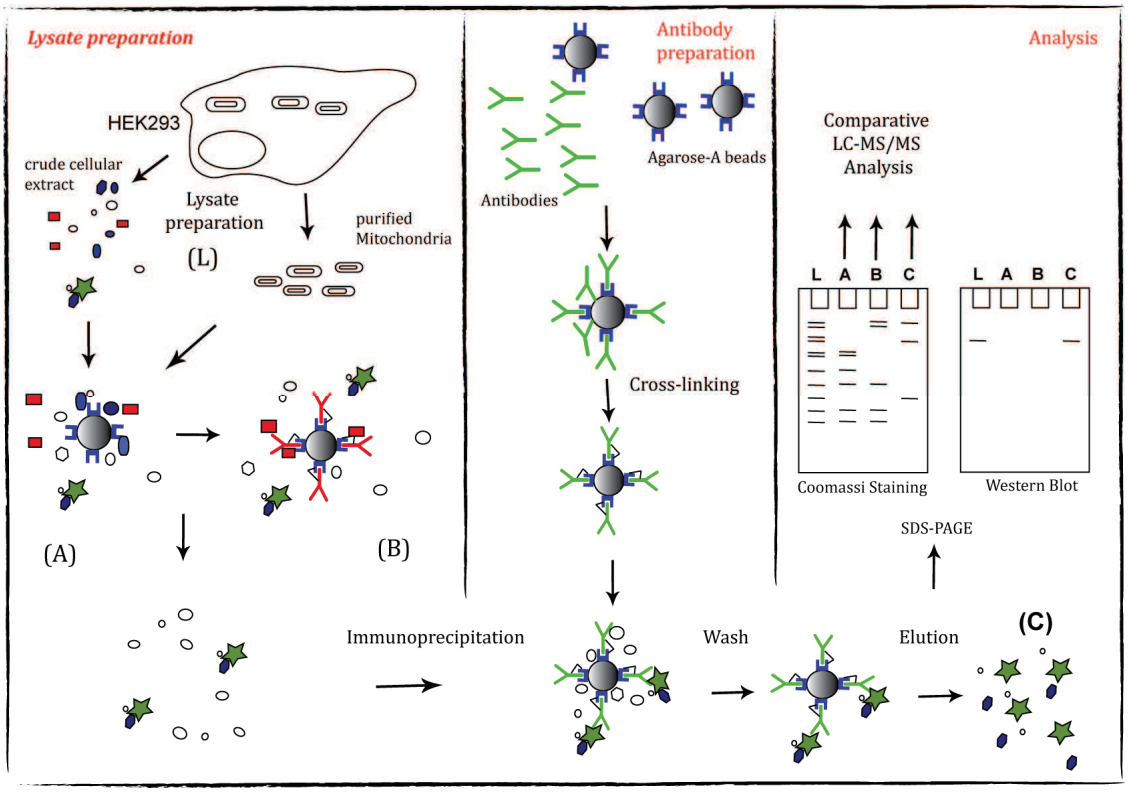


Figure 40: Schematized flowchart of co-immuno precipitation (co-IP). The lysate preparation, the antibody preparation and the analysis of co-IP are displayed. Lysate preparation (L) includes the cell lysis with or without mitochondria purification. The pre-clearing step of the lysate with agarose-A beads (A) and with beads on which unspecific IgG are coupled (B). Antibody preparation includes the cross-linking of specific antibodies on the agarose-A beads and a washing step. The pre-cleared lysate is incubated with the prepared specific antibodies covalently linked to agarose-A beads. After several washing steps, the captured proteins are released from the antibody bound to the agarose-A beads. Eluted proteins (C) are separated on SDS-PAGE and subsequent subjected to western blot or LC-MS/MS analysis.

Four experiments were independently leading to four sets of data. The two first (IP1 and IP2) were performed using either total cellular extracts (where cells were lysed and directly submitted to immuno-precipitation), and the two last (IP3 and IP4) were performed using lysates obtained from purified mitochondria. In parallel to each IP, one control was performed. To do so, the same volume of lysate was incubated with only the agarose beads and subsequently treated like the IP. After elution of the proteins, both the control mixture and the IP protein mixtures were separated on a 10% SDS-PAGE. Gels were either subjected to western blot detection or stained with coomassie blue. For comparative analyzes, only those bands that were visible in the lane of the IP were sliced out. As control, the area corresponding to the control lane was also sliced out and submitted to LC-MS/MS analysis. The summary of the number of obtained proteins is shown in **Table 6**.

Number of Identified Proteins	IP1	IP2	IP3	IP4
Total IP	50	75	1003	1949
Control	11	61	718	1372
IP	39	14	285	577
Corrected IP	24	6	36	11
Sliced Bands	16	30	-	-

Table 6: Summary of identified proteins by co-IP experiments. "Total IP" numbers of proteins represent the uncorrected amount of proteins obtained from the co-IP (containing proteins e.g. present in control and IP). The numbers of proteins identified in "control" are indicated. The numbers of identified proteins only present in the "IP" are indicated. The numbers of identified proteins in the "corrected IP" is manually corrected and redundant proteins are removed. Final numbers of identified proteins are illustrated in red. In addition, in IP 3 and IP4 only proteins with a mascot score above 35 are present in the "corrected IP" list. Of note, only the proteins in sliced out bands were analyzed. Proteins obtained in IP3 and IP4 were not separated and the complete protein extract obtained after elution were analyzed.

The summary of identified proteins obtained per IP experiment is given in **Table 7**. For each IP, proteins found in the control experiment ('control') were subtracted from those initially found ('total'). Remaining candidates ('IP') were then manually analyzed so that to keep only those ('cleaned IP') with a Mascot scores ≥ 35.

Uniprot-ID	Protein name	Size (aa)
IP1-total lysate		
ACLY_HUMAN	ATP-citrate synthase	1,101
ACTG_HUMAN	Actin, cytoplasmic 2	375
CH60_HUMAN	60 kDa heat shock protein, mitochondrial	573
DHSA_HUMAN	Succinate dehydrogenase [ubiquinone]	664
EF1A1_HUMAN	Elongation factor 1-alpha 1	462
ENOA_HUMAN	Alpha-enolase	434
GRP75_HUMAN	Stress-70 protein, mitochondrial	679
HS90B_HUMAN	Heat shock protein HSP 90-beta	724
HSP71_HUMAN	Heat shock 70 kDa protein 1A/1B	641
IF4A1_HUMAN	Eukaryotic initiation factor 4A-I	406
IGKC_HUMAN	Ig kappa chain C region	106
LDHB_HUMAN	L-lactate dehydrogenase B chain	334
PIMT_HUMAN	Protein-L-isoaspartate(D-aspartate) O-methylt...	227
TBB5_HUMAN	Tubulin beta chain	444
TCPZ_HUMAN	T-complex protein 1 subunit zeta	531
VIME_HUMAN	Vimentin	466
Q7Z759_HUMAN	CCT8 protein	497
Q8N532_HUMAN	TUBA1C protein	325
Q8N6N5_HUMAN	Tubulin, beta 2C	445
Q14222_HUMAN	EEF1A protein	227
Q5ZEY3_HUMAN	Glyceraldehyde-3-phosphate dehydrogenase	86
Q59EJ3_HUMAN	Heat shock 70kDa protein 1A variant	709
Q0QEN7_HUMAN	ATP synthase subunit beta	445
Q53HF2_HUMAN	Heat shock 70kDa protein 8 isoform 2 variant	49
IP2-total lysate		
FUBP2_HUMAN	Far upstream element-binding protein 2	711
PRDX4_HUMAN	Peroxiredoxin-4	271
Q96RE1_HUMAN	Elongation factor 1-alpha	398
Q53SY7_HUMAN	Putative uncharacterized protein CAD	2,151
Q59EJ3_HUMAN	Heat shock 70kDa protein 1A variant	709
Q53HF2_HUMAN	Heat shock 70kDa protein 8 isoform 2	493
IP3-mitochondrial lysate		
CNKR2_HUMAN	Connector enhancer of kinase suppressor of ra...	1,034
DAZP1_HUMAN	DAZ-associated protein 1	407
DDRGK_HUMAN	DDRGK domain-containing protein 1	314
FAKD5_HUMAN	FAST kinase domain-containing protein 5	764
FSIP2_HUMAN	Fibrous sheath-interacting protein 2	6,907
GEPH_HUMAN	Gephyrin	736
GIMA8_HUMAN	GTPase IMAP family member 8	665
GNPAT_HUMAN	Dihydroxyacetone phosphate acyltransferase	680
GOGA4_HUMAN	Golgin subfamily A member 4	2,230
IGHG1_HUMAN	Ig gamma-1 chain C region	330
IQGA1_HUMAN	Ras GTPase-activating-like protein IQGAP1	1,657
LAMP2_HUMAN	Lysosome-associated membrane glycoprotein 2	410

LCN1_HUMAN	Lipocalin-1	176
MYH14_HUMAN	Myosin-14	1,995
NSDHL_HUMAN	Sterol-4-alpha-carboxylate 3-dehydrogenase	373
PTCD3_HUMAN	Pentatricopeptide repeat domain-containing protein	689
QCR1_HUMAN	Cytochrome b-c1 complex subunit 1	480
RAB1B_HUMAN	Ras-related protein Rab-1B	201
RL11_HUMAN	60S ribosomal protein L11	178
RL14_HUMAN	60S ribosomal protein L14	215
RL19_HUMAN	60S ribosomal protein L19	196
RL34_HUMAN	60S ribosomal protein L34	117
RL8_HUMAN	60S ribosomal protein L8	257
RM16_HUMAN	39S ribosomal protein L16, mitochondrial	251
SAC1_HUMAN	Phosphatidylinositide phosphatase SAC1	587
SYIC_HUMAN	Isoleucine--tRNA ligase, cytoplasmic	1,262
SYLC_HUMAN	Leucine--tRNA ligase, cytoplasmic	1,176
SYMC_HUMAN	Methionine--tRNA ligase, cytoplasmic	900
SYNE2_HUMAN	Nesprin-2	6,885
SYRC_HUMAN	Arginine--tRNA ligase, cytoplasmic	660
TCPH_HUMAN	T-complex protein 1 subunit eta	543
TMED9_HUMAN	Transmembrane emp24 domain-containing protein...	235
TR61B_HUMAN	tRNA (adenine(58)-N(1))-methyltransferase, mi...	477
TRFL_HUMAN	Lactotransferrin	710
U17L6_HUMAN	Ubiquitin carboxyl-terminal hydrolase 17-like	398
UFL1_HUMAN	E3 UFM1-protein ligase 1	794
IP4-mitochondrial lysate		
ARI3A_HUMAN	AT-rich interactive domain-containing protein...	593
CC124_HUMAN	Coiled-coil domain-containing protein 124	223
CDK1_HUMAN	Cyclin-dependent kinase 1	297
KIF1A_HUMAN	Kinesin-like protein KIF1A	1,690
MPRIP_HUMAN	Myosin phosphatase Rho-interacting protein	1,025
ORNT1_HUMAN	Mitochondrial ornithine transporter 1	301
PABP1_HUMAN	Polyadenylate-binding protein 1	636
RT18A_HUMAN	28S ribosomal protein S18a, mitochondrial	196
SCMC3_HUMAN	Calcium-binding mitochondrial carrier protein...	468
TM177_HUMAN	Transmembrane protein 177	311
VDAC3_HUMAN	Voltage-dependent anion-selective channel pro...	283

Table 7: Summary of identified proteins from four different co-IPs. The UNIPROT_ID, the full name and the size (in aa) are displayed in the table. In blue: proteins of annotated cytosolic location; in green: proteins of annotated mitochondrial location; in purple: proteins of annotated endoplasmic reticulum location; and in black: proteins with unknown location.

Concerning the immuno-precipitation experiments IP1 and IP2 performed using cellular extracts, 24 and 6 proteins were identified, respectively. To further estimate the relevance of each of these proteins, we correlated the Mascot score value with the abundance of the

corresponding protein. Thus, we interpreted that high scores reflect a high quantity of a protein (**Figure 41**). As illustrated in this figure, the peptide coverage for mt-AspRS is very low in IP1, indicating that this experiment lead to a poor enrichment. On the contrary, we observe high peptide coverage of mt-AspRS for IP2, indicating a successful enrichment. As already mentioned, the two experiments were performed using crude cellular extracts, likely explaining the high recovery of cytosolic proteins (**Table 7**). Among these proteins, heat shock proteins, metabolic proteins and proteins involved in translation build the major clusters. From 24 proteins in IP1, five proteins were annotated of mitochondria localization, while in IP2 no mitochondrial proteins were retrieved.

Facing these unsatisfactory results, IPs were performed using lysates extracted from purified mitochondria (IP3 and IP4). Several pre-clearing steps were added to reduce non-specific interactions. To do so, the mitochondrial lysate was sequentially incubated with free agarose A beads (ctrl 1), with agarose beads coupled with non-specific IgG's (ctrl 2) before being incubated with beads coupled with specific antibodies against mt-AspRS. After incubation, the beads were washed and the proteins eluted. The eluted proteins were separated on a 10% SDS-PAGE. Instead of cutting only specific gel bands, we analyzed complete set of separated proteins. To do so, we systematically sliced out the lanes and performed a comprehensive analyzes by LC-MS/MS. The datasets were analyzed as described before. In total, we observed more than 1000 proteins for each experiment. Most of them were also identified within the control experiments (ctrl1+2) and thus not further considered.

We observed 36 possible partner proteins in IP3 and 11 in IP4. High mascot scores for the mt-AspRS indicated a high enrichment and high quality of both IPs (**Figure 41**). In IP3, we recovered 18 proteins from the cytosol, eight from the mitochondria, and five from the endoplasmic reticulum. The locations of six proteins have not yet been annotated. Of note, 30 % of the proteins are involved in translation. Among these proteins, cytosolic and mitochondrial ribosomal proteins and cytosolic aaRSs were identified. Also, one mitochondrial protein involved in mt-tRNA modification and one protein of the respiratory chain were identified. In IP4, we pulled down four proteins from the cytosol, five from the mitochondria and two of unknown localization. Of note, three of the mitochondrial proteins

are membrane proteins and have a function as membrane channel or carrier. In addition, one mitochondrial ribosomal protein was detected.

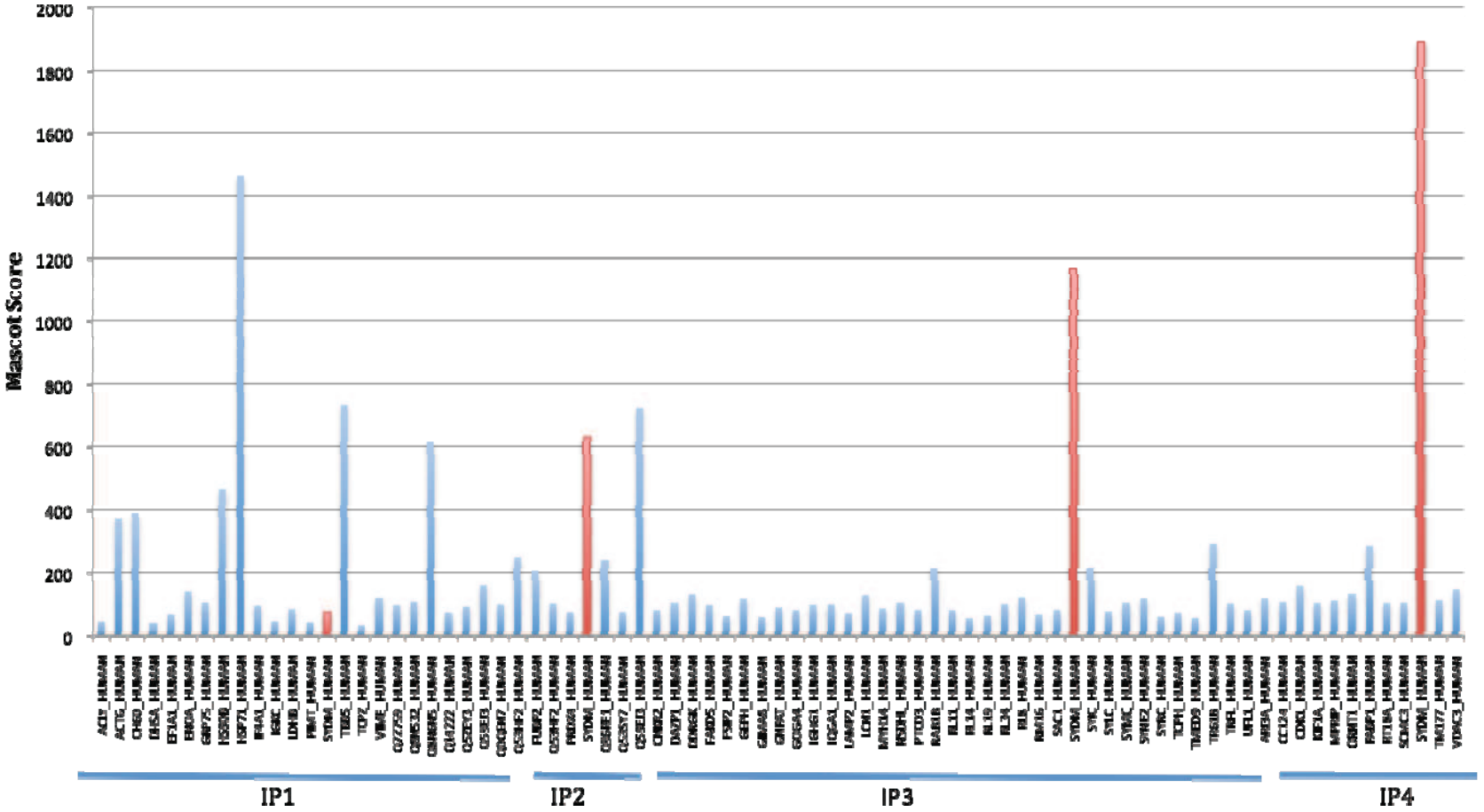


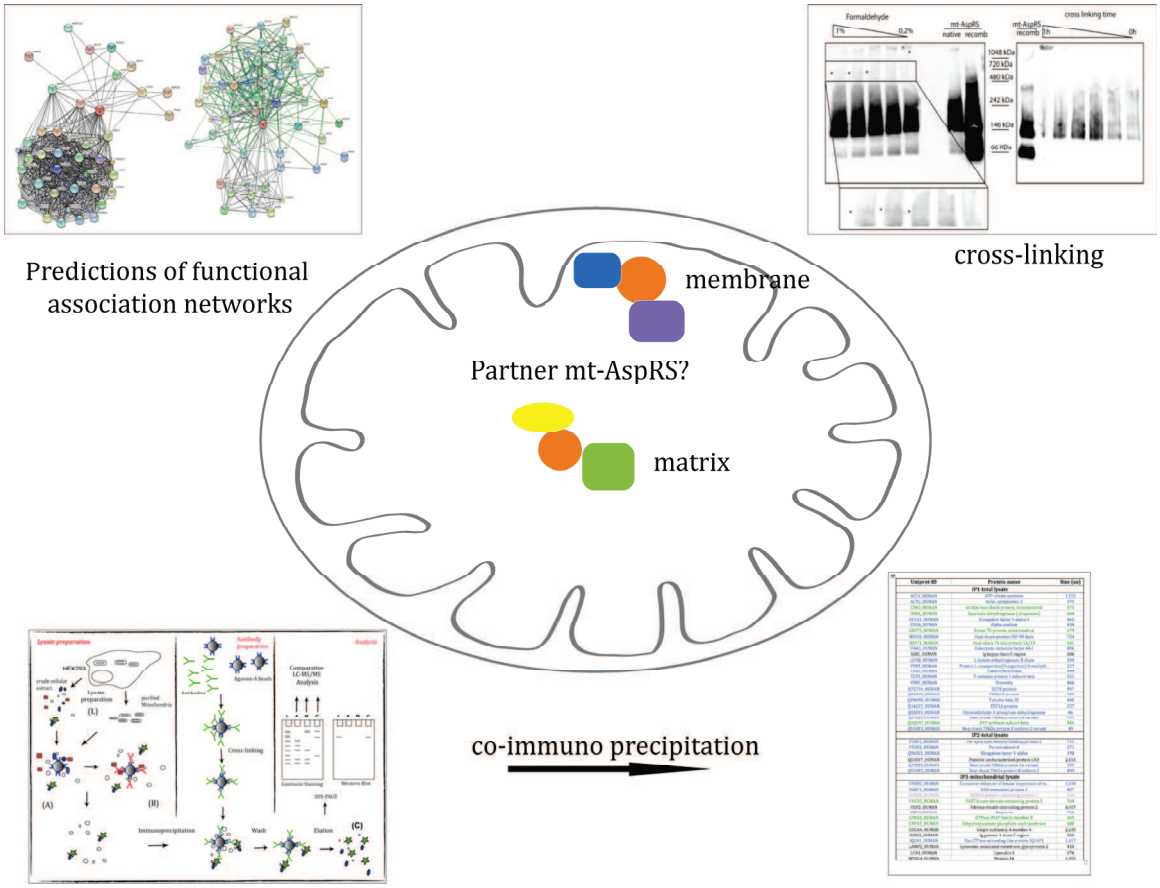
Figure 41: Mascot scores of the identified proteins. Proteins are sorted after the experiment. The mascot scores of the mt-AspRS for each experiment are highlighted in red.

In summary, despite the fact that the quality of the experiments was clearly improved (as reflected by the mt-AspRS enrichment), we could not identify one or more candidates appearing in all 4 independent IPs. Nevertheless, some interesting candidates were obtained promoting further investigations. First, we found in nearly all IPs proteins of the translation machinery of the cell. Second, we identified in two experiments mitochondrial ribosomal proteins as well as proteins of the inner mitochondrial membrane or mitochondrial respiratory chain.

4 Achivements at a glance

- We predict a functional association network of mt-AspRS
- We identify involvement of mt-AspRS in a putative complex of >1000kDa
- We achieve a set of putative interaction partners by co-IP

5 Graphical summary



Chapter V: Disorders of the human mitochondrial aminoacylation system

Introduction

Translational medicine is a field in medical research connecting basic biomedical science (e.g. molecular and cell biology) with the medical needs of patients. It describes a multi-disciplinary collaboration between medical doctors and researchers in order to make progress in the molecular definition and mechanistic elucidation of disease, with the aim of developing effective therapies.

Mitochondrial disorders were defined as pathologies with aberrant oxidative phosphorylation (OXPHOS). They were described in the late 80's as exclusively related to mutations within the mt-DNA and thus maternally inherited. Additional disorders were subsequently associated with mutations within nucleus genes coding for proteins of mitochondrial location, and thus following Mendelian inheritance. Mitochondrial disorders are nowadays classified according to the genetic origins of the involved-mutations: (i) the first category, actually the first reported, concerns mt-DNA-encoded RNAs (reviewed in e.g. Schon et al., 2012Schon et al. 2012); (ii) the second category concerns mt-DNA-encoded proteins (e.g. Schon et al., 2012); (iii) the third category, the most diverse one, concerns nucleus-encoded proteins of mitochondrial location (e.g. Scheper et al., 2007b). Within the last group, mutations can directly contribute to a defect in OXPHOS by affecting proteins of the RC, or can interfere indirectly with OXPHOS by affecting proteins involved in the mt-DNA maintenance and/or translation. In this chapter we describe case studies belonging the first or third category. The herein reported mutations were found in patients with mitochondrial diseases of unknown cause and were investigated in collaboration with other scientists or medical doctors. We investigated three cases with different methods in order to offer new insights in the molecular mechanism underlying each specific case. We were not able to completely explain the molecular mechanism behind the pathological manifestation but our

experiments offer some pieces of the puzzle towards the complete understanding of those particular cases. Part of the obtained results contributed to two publications.

In the three following parts we will describe our:

In vitro analyses of the aminoacylation properties of a mutant tRNAAsp found in patients with sideroblastic anemia

In vivo analysis of the aminoacylation status of tRNAAsn found in patients harboring mutation in mt-AsnRS

In vitro and in vivo analyses of mutations found in mt-AspRS, their impact on solubility, mitochondrial import, dimerization (oligomery) and submitochondrial localization

1 *In vitro* analyses of the aminoacylation properties of a mutant $tRNA^{Asp}$ found in patients with sideroblastic anemia

The sideroblastic anemias are related to a group of disorders, in which the synthesis of heme is defective, leading to pathological accumulations of iron in the mitochondria of erythroid precursors. Patients with this disease suffer from a wide range of symptoms, including fatigue, dizziness, a rapid heartbeat, pale skin, and an enlarged liver and spleen (hepatosplenomegaly). In addition, patients are at risk of heart disease and liver damage (cirrhosis) due to an excess of iron in these organs (U.S. National Library of Medicine). The most common sideroblastic anemia is X-linked (XLSA), which is characterized by mutations in the erythroid-specific ALA synthase gene (ALAS2). This gene codes for a protein involved in the first and rate-limiting step of the heme synthesis, which occurs in the mitochondrial matrix. Over a decade ago it was hypothesized that primary mitochondrial defects may be involved in the pathogenesis of some forms of the sideroblastic anemia) as defects in mitochondrial translation have been associated with Pearson marrow-pancreas syndrome (large mtDNA rearrangment) (Rötig et al., 1990) or Myopathy, Lactic Acidosis, and Sideroblastic Anemia—MLASA Syndrome (mutations in mt-TyrRS) (Riley et al., 2010). This study is a part of a project aimed at analyzing the correlation between accumulations of mutated mt-tRNAs in patients with symptoms of a sideroblastic anemia (in collaboration with Mario Mörl, Leipzig, Germany).

Functional effects of the mutations on tRNA CCA-processing, tRNA aminoacylation and tRNA structure are analyzed. As we have expertise in the field of human mitochondrial aminoacylation system, we focused on aminoacylation of the tRNA and in particular of the tRNAAsp harboring the transition mutation A26G at a position known to be conserved as a purine (Helm et al., 2000). The aim of the investigation was it to decipher the kinetic parameters of aminoacylation of the mutant tRNA and compare them to those of the wild type tRNA. To do so, we cloned the genes corresponding to the wild type and mutant tRNA into a puc18 vector containing an upstream hammerhead ribozyme and a highly active T7 RNA Polymerase promoter

(**Figure 42**). A BstN1 sequence was included at the 3'-end of the sequence so that an accurate tRNAs CCA-end was generated after run-off transcription (crucial for recognition by the mt-aaRS). The tRNA was in vitro transcribed with a purified T7-Polymerase (gift from Joelle Rudinger, IBMC, Strasbourg).

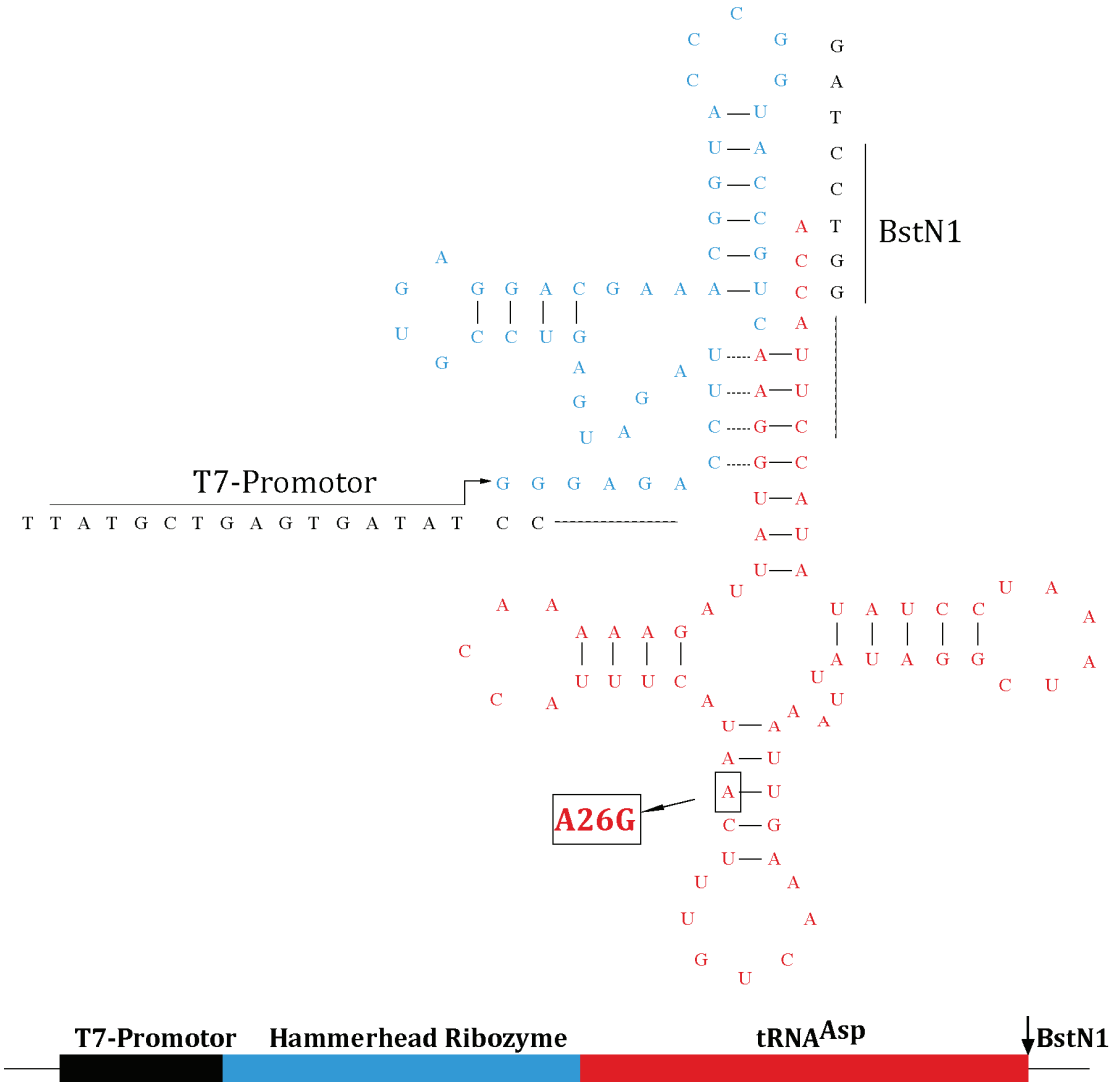


Figure 42: Construction of the tRNA^{Asp} transcript. On top: Schematized cloverleaf structure of the tRNA transcript (red) with the 5'-hammerhead ribozyme (blue). The T7-promotor is displayed as DNA sequence in black. The BstN1 digestion site is also displayed as DNA sequence in black. The position of the point mutation is indicated.

The generated "transzyme" (hammerhead ribozyme + tRNA transcript) was autocatalytically cleaved. Both the cleaved hammerhead ribozyme and the tRNA were separated on a preparative denaturating polyacrylamide gel and purified. The purified tRNA was used for aminoacylation assays. The kinetic parameters were determined

over three independent experiments. We showed that the $tRNA^{Asp}$(A26G) has a catalytic efficiency reduced by a factor of 2.2-fold compared to the wild-type tRNAAsp (**Table 8**). Of note, the affinity of the enzyme for its tRNA is not drastically impacted by the mutation (K_M values are similar) while the catalytic activity is reduced (k_{cat} values reduced 1.85 fold).

Construct	K_M (µM)	V_{Max} (nMs^{-1})	k_{cat} ($10^{-3}s^{-1}$)	Catalytic efficiency ($10^{-3}\mu Ms^{-1}$)	Loss
$tRNA^{Asp}$	1.53±0.72	0.46±0.5	9.07	**10.1**	-
$tRNA^{Asp}$(A29G)	1.64±1.01	0.24±0.13	4.88	**4.54**	2.2x
$tRNA^{Asp}$ *	0.72	-	27	37,5	-

Table 8: Kinetic parameters of wild type and mutant $tRNA^{Asp}$. Kinetic data were obtained using different amounts of tRNA and a constant concentration of recombinant mt-AspRS. Data are mean values of three independent assays. Line marked with asterisks * are extracted from (Scheper et al., 2007)

Of note, the catalytic efficiency of the wild-type $tRNA^{Asp}$, as determined in the present study, is ~3.5 fold lower than that determined by Scheper et al. (Scheper et al., 2007). However, the effect of the A26G mutation on aminoacylation kinetics are not significant compared to the drastic decrease (5000- to 7000-fold) observed in lysylation efficiencies of some $tRNA^{Lys}$ mutants found in patients with MERRF syndrome (Sissler et al., 2004). We thus conclude that a 2.2 fold loss in aspartylation efficiency is unlikely to be responsible for the phenotypic manifestation of the disease. Other step(s) of tRNA life cycle such as the 3'-end maturation, CCA-addition or the subsequent recognition of the aspartyl-$tRNA^{Asp}$ by Ef-Tu or the ribosome could be affected. Impacts of mutation on some of those steps (e.g. CCA-addition) are currently under investigation in the partner laboratory (Mario Mörl, Leipzig, Germany). A cumulative (mild) effect (s) of two or several of those steps cannot be excluded as well (reviewed in e.g. Florentz et al., 2003, Levinger et al., 2004). Further investigation should include the determination of the *in vivo* aminoacylation level of tRNA extracted from patient cells to decipher if cumulative defects lead to decrease *in vivo* aminoacylation level.

2 Analysis of mutations found in patients with a Leigh-like syndrome

In this section, we present our contribution to investigations concerning three patients, from apparently non-consanguineous parents, having a Leigh-like syndrome of unknown cause (in collaboration with Vincent Procaccio, Anger, France). Respiratory chain analysis of muscle homogenates revealed a functional defect of the NADH-cytochrome C reductase (Complex I→III) when normalized against citrate synthase activity. The analysis of protein content of the RC complexes by BN-PAGE and western blotting showed a reduced amount of complex I suggesting a defect in mitochondrial translation. Further analysis of cybrid cell lines, with an unrelated nuclear background and patient's mitochondria, showed no defect in mt-translation suggesting that the defect in mitochondrial translation observed in patients is of nuclear origin (personal communication, Vincent Procaccio). However, both the exome and mt-DNA from patients and relatives were sequenced and identified mutations are highlighted in the family pedigree in **Figure 43**.

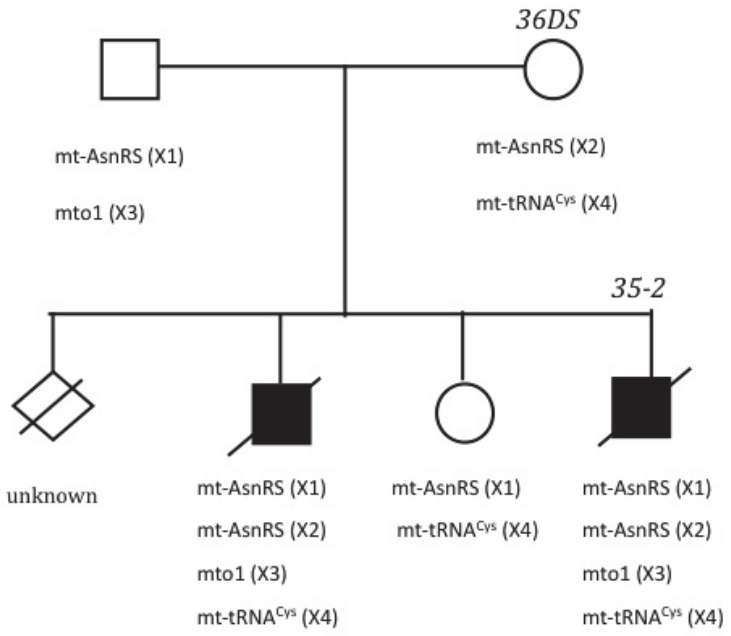


Figure 43: Family pedigree of the affected family. Mutations identified in different family members are highlighted. Four different mutations were identified in three different genes. Family members showing a

phenotype are indicated by filled squares. Deceased family members are indicated by a slash. Due to the confidentiality of these data, mutations are numbered as X1-X4.

The exome sequencing revealed that the two affected boys are carriers of the maternally inherited mt-AsnRS(X2) and paternally inherited mt-AsnRS(X1) mutations. In addition, they carry a mutation (X3) in the mto1 gene. Mt-DNA sequencing also revealed a maternally inherited heteroplasmic mutation in $tRNA^{Cys}$(X4). The unaffected girl is heterozygous carrier of the mt-AsnRS(X1) mutation and carry the heteroplasmic mutation $tRNA^{Cys}$ (personal communication, Vincent Procaccio).

On the basis of the above-mentioned sequence data, we tested the hypothesis that the translational defect could result from a reduced mt-AsnRS activity in patient cells. To test this, we analyzed the aminoacylation charging level of $tRNA^{Asn}$, extracted from fibroblasts cultivated from patient samples. We extracted total RNA under acidic conditions (to keep the ester bound between tRNA and amino acid) (Varshney et al., 1991) from cultured fibroblasts of one affected child (35-2), the mother (36DS), one unrelated affected patient (index case, 70JS) and a fibroblast cell line (as a control). RNAs were separated on a 40x30cm, 6,5% denaturating polyacrylamide gel in 0,1M Na-acetate (pH=5).

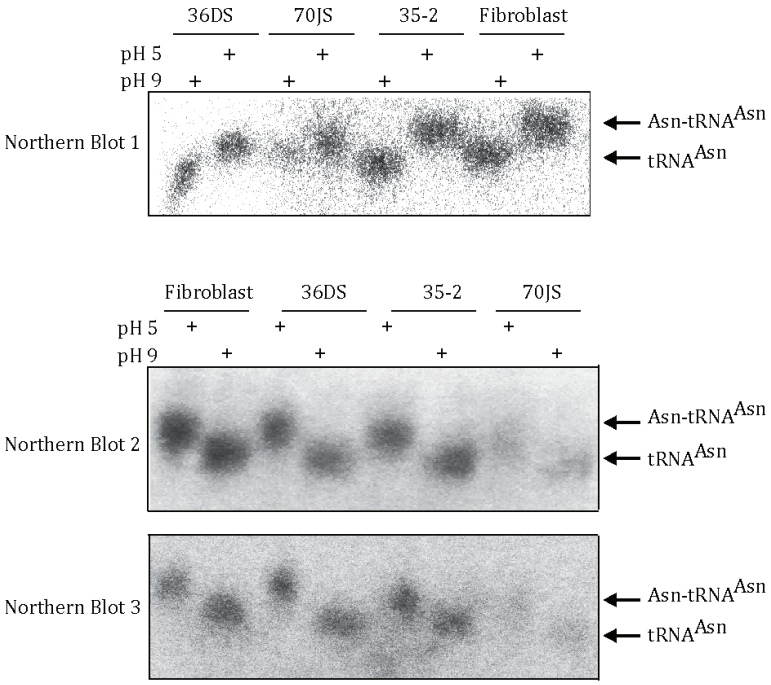


Figure 44: Analysis of asparaginyl-$tRNA^{Asn}$ by northern blot hybridization. Total RNA was extracted from different samples (see text). Equal amounts of RNA were analyzed. RNA were treated at

pH 9 to deacylate the tRNA or treated at pH 5 to keep the acylated tRNA. Three representative northern blots are displayed. Arrows indicate the acylated and deacylated forms of $tRNA^{Asn}$.

Northern hybridizations were performed with 5' radiolabeled oligonucleotide probes annealing either to tRNAAsn, $tRNA^{Lys}$ or to $tRNA^{Asp}$ (Enriquez and Attardi, 1996). As a first outcome (**Figure 44**), we observed that all detected tRNAAsn were aminoacylated and that none of them were non-aminoacylated under acidic conditions. This observation, confirmed by three independent preparations of tRNAs from patients cells, suggests that the efficiency of the aminoacylation reaction is not affected. The relative amount of Asn-$tRNA^{Asn}$ normalized against the control $tRNA^{Lys}$ or $tRNA^{Asp}$ were further quantified (**Figure 45**). Of note, while three independent northern blots were performed, the amount of tRNA could be quantified for only two of the three cases for sample 70JS.

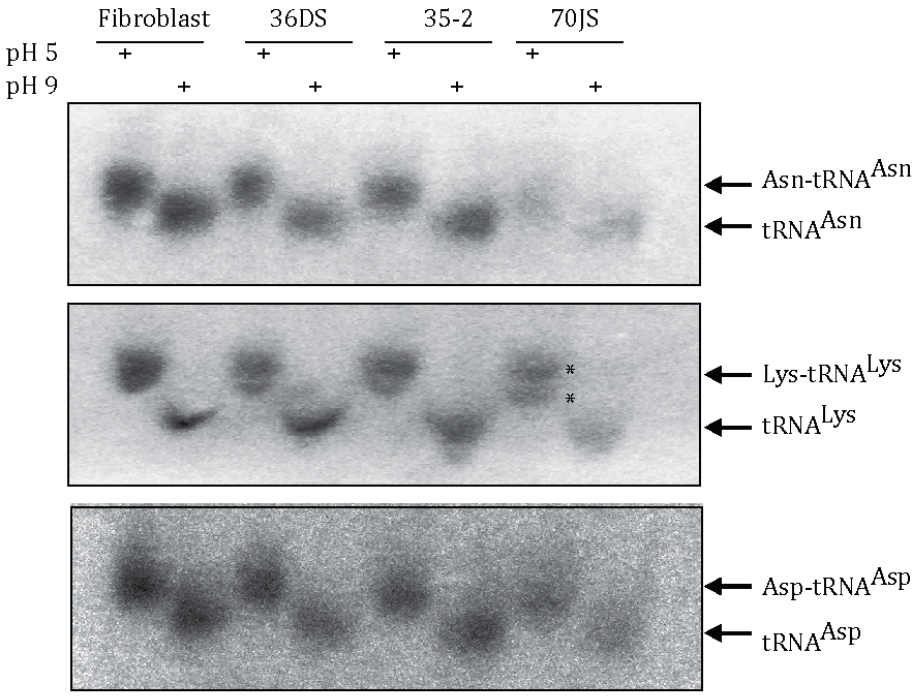


Figure 45: Analysis of asparaginyl-$tRNA^{Asn}$ by northern blot hybridization. Probes against $tRNA^{Lys}$ and $tRNA^{Asp}$ were used to further normalize the deposit of all samples (see **Figure 46).** RNAs were treated at pH 9 to deacylate the tRNA or treated at pH 5 to keep the tRNA acylated. Three representative northern blots are displayed. Arrows indicate the acylated and deacylated forms of tRNAs. Of note we observed two bands for the $tRNA^{Lys}$, which likely correspond to post-transcriptionally modified (lower band) and unmodified (upper band) tRNAs (Umeda et al., 2005).

The level of $tRNA^{Asn}$ is significantly ($p<0,05$) reduced in the cells from 36DS, 35-2 and 70JS as compared to control fibroblasts. The difference between the mother (36DS) and the child (35-2) is however not significant ($p>0,05$). Because of high experimental standard deviations observed with 70JS, no statistically relevant differences to 36DS or 35-2 could be stated. However, a tendency to a decreased amount of $tRNA^{Asn}$ in 70JS is observed.

As a first summary, no significant impact on the $tRNA^{Asn}$ aminoacylation levels could be demonstrated. In spite of this, a decrease in the total amount of $tRNA^{Asn}$ was observed in all three samples as compared to fibroblast control cells, with however no statistical relevance for the relative amounts within the three samples (because of high experimental standard deviations).

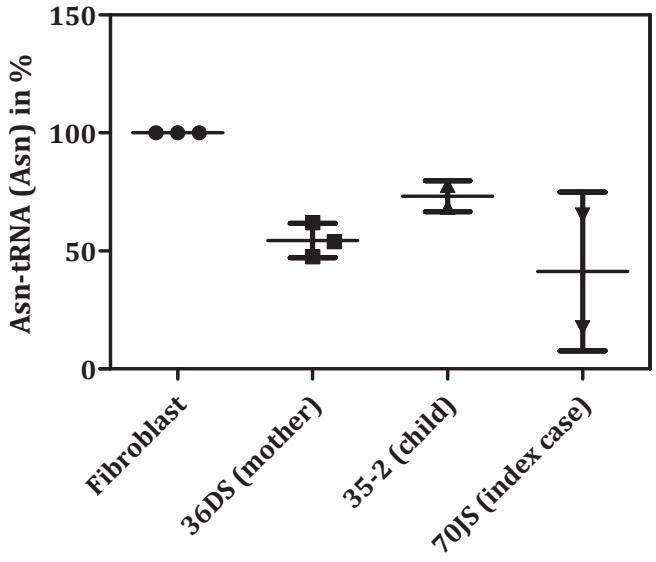


Figure 46: Quantification of asparaginyl-$tRNA^{Asn}$ northern blot hybridizations (from Figure 45). Amounts of aminoacylated $tRNA^{Asn}$ were quantified and normalized against the $tRNA^{Lys}$. Quantification is the result of at minimum of two independent northern blot experiments.

These reductions could reflect a retrograde regulation of the $tRNA^{Asn}$ level in response to the mt-AsnRS(X2) mutation. This mutation, present in all three samples including the 'healthy' mother, introduces a premature stop codon and thus leads to a likely unstable truncated product. Such a retrograde effect on tRNA amount has already been observed in cells from patients with defects in mt-ArgRS (Edvardson et al., 2007)

or from patients with defects in mt-SerRS (Belostotsky et al., 2011). These reductions have however to be considered with caution since they were only observed when compared to the fibroblast control cells of unrelated filiations.

The family pedigree revealed also the $tRNA^{Cys}$ (X4) mutation within the mt-DNA. This mutation is present in the two affected boys, but also in the 'healthy' mother and 'healthy' girl. Preliminary northern blot experiments under basic and acidic conditions did not reveal significant differences in the two samples (36DS and 35-2) when compared to the fibroblast control cells. It however revealed a significant reduction of the $tRNA^{Cys}$ level in samples obtained from the index case (70JS) (with no visible reduction of $tRNA^{Lys}$ and $tRNA^{Asp}$ amounts, as controls). In the absence of information about the family history of this patient further discussion remains speculative.

Another mutation was identified in the coding sequence of mto1 (**Figure 43**). The corresponding protein is responsible for the 5-carboxymethylaminomethyl modification (of the mnm5s2U34 hyper-modification) of uridine at the wobble position in mitochondrial tRNAs (Li et al., 2002). This modification is found in e.g. bacterial tRNAs specific for glutamate, lysine, and glutamine and plays import roles in tRNA structure stabilization, as an identity element in aminoacylation or mRNA codon recognition (Yarian et al., 2002). Mutations in mto1 were already reported to cause hypertrophic cardiomyopathy and lactic acidosis (Ghezzi et al., 2012). Our aim was to test the impact of this mutation on the post-transcriptional modification profile of the $tRNA^{Lys}$ and $tRNA^{Glu}$. Indeed, additional bands (asterisks) were observed in **Figure 45** (migration under acidic conditions), and were assumed to represent modified and unmodified forms of the tRNA (similar separation of modified and unmodified tRNA were also observed in chapter II, migration under non-acidic conditions). Unfortunately, as shown in **Figure 47**, the separation under non-acidic conditions could not resolve modified tRNAs from unmodified tRNAs, neither in the control fibroblast nor in samples obtained from patients. Thus a comparative analysis was not possible.

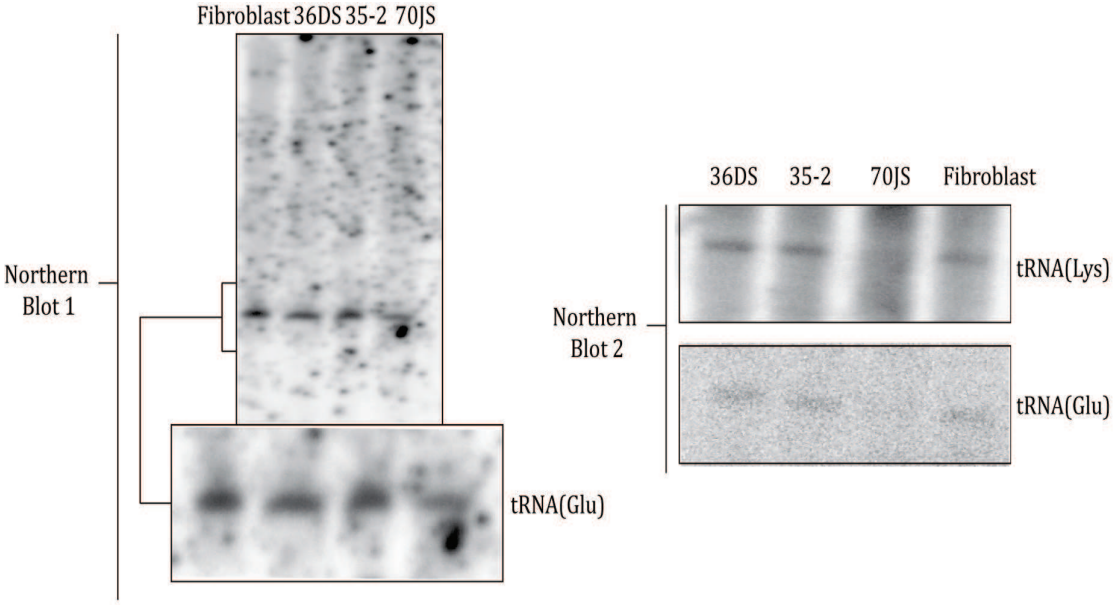


Figure 47: Analysis of modified tRNAs by northern blot hybridization. Total RNA was extracted from different samples (see text). Equal amounts of RNA were analyzed. Two northern blots are displayed. In northern blot 1 a picture of the whole membrane is shown and enlarged picture of the band corresponding to tRNAs. In northern blot 2 only the enlarged pictures are shown.

To overcome this problem, it is reported in the literature that the degree of thiolated tRNA can be assessed by retardation of electrophoresis mobility in polyacrylamid gels containing APM (N-acryloylamino)phenylmercuric chloride) (Igloi, 1988). This experiment has not been performed so far, as we were unable to obtain APM yet (not commercially available). However, the effect of the mutation in mto1 remains open since the unaffected father and affected child are both carriers in a heterozygous compound state. In addition, it was shown for yeast mto1 knock out strains, that the *mto1* mutation alone is insufficient to produce a respiratory deficient phenotype (Umeda et al., 2005). But a large proportion of incompletely modified mt-tRNAs were observed and associated with increase degradation rate of the tRNAs (Umeda et al., 2005; Wang et al., 2010).

In summary, we investigated several hypotheses to try to explain the pathological phenotype of a distinct set of patients with Leigh-like syndrome harboring mutations in mt-AsnRS, $tRNA^{Cys}$, and mto1. We observed no significant impact on the $tRNA^{Asn}$ or $tRNA^{Cys}$ aminoacylation levels caused by mutation in mt-AsnRS and $tRNA^{Cys}$, respectively. In addition, we could not so far observe the impact of a mutation in mto1-gene on the modification of $tRNA^{Glu}$ and $tRNA^{Lys}$. Nevertheless, we observed a

decrease level of the quantified tRNAs in the unaffected mother and affected child compared to fibroblast obtained from unrelated control. It is speculated that unmodified or uncharged tRNAs undergo a degradation. Indeed, several studies showed that mutations in mto1 or mt-aaRSs can lead to partially modified tRNAs, which are associated with an increased degradation rate. It thus can be speculated/proposed that cumulative (mild) effects may play a role in the phenotypic manifestation of this disease.

Of note, the phenotypic manifestations were observed in muscle cells. For sake of feasibility, patient samples are derived from fibroblasts. Thus, it can not be excluded that the mutation in genes coding for mto 1, mt-AsnRS or $tRNA^{Cys}$ have not the same impact in fibroblasts than in muscle or neuronal cells.

3 Pathology-related mutations in mt-AspRS: biophysical properties and in vivo characterization

3.1 Biophysical characterization of mutations in mt-AspRS

The first identification of mutations within a nucleus-encoded mt-aaRSs as cause of a mitochondrial disease dates to 2007. It was noticed that mutations in the DARS2 gene, coding for the mt-AspRS, correlated with cerebral white matter abnormalities of unknown origin (Scheper et al., 2007). These abnormalities were part of a childhood-onset disorder called Leukoencephalopathy with Brain stem and Spinal cord involvement and Lactate elevation (LBSL) (van der Knaap et al., 2003). Since this first discovery, an additional 9 mt-aaRSs have been correlated with mitochondrial diseases (reviewed in bookchapter #2). Over 65 mutations are reported so far (April 2013). Among these, 28 were found in the mt-AspRS. This aaRS is today the most intensively explored one in connection with a disease. The initial set of mutations was analyzed in regard to their aminoacylation activity (Scheper et al., 2007). In vitro analysis showed that some mutations severely impact the aminoacylation and some others do not. Later investigations revealed impacts on pre-mRNA splicing, protein expression, or protein

dimerization (van Berge et al., 2012a; van Berge et al., 2012b). The data obtained for six of the mutants are summarized in **Table 9**.

Mutation		Protein	Expression	Activity	Localization	Dimerization Homo	Dimerization Hetero
R58G	missense mutation		OK	OK	OK	↓	↓
T136S	missense mutation		OK	OK	OK	↓	↓
R76SfsX5	splicing defect	truncated protein	↓	OK	OK	OK	↓↓
Q184K	missense mutation		↓	OK	OK	↓	↓↓
R76SfsX5	splicing defect	truncated protein	↓	OK	OK	OK	↓
R263Q	missense mutation		OK	↓↓	OK	↓	↓
L613F	missense mutation		OK	OK	OK	OK	OK
L626Q	missense mutation		OK	↓↓	OK	↓	OK

Table 9: Summary of previously reported effects of missense mutations in mt-AspRS (addapted from Van Berge et al., 2012a). The heterozygous compound status is illustrated by pairwise display of mutants; Arrows indicate a reduction compared to wild-type. A double arrow indicates a strong reduction. Red cross indicate point mutation.

It is worth recalling that those mutations are found in a compound heterozygous state and that patients harbored the following combination of mutations: R58G with T136S; L613F with L626Q; Q184K with the exon3 splicing defect (R76SfsX5); and R263Q with the same exon3 splicing defect. For the two last cases, a mutation in a polypyrimidine tract at the 3'-end of intron 2 is found in one allele. This mutation affects the correct splicing of the third exon, which leads to a frameshift and a premature stop. This mutation is 'leaky', meaning it leads to a decreased, but not zero, expression of full-length mt-AspRS (Scheper et al., 2007). A consequence of heterozygosity is the production of two distinct polypeptide chains that can randomly associate to build the dimer, which theoretically leads to an equal proportion between the two possible homodimers (each harboring the same mutation) and the heterodimer (where each constitutive momomer is harboring a different mutation) (reviewed in bookchapter #2).

The mutations listed in **Table 9** were further analyzed in our group. The goal of this study was to shed light on the impact of the mutations on physical-chemical properties of the protein using complementary approaches including Dynamic (DLS)

and Static (SLS) Light Scattering, Differential Scanning Fluorimetry (DFS), Synchrotron radiation Circular Dichroïsm (SRCD) and Small-Angle X-ray Scattering (SAXS). I was in charge of a further investigation of the oligomeric state and solubility of the mutants using biochemical and cellular biology methods. The major outcome of this study is presented in article #3 (under preparation). The conclusion of this study was that the mutation that introduces the replacement of a Gln into a Lys at position 184 leads to the most drastic structural perturbation and very likely to aggregation of the protein both *in vitro* and *in vivo*. The manuscript being under preparation, the current draft is proposed below, and the experiments directly under my responsibility are further discussed here.

Some of the selected mutations are close to the dimerization interface of the mt-AspRS (**Figure 48**). It has indeed been shown by 'dimerization assays' (van Berge et al., 2012a) that oligomery of the enzyme is affected mostly by mutations R58G, Q184K and R263Q (summarized in **Table 9**). The R58Q showed a reduced dimerization of the homo- (with itself) and heteromer (with T136S). The mutants Q184K and R263Q showed a reduced dimerization as a homomer or heteromer (with wild type protein) (van Berge et al. 2012a). In these experiments, the authors transfected HEK cells with Flag- and Gluc-tagged mt-AspRS constructions. After co-immuno precipitation using an anti-flag antibody, the luminescence of the co-immuno precipitated mt-AspRS Gluc fusion proteins was measured. While somehow informative, this experiment does not take into account the solubility of the expressed mutants and the possible influence of the solubility on the dimerization. Of note, an inspection of the experimental conditions of this report suggests that putative aggregates were removed together with cell debris and nuclei after cell lysis by a centrifugation step.

We thus performed an experiment that combined both parameters: solubility and dimerization. To do so, we used the engineered expression system using MVA-infected hamster cell-lines (Jester et al., 2011; article #2).

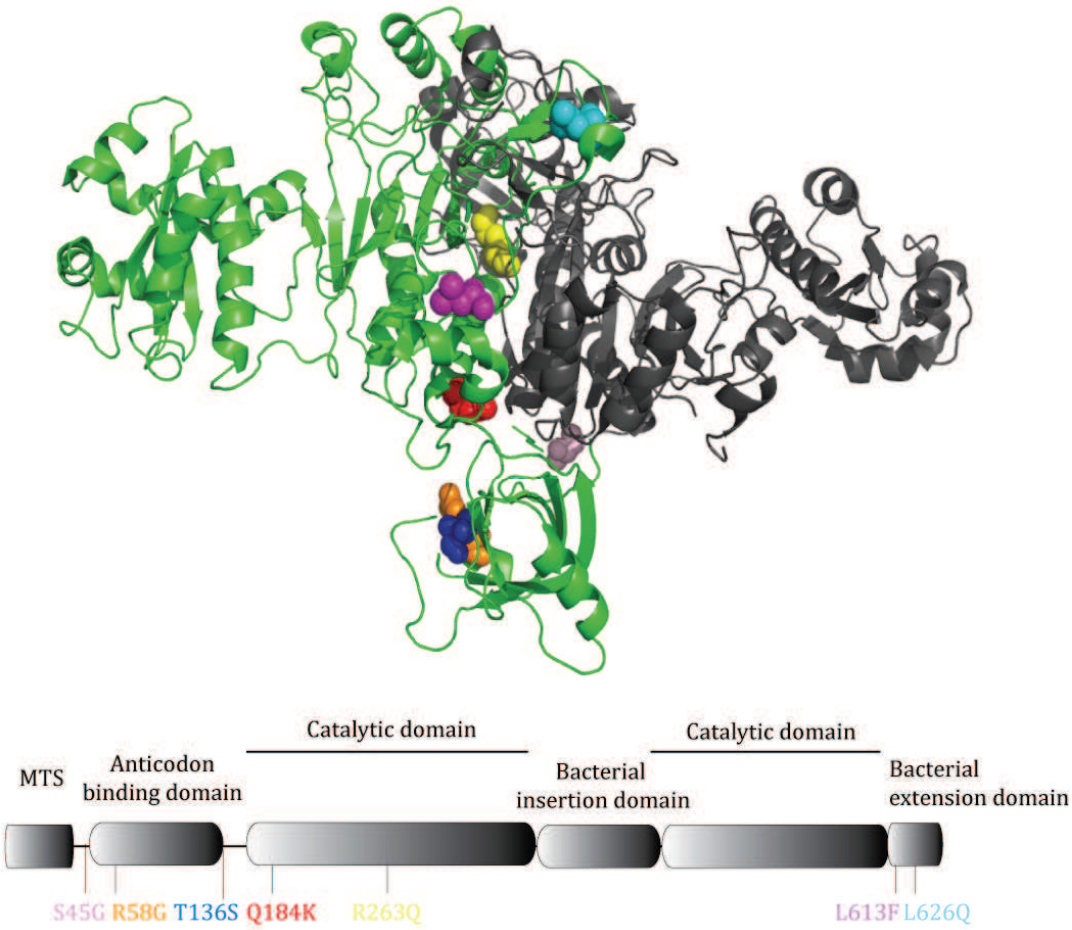


Figure 48: Localization of mutations within the crystallographic structure of human mt-AspRS (Structure from Neuenfeldt et al., 2012; PDB 4AH6). The mt-AspRS is a dimer. Monomers are distinguished by the green and grey colors. The mutations S45G (light pink), R58G (orange), T136S (blue), Q184K (red), R263Q (yellow), L613F (magenta) and L626Q (cyan) are displayed. On bottom: schematic illustration of the 2-D organization of the human mt-AspRS. Locations of mutations are indicated.

Point mutations were introduced inside the coding sequence of mt-AspRS by mutagenesis PCR. The different constructions (wt and mutants) were expressed in BHK21 cells. After disruption of the cells, crude lysates were further separated into "soluble" and "insoluble" protein fractions by centrifugation (**Figure 49**). The soluble protein fraction was loaded and separated on a BN-PAGE. Separated proteins were transferred to PVDF membrane and the recombinant expressed mutants were detected by immuno detection with an anti-flag antibody (**Figure 49**A). The remaining insoluble fraction was further first treated with a mild nonionic detergent (2% Trition) and subsequently treated with a protein denaturant (urea 7M). The two resolubilized fraction

were resolved on a SDS-PAGE and further transferred to PVDF membrane. As previously, mutants mt-AspRS were detected by immuno detection with an anti-flag antibody (**Figure 49**B). The combination of all allows detection of all expressed proteins, either in a soluble state, or in an insoluble state (membrane-anchored or aggregated).

We observed that:

1. all mutants were expressed (**Figure 49**A) as soluble or insoluble protein
2. all soluble expressed mutants form dimers
3. wild-type protein, R58G, L613F and L626Q mutants are soluble in triton and Q184K and R263Q not.
4. mutant Q184K has a reduced expression and is only soluble in urea suggesting aggregates.
5. also the R263Q and the L626Q partly formed aggregates

It appeared that Q184 has a reduced expression and is insoluble. This result is in line with the strong tendency of the mutant Q184K to aggregate, as demonstrated by the different biophysical approaches (See article #3). Concerning the mutation R263Q, it remains unclear whether the mutations have a direct impact on the folding of the monomer or on the dimerization of the two folded monomers. For L626Q, the increased expression level may fortify the aggregation of a part of the protein as artifact of overexpression. Of note, experiments performed on purified and fractionated mitochondria indicate that all mutants are at least targeted to mitochondria, since they co-purified with the organelle (see chapter V). Of note, double bands are observed in our western blots for the different mt-AspRS constructs. They may either correspond to artifacts of the SDS-gel electrophoresis or be of biological relevance. Indeed, similar pattern of double bands were already shown for mt-AspRS in totally independent experiments (*in vitro* import assay, (Messmer et al., 2011). Those two bands were designed as maturated products, cleaved upon internalization into the mitochondria. Their migration profiles were distinguished (on SDS-PAGE) from the one of precursor mt-AspRS. At this time, the appearance of the two bands was suggested to be the

signature of the processing of mt-proteins localized within the inner membrane (Nett at al., 1996). Nevertheless, no data are presently favoring either of these hypotheses.

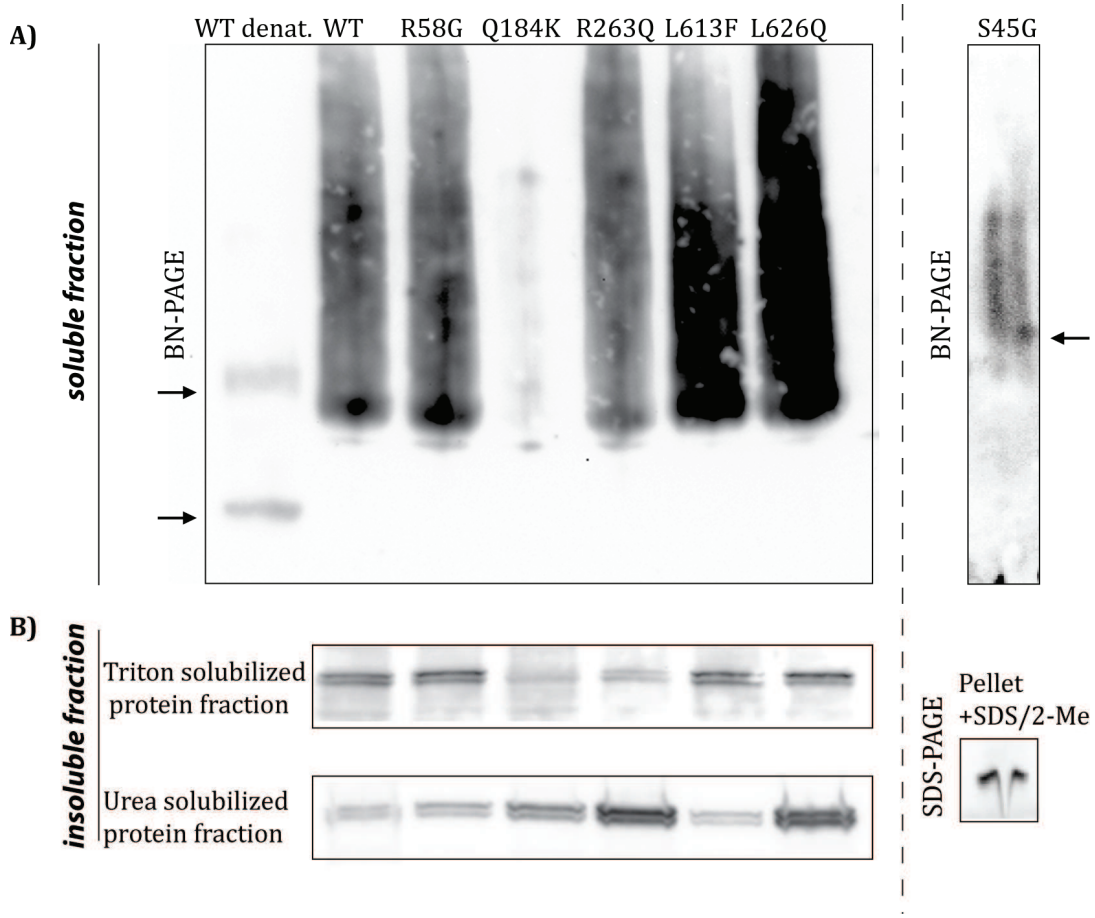


Figure 49: Solubility test of recombinant expressed mutants. A) Soluble cell lysate fraction was loaded on BN-PAGE. Western blot detections were performed using anti-Flag antibody. Monomeric and dimeric forms are indicated with arrows. B) Insoluble cell lysate fraction (pellet) was sequentially treated with 2% Triton and Urea/2-Mercaptoethanol (2-Me). Solubilized proteins were separated on SDS-PAGE and detected by western blot analysis. Of note, mutant T136S was not investigated at this point. Picture on the right side of the dashed line referred to mutant S45G discussed in a later paragraph. WT stands for wild type protein. Denat. stands for a denaturating treatment with SDS, 2-Me and a 8 min cooking step

In summary, we provide additional value to the previously obtained data. In addition to this report, we showed that all soluble expressed proteins form dimers while the Q184K mutant has a reduced expression and mainly forms aggregates. We also presented that R263Q forms aggregates, which was previously revealed to have reduced dimerization behavior.

3.2 Article # 3

Comparative biophysical investigations of pathogenic mutants of human mitochondrial aspartyl-tRNA synthetase

Claude SAUTER[1†]*, Bernard LORBER[1†], Agnès GAUDRY[1], Hagen SCHWENZER[1], Frank WIEN[2], Pierre ROBLIN[2,3], Catherine FLORENTZ[1] and Marie SISSLER[1]*

2013

(in preparation).

Abstract:

Since the discovery of pathology-related mutants of human mitochondrial aspartyl-tRNA synthetase (mt-AspRS), the list of single point mutations affecting nuclear genes of aminoacyl-tRNA synthetases (mt-aaRSs) has been growing rapidly. These mutations are associated with a variety of neurodegenerative disorders but the way they affect mt-aaRSs in their structure and/or function remains to be established. To gain an insight into their link with mitochondrial disorders we focused on human wild type mt-AspRS and a selection of six mutants spread all over the enzyme sequence. The structural characterization of the native mt-AspRS and its comparison with the *E. coli* enzyme highlighted a common 3D architecture. Despite this structural similarity, functional differences in stability, substrate selectivity and processing were detected, indicating a higher plasticity of the mitochondrial enzyme. To detect the effects of pathology-related mutations on mt-AspRS properties in solution we used an integrated structural strategy based on biochemical and biophysical approaches including enzymology, dynamic light scattering (DLS), size exclusion chromatography (SEC), differential temperature scanning (DTS), circular dichroism (SRCD), small angle X-ray scattering (SAXS), 3D modeling, and *in vivo* methods. While the mt-AspRS fold and activity do not seem to be dramatically altered, these mutants show pleiotropic effects on the enzyme solubility, thermal stability, or a strong tendency to aggregate. The biological relevance of these observations will be discussed in relation with mitochondrial disorders.

3.3 In vivo characterization of mutants of mt-AspRS

3.3.1 Pathology related mutations disturb the mitochondrial import and sub mitochondrial localization

A significant part of my PhD-thesis was to establish the sub-organization of the aaRSs within mitochondria. The aim was to contribute to a broader knowledge about the mitochondrial aaRSs and to "integrate" them into cellular and functional networks. This increased knowledge should also serve to better understand the impact of some pathology-related mutations. The mutations R58G, Q184K, R263Q, L613F and L626Q were intensely investigated in regard to their classical properties, e.g. aminoacylation but also in regard to their expression, solubility, dimerization and biophysical properties (van Berge et al., 2012a and article#3). In addition, the translocation of the mutants to mitochondria has been investigated using flag-tagged proteins transfected in HEK293T cells and immuno cytochemical detection methods. It was shown that all mutants have normal mitochondrial localization, indicating a proper targeting to mitochondria (van Berge et al., 2012a). This method is widely used to analyze the mitochondrial localization of proteins but its possible outcomes remain restrained. It is for instance not possible to distinguish proteins localized at the periphery of the mitochondria or that have been blocked from translocation, from proteins that have been internalized inside mitochondria. There exist two methods to overcome this problem and they both require a mitochondria purification step. The first method uses purified mitochondria for *in vitro* import assays (for the investigation of the S45G mutant of mt-AspRS, article #4), and the second method uses purified mitochondria for subsequent fractionation (for the investigation of mutants R58G, Q184K, R262Q, L613F and L626Q of mt-AspRS, see below).

3.3.2 A human pathology-related mutation S45G prevents import of the mt-AspRS into mitochondria

The mutation S45G was found in some patients with LBSL (Scheper et al., 2007). It is located inside the theoretically predicted MTS of mt-AspRS (Bonnefond et al., 2005) at close vicinity of the experimentally determined cleavage site (L33, article #2). To recall, the MTS is necessary for the import of proteins into mitochondria. In the present study, we have investigated the effect of the S45G mutation on targeting, binding, translocation and processing steps of mt-AspRS. Confocal microscopy imaging revealed that mutation S45G affects neither the targeting nor the binding of the protein to the mitochondria. However, by combining *in vitro* import and processing assays, we found the mutation impaired the translocation step. We showed that this pathogenic mutation strongly affects mt-AspRS translocation through the mitochondrial membranes. This is the first time that an import defect of a mitochondrial translation machinery factor is associated with a severe human brain disorder. These results are summarized in article #4.

In addition, we expressed the S45G mutant in BHK21 cells, and observed a abberant running pattern on BN-PAGE (**Figure 49**). The soluble protein showed an size of higher molecular weight compared to the wild type mt-AspRS dimer. We assume that the size difference represents the unprocessed forms of the cytosol or the unfolded proteins having a different migration behavior. This mutant was not further detected inside mitochondria after fractionation and separation of the matrix, peripheral and integral fractions (**Figure 50**). This confirmed the impact of the mutation on the translocation steps and further validates this method (as suggested in article #2) to investigate pathology-related mutants.

The S45G has not previously been shown to affect mitochondrial activity. Of note, this mutation was found in compound-heterozygous state together with R76SfsX5 mutation. This mutation is situated in intron 2 and affects the correct splicing of exon 3, which causes a frameshift and generates a premature stop codon. The leaky nature of the splicing defect leads to the expression of a small amount of wild type mt-AspRS being apparently sufficient to support mitochondrial translation in the tested cells but not sufficient in neuronal cells. This hypothesis remains to be tested. In addition, the

discrepancy between the molecular impact of the mutation and phenotypic expression could also indicate that, a yet unknown, non-translational function of the mt-AspRS is affected.

3.3.3 Article #4

A human pathology-related mutation prevents import of an aminoacyl-tRNA synthetase into mitochondria

Marie MESSMER, Catherine FLORENTZ, Hagen SCHWENZER, Gert C. SCHEPER, Marjo S. VAN DER KNAAP, Laurence MARECHAL-DROUARD and Marie SISSLER

Biochem. J. (2011) ***433****, 441–446 (Printed in Great Britain)*

BJ www.biochemj.org

This is a data-enriched, interactive PDF that provides the gateway to a world of information when opened in Utopia Documents
Download FREE software now

BJ Disease

Biochem. J. (2011) **433**, 441–446 (Printed in Great Britain) doi:10.1042/BJ20101902

ACCELERATED PUBLICATION

A human pathology-related mutation prevents import of an aminoacyl-tRNA synthetase into mitochondria

Marie MESSMER*, Catherine FLORENTZ*, Hagen SCHWENZER*, Gert C. SCHEPER†, Marjo S. VAN DER KNAAP†, Laurence MARÉCHAL-DROUARD‡[1] and Marie SISSLER*[1]

*Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 rue René Descartes, F-67084 Strasbourg, France, †Department of Pediatrics and Child Neurology, VU University Medical Center, 1081 HV Amsterdam, The Netherlands, and ‡Institut de Biologie Moléculaire des Plantes, UPR 2357, CNRS, conventionné avec l'Université de Strasbourg, 12 rue du Général Zimmer, F-67084 Strasbourg, France

Mutations in the nuclear gene coding for the mitochondrial aspartyl-tRNA synthetase, a key enzyme for mitochondrial translation, are correlated with leukoencephalopathy. A Ser^{45} to Gly^{45} mutation is located in the predicted targeting signal of the protein. We demonstrate in the present study, by *in vivo* and *in vitro* approaches, that this pathology-related mutation impairs the import process across mitochondrial membranes.

Key words: aminoacyl-tRNA synthetase, organelle, pathology-related mutation, protein import, translation machinery, translocation.

INTRODUCTION

The mitochondrion has many fundamental functions, e.g. in metabolic pathways, redox processes, energy production or apoptosis. The link between mitochondrial energetic dysfunction and cancer, aging phenomena and a broad range of metabolic and degenerative diseases is becoming more and more recognized [1]. Mitochondrial disorders can be caused by mutations in mtDNA (mitochondrial DNA) genes encoding either a core protein of an oxidative phosphorylation complex, or rRNAs and tRNAs required for mitochondrial translation. Additionally, mutations in nuclear genes encoding mitochondrial proteins are increasingly found to be associated with mitochondrial disorders, with the hypothesis that they would even be more common than mtDNA mutations [2].

An emerging field of neuromuscular and neurodegenerative disorders is linked to mutations in nuclear-encoded factors of the mitochondrial translation machinery [3] and especially mt-aaRS (mitochondrial aminoacyl-tRNA synthetases). LBSL (leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation) is a monogenic disease associated with a large variety of mutations affecting the human nuclear gene *DARS2*, encoding mt-AspRS (mitochondrial aspartyl-tRNA synthetase) [4], and novel mutations in this gene have been reported [5,6]. Deleterious mutations in *RARS2* [mt-ArgRS (mitochondrial arginyl-tRNA synthetase) gene] are associated with pontocerebellar hypoplasia [7] and mutations in *YARS2* [mt-TyrRS (mitochondrial tyrosyl-tRNA synthetase) gene] are associated with myopathy, lactic acidosis and sideroblastic anaemia [8]. Missense mutations in *GARS* [encoding both cytosolic and mitochondrial GlyRS (glycyl-tRNA synthetase)] cause a variant of Charcot–Marie–Tooth disease [9,10].

Although mutations in *RARS2* and *YARS2* alter aminoacylation properties of mt-ArgRS and mt-TyrRS, consequences of the mutations affecting mt-AspRS are intriguingly different. Indeed, *in vitro* activities of the subset of mutants tested so far only revealed a limited loss of their aminoacylation properties. Additionally, when respiratory chain complex activities were examined in fibroblast and lymphoblasts, no apparent dysfunctions were observed. Tissue-specific parameters and/or consequences on alternative functions of the aaRS have to be considered [4,5].

In the present paper, we are interested in the molecular consequences of the S45G change in mt-AspRS that was identified in LBSL patients. Ser^{45} resides within the predicted MTS (mitochondrial-targeting sequence) of mt-AspRS [11] (Figure 1). Most nuclear-encoded mitochondrial proteins possess a unique MTS, necessary for their presence in mitochondria. Mitochondrial proteins synthesized in the cytosol are first targeted to the surface of mitochondria via their interaction with chaperones. The MTSs are then specifically recognized by mitochondrial outer membrane receptors before guiding the precursor proteins through translocase complexes of the outer and inner membrane. Finally, MTSs are identified and cut-off by mitochondrial-processing peptidases, releasing mature proteins (reviewed in [12–16]). In the present study, we have investigated the consequences of the S45G mutation on targeting, binding, translocation and processing steps of mt-AspRS. We show that this pathogenic mutation strongly affects mt-AspRS translocation through the mitochondrial membranes. This is the first time that an import defect of a mitochondrial translation machinery factor is associated with a severe human brain disorder.

EXPERIMENTAL

Cloning

Constructs used in the confocal microscopy experiments were cloned into pcDNA3.1/CT-GFP-TOPO® vector (Invitrogen) following the manufacturer's protocol. PCR fragments corresponding to MTS–GFP (green fluorescent protein),

Abbreviations used: (mt-) aaRS, (mitochondrial) aminoacyl-tRNA synthetase; (mt-) ArgRS, (mitochondrial) arginyl-tRNA synthetase; (mt-) AspRS, (mitochondrial) aspartyl-tRNA synthetase; GFP, green fluorescent protein; (mt-) GluRS, (mitochondrial) glutamyl-tRNA synthetase; (mt-) GlyRS, (mitochondrial) glycyl-tRNA synthetase; HEK, human embryonic kidney; LBSL, leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation; mtDNA, mitochondrial DNA; MTS, mitochondrial-targeting sequence; mt-TyrRS, mitochondrial tyrosyl-tRNA synthetase.

[1] Correspondence may be addressed to either of these authors (email laurence.drouard@ibmp-cnrs.unistra.fr or m.sissler@ibmc-cnrs.unistra.fr).

© The Authors Journal compilation © 2011 Biochemical Society

Biochemical Journal

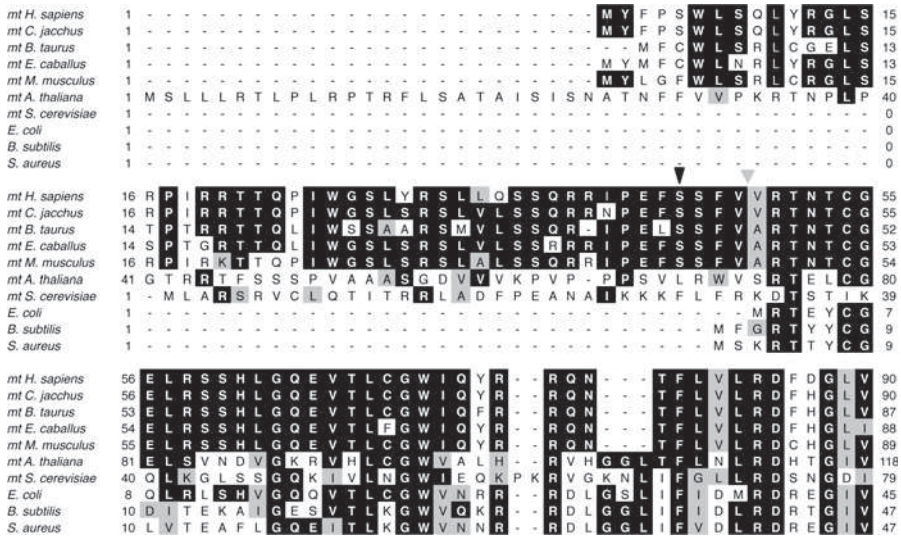


Figure 1 Multiple sequence alignment of the N-terminal region of human mt-AspRS with a selection of known AspRSs

NCBI accession numbers of sequences are mt *H. sapiens* (mitochondrial *Homo sapiens*), NP_060592.2; mt *C. jacchus* (mitochondrial *Callithrix jacchus*), XP_002760400; mt *B. taurus* (mitochondrial *Bos taurus*), NP_001095692; mt *E. caballus* (mitochondrial *Equus caballus*), XP_001493136; mt *M. musculus* (mitochondrial *Mus musculus*), NP_766232; mt *A. thaliana* (mitochondrial *Arabidopsis thaliana*), NP_195102; mt *S. cerevisiae* (mitochondrial *Saccharomyces cerevisiae*), NP_015221; *E. coli* (*Escherichia coli*), NP_416380; *B. subtilis* (*Bacillus subtilis*), NP_390633; and *S. aureus* (*Staphylococcus aureus*), NP_372154. Black and grey boxes highlight identical and similar residues respectively. The predicted cleavage position of the mt-AspRS MTS is indicated by a grey arrowhead. The serine residue mutated into glycine at position 45 in the human mt-AspRS is indicated by a black arrowhead.

MTS-mt-AspRS−GFP and MTS(S45G)-mt-AspRS−GFP were generated using the pEGFP-N1-DARS2 and pEGFP-N1-(S45G)-DARS2 plasmids. These plasmids contain the full-length mt-ApsRS-coding sequence (either wild-type or containing the mutation corresponding to S45G) cloned into pEGFP-N1 (Clontech). The mt-AspRS amplicon was amplified from pQE70-mt-AspRS [11]. Constructs used in the *in vitro* import and processing assays, PCR-generated truncated mt-AspRS versions (from amino acids 1 to 313, with or without the mutation), were cloned into pCR®2.1 vector (Invitrogen) downstream of a T7 promoter. A Kozak consensus sequence (5′-GCCATG-3′) was introduced during the PCR for the translation of the corresponding proteins. The plasmid expressing the first 81 amino acids of *Arabidopsis thaliana* mt-GluRS (mitochondrial glutamyl-tRNA synthetase) upstream of the GFP sequence has been published previously [17].

Cell culture and visualization of the mt-AspRS−GFP fusion protein variants

HEK (human embryonic kidney)-293T cells were maintained in DMEM (Dulbecco's modified Eagle's medium, Invitrogen) supplemented with 10 % FBS (fetal bovine serum), 100 units/ml penicillin and 100 μg/ml streptomycin at 37 °C, in a 10 % CO_2 humidified incubator. Cells were seeded on to coverslips at 3×10^5 cells per 35 mm dish. At semi-confluence, cells were transiently transfected by the calcium phosphate method with 500 ng of plasmids. After 36 h, MitoTracker Red CM-H_2XRos (Invitrogen) was added to a final concentration of 100 nM for 30 min. Images were obtained using confocal microscopy as described previously [17].

Human mitochondria purification

Mitochondria were purified from HEK-293T cell lysates (Waring blender) by differential centrifugation at 1500 ***g*** followed by a final step at 20500 ***g*** for 25 min in 5 mM Hepes/KOH (pH 7.5), 210 mM mannitol, 70 mM sucrose, 2 mM EDTA, 0.5 % BSA and 2 mM 2-mercaptoethanol. The pellet of mitochondria was washed twice either in the import buffer [10 mM Hepes/KOH (pH 7.5), 25 mM sucrose, 75 mM sorbitol, 100 mM KCl, 10 mM KH_2PO_4/K_2HPO_4, 0.05 mM EDTA and 5 mM $MgCl_2$] or in the import buffer without EDTA for the processing assay, and then kept on ice. The concentration of mitochondrial proteins was measured using the Bradford assay.

In vitro assays

In vitro synthesis of proteins in the presence of [^{35}S]methionine (1000 Ci/mmol, Amersham) was carried out with the TNT® T7 Coupled Reticulocyte Lysate System (Promega) following the manufacturer's protocol. Import assays were performed as described previously [17]. The incubation time (30 min) is sufficient to cover the full process of import and maturation [18]. *In vitro* processing assays were performed as described previously [19]. Briefly, sonicated HEK-293T mitochondria (3 rounds of sonication for 10 s, medium speed, bioruptor, Diagenode) were diluted 2-fold in a processing buffer containing 40 mM Hepes (pH 7.3), 2 mM $MnCl_2$, 0.2 % Tween, 2 mM DTT (dithiothreitol) and 1×EDTA-free protease inhibitor cocktail. Then 5 μl of ^{35}S-labelled fusion proteins were added to 30 μl of mitochondrial suspension. After a 3 h incubation at 30 °C, samples were analysed by SDS/PAGE and autoradiography.

RESULTS AND DISCUSSION

Mutated mt-AspRS co-localizes with mitochondria

To determine the possible consequences of the mutation on any of the mitochondrial protein import steps, targeting to mitochondria was first explored. HEK-293T cells, transfected with GFP-fusion protein constructs, were visualized by confocal microscopy. Although mt-AspRS lacking its theoretical MTS

© The Authors Journal compilation © 2011 Biochemical Society

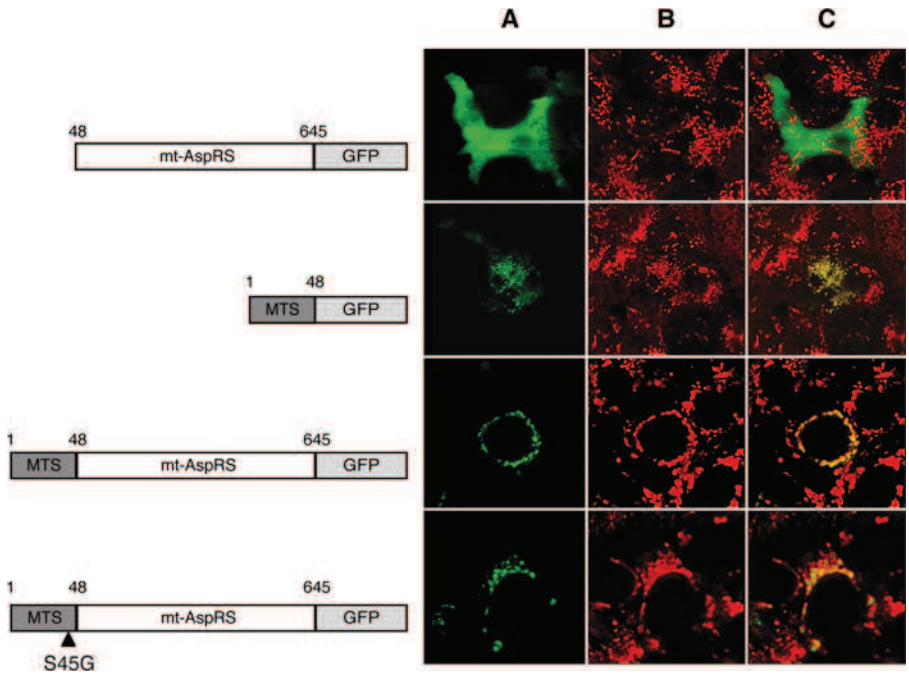


Figure 2 Subcellular localization of GFP-fused mt-AspRS derivatives in human HEK-293T cells

Schematic representations of the four GFP-fused protein constructs are shown. Cells were transiently transfected with full-length (amino acids 1–645) or partial (amino acids 1–48 or 48–645) sequences of mt-AspRS and fluorescence signals were visualized by confocal microscopy. The S45G mutation is highlighted with an arrowhead. (**A**) GFP fluorescence; (**B**) Mitochondrial marker (MitoTracker Red CM-H_2XRos); and (**C**) merge. The integrity of nuclei of HEK-293T cells following transient expression of MTS-mt-AspRS–GFP or MTS(S45G)-mt-AspRS–GFP was verified by staining with DAPI (4′,6-diamidino-2-phenylindole) (results not shown).

(residues 1–48) is exclusively present within the cytosol, both mt-AspRS and (S45G)mt-AspRS possessing their MTS co-localized with mitochondria (Figure 2). It is worth noting that GFP fused to the mt-AspRS MTS alone also co-localized with the mitochondrial-specific marker, showing that the N-terminal 48 amino acids of mt-AspRS are sufficient to guide this protein to mitochondria *in vivo*. In addition, S45G impaired neither targeting nor binding of mt-AspRS to mitochondria. Next, import through the mitochondrial membranes was analysed.

Mutated MTS prevents import of mt-AspRS into mitochondria

The impact of the mutation on the import process was tested by *in vitro* import assays of ^{35}S-radiolabelled proteins into mitochondria purified from HEK-293T cells. mt-GluRS–GFP (N-terminal 81 amino acids of *A. thaliana* GluRS fused to GFP) was used as a positive control. This protein was previously successfully used for assays studying import and processing into isolated plant mitochondria under well-defined experimental conditions [17]. Figure 3(A) shows that this protein is also imported and processed into purified human mitochondria.

Comparable with the *A. thaliana* mt-GluRS variant, a truncated version of mt-AspRS was used (N-terminal 313 amino acids including the MTS) for good resolution of precursor and processed proteins on SDS/PAGE. Processed forms of mt-AspRS, resistant to proteinase K, were observed, confirming proper import of this protein into mitochondria (Figure 3B). The size difference of ~6 kDa between the precursor and processed products is in agreement with the expected removal of the theoretical MTS. Import was restrained as expected in the presence of valinomycin, an uncoupling agent inhibiting protein import [17]. Two observations were made for (S45G)mt-AspRS. First, the mutated precursor protein bound to mitochondria as efficiently as the wild-type form since both proteins were recovered with mitochondria after a sucrose cushion purification step (Figures 3B and 3C, lane 2). This binding is consistent with the confocal microscopy experiments (Figure 2). Secondly, in contrast with mt-AspRS, no processed versions of the protein were visible (Figure 3C, lanes 2 and 3). This absence of mature product strongly suggests that the S45G mutation affects the import process, either at the translocation step through the double mitochondrial membrane or at the MTS-processing step. Actually, it has been shown that a non-processed protein is thermally less stable in the mitochondrial matrix and more rapidly degraded than a mature protein [20]. Therefore the effect of mutation on those two steps has been investigated further. In addition, incomplete degradation of mt-AspRS or (S45G)mt-AspRS by proteinase K (Figures 3B and 3C, lane 3) suggests that *in vitro* import of these human proteins is slower or less efficient than the mt-GluRS–GFP fusion protein (Figure 3A, lane 3).

Mutated AspRS precursor can be cleaved off by mitochondrial peptidases, but is not translocated through the mitochondrial membranes

To discriminate between the two above-mentioned steps, an *in vitro* processing experiment was performed. The same radiolabelled proteins [mt-GluRS–GFP, mt-AspRS and (S45G)mt-AspRS] were incubated with a mitochondrial enzymatic extract from HEK-293T cells containing active mitochondrial-processing peptidases [19]. No processing products were visible in the absence of the mitochondrial enzymatic extract for any of the three proteins (Figure 4). Incubation with the mitochondrial extract led to the expected processed products, with identical differences in size between precursor and processed forms with the ones obtained in the *in vitro* import assay. Interestingly, the processing

© The Authors Journal compilation © 2011 Biochemical Society

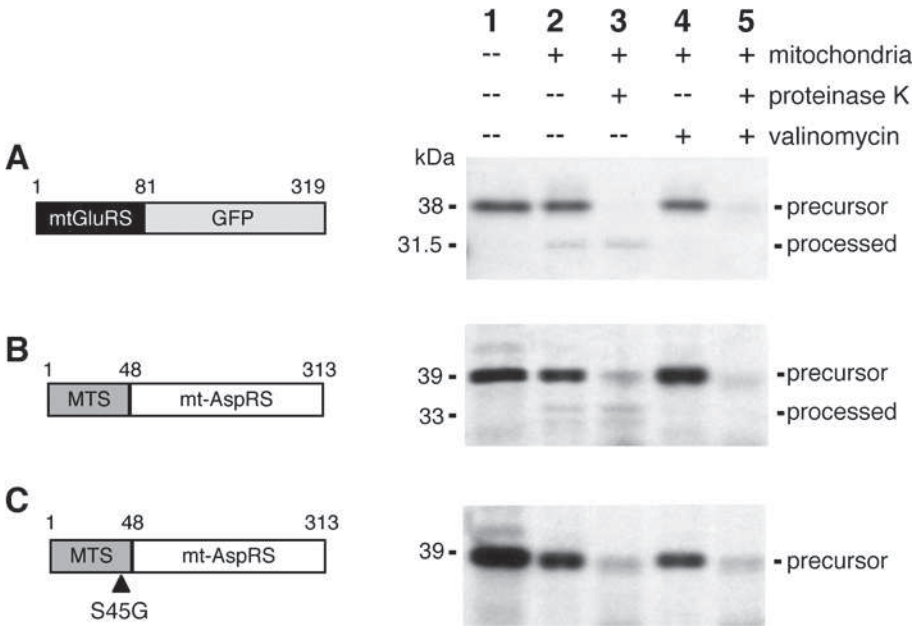


Figure 3 ***In vitro*** **import experiments with purified human mitochondria**

Experiments were performed with three constructs: (**A**) as a control, mt-GluRS–GFP (first 81 amino acids of *A. thaliana* mt-GluRS upstream of the GFP sequence) [17]; (**B**) human mt-AspRS, the first 313 amino acids; and (**C**) (S45G)mt-AspRS. Radioactive full-length precursors obtained by *in vitro* transcription/translation (lane 1), were incubated with mitochondria purified from HEK-293T cells (lanes 2–5) for 30 min, and further purified on a sucrose gradient before separation by SDS/PAGE and autoradiography. Proteinase K degrades proteins present at the mitochondrial surface, but does not affect proteins located inside mitochondria. Pre-incubation with valinomycin inhibits protein import. The molecular mass in kDa is indicated on the left-hand side; molecular masses were calculated based upon marker proteins. Each assay was independently performed at least three times, with no significant variations.

assay shows that the pathogenic mutation S45G did not prevent removal of MTS by mitochondrial peptidases. Therefore, if the mutated (S45G)mt-AspRS is present inside the mitochondria, its MTS could be processed by mitochondrial peptidases. Combining these results indicates that the absence of a processed form of (S45G)mt-AspRS (Figure 3C) appears to be due to non-translocation of the protein and not to some inherent defect in its ability to be processed.

Mitochondrial import failure of an aaRS in LBSL

Many mutations in *DARS2* have been associated with LBSL. In the present study, we have analysed the sole mutation known to be located in the MTS and nearby the predicted processing site (Figure 1). Interestingly, although the import machineries are well conserved between species of the same kingdom, each mitochondrial protein has its own MTS. Consequently, rules regarding the nature and the position of crucial residues for mitochondrial import cannot be precisely defined. The identity of those amino acids is likely to be studied on a case-by-case protein basis, and molecular consequences of mutations within the MTS cannot be anticipated. For instance, in *A. thaliana*, an alanine-to-serine mutation at MTS position 2 of mt-GluRS and mt-MetRS (mitochondrial methionyl-tRNA synthetase), or a threonine-to-glycine change of mt-PheRS (mitochondrial phenylalanyl-tRNA synthetase) respectively enhances, inhibits or does not affect their mitochondrial import [21]. Thus the consequence of the mutation S45G on the import process of mt-AspRS could not be foreseen before the present study; however, the strong conservation of Ser^{45} within the MTS of mammalian mt-AspRSs (Figure 1) suggested a role in their import. In the present study we show that S45G has a clear effect on the import process. Confocal microscopy imaging reveals that it does not affect the targeting, or the binding to mitochondria. *In vitro* import and processing assays show the import, and more precisely the translocation step, to be impaired by the mutation. This is a novel point of impact for a pathology-related mutation regarding a translation machinery protein. Strikingly, S45G has been described to not have any consequence on mitochondrial activity [4]. This remains puzzling, but an answer may arise from the fact that patients are compound-heterozygous for this mutation, the other mutation leading to only a weak expression of wild-type protein. This low level is apparently sufficient to support mitochondrial translation in the cells that were tested. Whether this is different in brain white matter cells has not been studied. It could also be that the mutation affects another, but still unknown, function of the protein [22]. Remarkably, and for unknown reasons, this compound-heterozygosity phenomenon is common to all of the *DARS2* mutations so far reported [5].

Mitochondrial import failure and diseases

Two major types of mitochondrial import defects have been reported to be associated with disorders (for a review, see [23]). Either the components of the mitochondrial import and/or processing machinery are directly affected or mutations are present in the MTS, reducing the mitochondrial import of the corresponding protein. For example, deafness dystonia syndrome as well as a form of cardiomyopathy are caused by mutations in the genes encoding small proteins of the inner mitochondrial membrane [24,25]. Only two cases have been so far reported with mutations affecting the MTS: a pyruvate dehydrogenase deficiency leading to abnormalities in the central nervous system [26], and partial arrest of manganese superoxide dismutase precursor protein within the inner mitochondrial membrane probably modulating susceptibility to various diseases [27]. For LBSL we describe a novel finding: the impairment of the import of a crucial translation machinery protein is correlated with a white

© The Authors Journal compilation © 2011 Biochemical Society

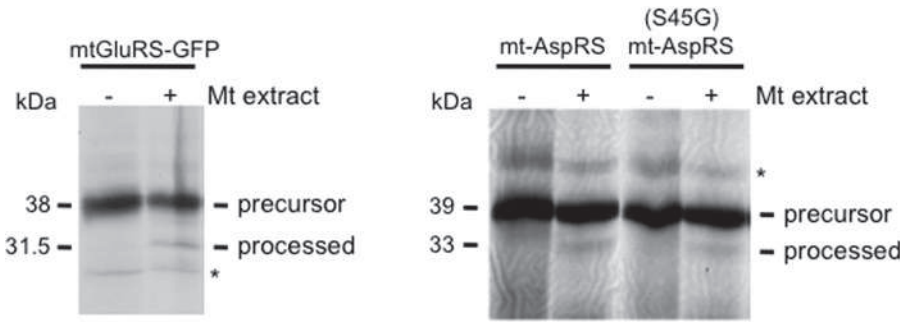


Figure 4 Processing of precursor proteins by a human mitochondrial peptidase extract

The experiment was performed with the same constructs as described in Figure 3. Radioactive full-length precursor proteins of mt-GluRS—GFP, MTS-mt-AspRS and (S45G)MTS-mt-AspRS obtained by *in vitro* transcription/translation were incubated in the absence (−) or in the presence (+) of an HEK-293T mitochondrial enzymatic extract. Three independent sets of experiments were performed and showed identical patterns. Precursor and processed forms are indicated. The molecular mass in kDa is indicated on the left-hand side; molecular masses were calculated based upon marker proteins. *Non-specific products synthesized with the TnT® T7 Coupled Reticulocyte Lysate System.

matter disorder. It can be anticipated that new disease-related mutations will be found in the MTS of proteins involved in any aspect of mitochondrial biogenesis.

AUTHOR CONTRIBUTION

Marie Messmer, Hagen Schwenzer, Gert Scheper and Laurence Maréchal-Drouard performed experiments. Marie Messmer, Laurence Maréchal-Drouard and Marie Sissler designed the research. Laurence Maréchal-Drouard and Marie Sissler supervised the present work. Marie Messmer, Catherine Florentz, Laurence Maréchal-Drouard and Marie Sissler analysed the data and wrote the manuscript. Hagen Schwenzer, Gert Scheper and Marjo van der Knaap critically read the manuscript prior to submission.

ACKNOWLEDGEMENTS

We thank J. Mutterer (Microscopy and Imaging Platform, IBMP, Strasbourg) for help with confocal microscopy imaging. We thank Dr J.C. Paillart (IBMC, Strasbourg) for advice, Dr N. Entelis, Dr I. Tarassov (IPCB, Strasbourg), and Dr F. Sieber (IBMP, Strasbourg) for gifts of human mitochondria, and Dr N. Saod (IBMC Strasbourg) for help in cloning MTS–GFP into pcDNA3.1/CT-GFP-TOPO vector.

FUNDING

This work was supported by the Association Française contre les Myopathies (AFM); the Centre National de la Recherche Scientifique (CNRS); the Université de Strasbourg; and the The Prinses Beatrix Fonds. M.M. was supported by a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche and by l'Agence Nationale de la Recherche [ANR-07-BLAN-0207-01] and H.S. was supported by a fellowship from Région Alsace.

REFERENCES

1 Tuppen, H. A., Blakely, E. L., Turnbull, D. M. and Taylor, R. W. (2010) Mitochondrial DNA mutations and human disease. Biochim. Biophys. Acta **1797**, 113–128
2 Wallace, D. C. (2010) Mitochondrial DNA mutations in disease and aging. Environ. Mol. Mutagen. **51**, 440–450
3 Jacobs, H. T. and Turnbull, D. M. (2005) Nuclear genes and mitochondrial translation: a new class of genetic disease. Trends Genet. **21**, 312–314
4 Scheper, G. C., Van Der Klok, T., van Andel, R. J., van Berkel, C. G., Sissler, M., Smet, J., Muravina, T. I., Serkov, S. V., Uziel, G., Bugiani, M. et al. (2007) Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. Nat. Genet. **39**, 534–539
5 Lin, J., Faria, E. C., Da Rocha, A. J., Masruha, M. R., Vilanova, L. C., Scheper, G. C. and Van Der Knaap, M. S. (2010) Leukoencephalopathy with brainstem and spinal cord involvement and normal lactate: a new mutation in the DARS2 gene. J. Child. Neurol. **25**, 1425–1428
6 Isohanni, P., Linnankivi, T., Buzkova, J., Lönnqvist, T., Pihko, H., Valanne, L., Tienari, P. J., Elovaara, I., Pirttilä, T., Reunanen, M., Koivisto, K., Marjavaara, S. and Suomalainen, A. (2010) DARS2 mutations in mitochondrial leucoencephalopathy and multiple sclerosis. J. Med. Genet. **47**, 66–70
7 Edvardson, S., Shaag, A., Kolesnikova, O., Gomori, J. M., Tarassov, I., Einbinder, T., Saada, A. and Elpeleg, O. (2007) Deleterious mutation in the mitochondrial arginyl-transfer RNA synthetase gene is associated with pontocerebellar hypoplasia. Am. J. Hum. Genet. **81**, 857–862
8 Riley, L. G., Cooper, S., Hickey, P., Rudinger-Thirion, J., McKenzie, M., Compton, A., Lim, S. C., Thorburn, D., Ryan, M. T., Giegé, R., Bahlo, M. and Christodoulou, J. (2010) Mutation of the mitochondrial tyrosyl-tRNA synthetase gene, YARS2, causes myopathy, lactic acidosis, and sideroblastic anemia—MLASA syndrome. Am. J. Hum. Genet. **87**, 52–59
9 Antonellis, A., Ellsworth, R. E., Sambuughin, N., Puls, I., Abel, A., Lee-Lin, S. Q., Jordanova, A., Kremensky, I., Christodoulou, K., Middleton, L. T. et al. (2003) Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. Am. J. Hum. Genet. **72**, 1293–1299
10 Xie, W., Nangle, L. A., Zhang, W., Schimmel, P. and Yang, X. L. (2007) Long-range structural effects of a Charcot-Marie-Tooth disease-causing mutation in human glycyl-tRNA synthetase. Proc. Natl. Acad. Sci. U.S.A. **104**, 9976–9981
11 Bonnefond, L., Fender, A., Rudinger-Thirion, J., Giegé, R., Florentz, C. and Sissler, M. (2005) Towards the full set of human mitochondrial aminoacyl-tRNA synthetases: characterization of AspRS and TyrRS. Biochemistry **44**, 4805–4816
12 Gakh, O., Cavadini, P. and Isaya, G. (2002) Mitochondrial processing peptidases. Biochim. Biophys. Acta **1592**, 63–77
13 Neupert, W. and Herrmann, J. M. (2007) Translocation of proteins into mitochondria. Annu. Rev. Biochem. **76**, 723–749
14 Baker, M. J., Frazier, A. E., Gulbis, J. M. and Ryan, M. T. (2007) Mitochondrial protein-import machinery: correlating structure with function. Trends Cell Biol. **17**, 456–464
15 Bolender, N., Sickmann, A., Wagner, R., Meisinger, C. and Pfanner, N. (2008) Multiple pathways for sorting mitochondrial precursor proteins. EMBO Rep. **9**, 42–49
16 Van Der Laan, M., Hutu, D. P. and Rehling, P. (2010) On the mechanism of preprotein import by the mitochondrial presequence translocase. Biochim. Biophys. Acta **1803**, 732–739
17 Duchêne, A. M., Giritch, A., Hoffmann, B., Cognat, V., Lancelin, D., Peeters, N. M., Zaepfel, M., Maréchal-Drouard, L. and Small, I. D. (2005) Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U.S.A. **102**, 16484–16489
18 Hammen, P. K., Waltner, M., Hahnemann, B., Heard, T. S. and Weiner, H. (1996) The role of positive charges and structural segments in the presequence of rat liver aldehyde dehydrogenase in import into mitochondria. J. Biol. Chem. **271**, 21041–21048
19 Ou, W. J., Ito, A., Okazaki, H. and Omura, T. (1989) Purification and characterization of a processing protease from rat liver mitochondria. EMBO J. **8**, 2605–2612
20 Mukhopadhyay, A., Yang, C. S., Wei, B. and Weiner, H. (2007) Precursor protein is readily degraded in mitochondrial matrix space if the leader is not processed by mitochondrial processing peptidase. J. Biol. Chem. **282**, 37266–37275
21 Pujol, C., Maréchal-Drouard, L. and Duchêne, A. M. (2007) How can organellar protein N-terminal sequences be dual targeting signals? *In silico* analysis and mutagenesis approach. J. Mol. Biol. **369**, 356–367
22 Guo, M., Yang, X. L. and Schimmel, P. (2010) New functions of aminoacyl-tRNA synthetases beyond translation. Nat. Rev. Mol. Cell Biol. **11**, 668–674
23 MacKenzie, J. A. and Payne, R. M. (2007) Mitochondrial protein import and human health and disease. Biochim. Biophys. Acta **1772**, 509–523

© The Authors Journal compilation © 2011 Biochemical Society

24 Koehler, C. M., Leuenberger, D., Merchant, S., Renold, A., Junne, T. and Schatz, G. (1999) Human deafness dystonia syndrome is a mitochondrial disease. Proc. Natl. Acad. Sci. U.S.A. **96**, 2141–2146
25 Davey, K. M., Parboosingh, J. S., McLeod, D. R., Chan, A., Casey, R., Ferreira, P., Snyder, F. F., Bridge, P. J. and Bernier, F. P. (2006) Mutation of DNAJC19, a human homologue of yeast inner mitochondrial membrane co-chaperones, causes DCMA syndrome, a novel autosomal recessive Barth syndrome-like condition. J. Med. Genet. **43**, 385–393
26 Takakubo, F., Cartwright, P., Hoogenraad, N., Thorburn, D. R., Collins, F., Lithgow, T. and Dahl, H. H. (1995) An amino acid substitution in the pyruvate dehydrogenase E1 alpha gene, affecting mitochondrial import of the precursor protein. Am. J. Hum. Genet. **57**, 772–780
27 Sutton, A., Khoury, H., Prip-Buus, C., Cepanec, C., Pessayre, D. and Degoul, F. (2003) The Ala16Val genetic dimorphism modulates the import of human manganese superoxide dismutase into rat liver mitochondria. Pharmacogenetics **13**, 145–157

Received 17 November 2010/1 December 2010; accepted 1 December 2010
Published as BJ Immediate Publication 1 December 2010, doi:10.1042/BJ20101902

© The Authors Journal compilation © 2011 Biochemical Society

3.3.4 Impact of mutations in mt-AspRS on sub mitochondrial localization

The sub-mitochondrial localization (or distribution) of the wild type mt-AspRS was intensively investigated either in HEK293 or ectopic protein in BHK21 cells. We used the expression system in BHK21 to investigate the import and sub-mitochondrial localization of the mutant R58G, Q184K, R263Q, L613F and L626Q of mt-AspRS (**Figure 50**). These last experiments were performed by Loukmane Karim (IBMC, Strasbourg, France).

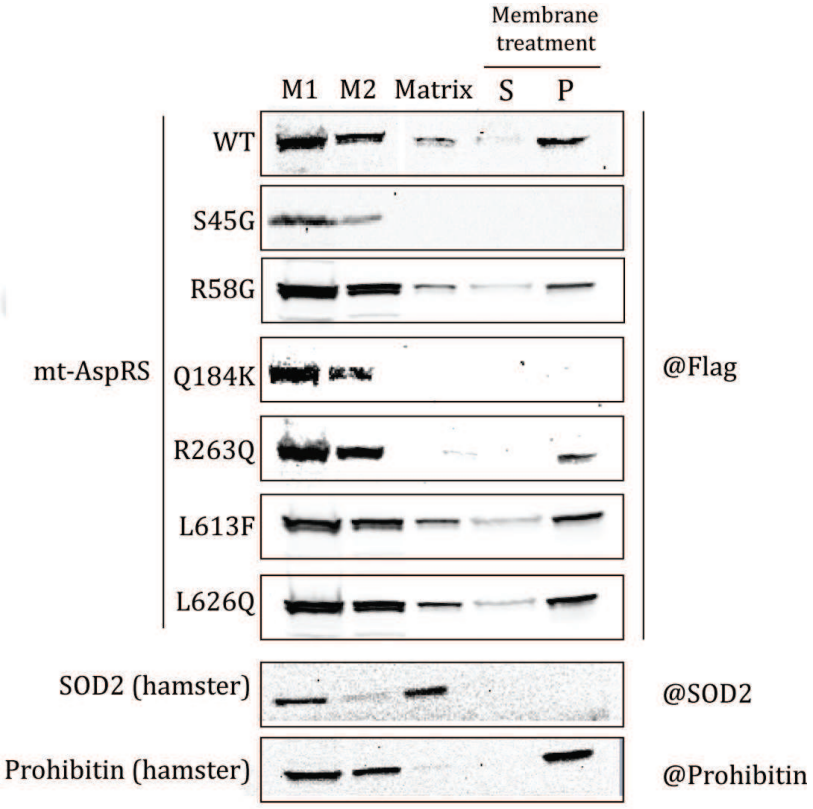


Figure 50: Sub-mitochondrial localization of mutant AspRSs. Mutants of mt-AspRS were recombinantly expressed in BHK21 cells, mitochondria purified and fractionated. Aliquots of the samples were separated on SDS-PAGE. Mutants of mt-AspRS were detected using anti-Flag antibody. Antibodies against SOD2 and Prohibitin were used to detect endogenous SOD2 and Prohibitin as control of the quality of the fractionation. M1 and M2 stand for mitochondria before and after percol gradient centrifugation, respectively. S and P stand for supernatant and pellet fraction obtained after Na_2CO_3 treatment.

We performed mitochondrial purification from BHK21 cells, expressing the mutant mt-AspRSs, and fractionated the mitochondria into matrix, peripheral and

integral membrane fractions. The sub-mitochondrial distribution of all proteins was determined by western blot analysis using an anti-flag antibody (Figure 8). While the pattern was unaffected by mutations R58G, L613F and L626Q, the mutation R263Q localized the mt-AspRS to the membrane fraction, and the mutation Q184K completely abolished the sub-mitochondrial distribution of mt-AspRS. The quality of the fractionation was assessed using antibodies against SOD2 (matrix) and Prohibitin (membrane) proteins.

Concerning the Q184K mutation, the corresponding protein is visible only in fractions containing intact mitochondria (M1 and M2; **Figure 50**), but not in the matrix or the membrane fractions. We have demonstrated in the previous part that this mutation not only affects the expression yield but also the solubility of mt-AspRS (**Figure 49**), leading to a high tendency to aggregate (article #3). It however remains unclear whether mt-AspRS Q184K aggregates before or after the entry to the mitochondria. To distinguish if the mutant protein is blocked as an aggregated protein at the level of the outer membrane or aggregates upon entry to mitochondria, a 'mitochondrial protection assay' needs to be performed. The principle of this experiment is to submit purified mitochondria to protease degradation (e.g. Proteinase K or Trypsin; (Shalak et al., 2009) and to investigate whether the mutant mt-AspRS is protected or not by the mitochondrial double membrane.

The mutant R263Q was found in the soluble fraction of the lysate forming a dimer but also a significant fraction was found in the insoluble fraction indicating aggregation (**Figure 49**). After fractionation, the mutant was detected solely on the membrane fraction. It is worth recalling that precursor proteins pass the mitochondrial membranes as unfolded proteins, and are folded only after processing inside the mitochondria (reviewed in e.g. Voos, 2013). It is thus very likely that proteins shown at the level of the membrane correspond to folded dimer of the mt-AspRS rather than aggregates. The above described 'mitochondrial protection assay' will confirm this assumption (experiments under way). This is the first example of a mutation that affects the correct sub-mitochondrial distribution of the protein. The molecular mechanism(s) underlying this observation remains to be seen. We may be able to answer this question only when the functional contribution (role) of the mt-AspRS in each of these fractions is known.

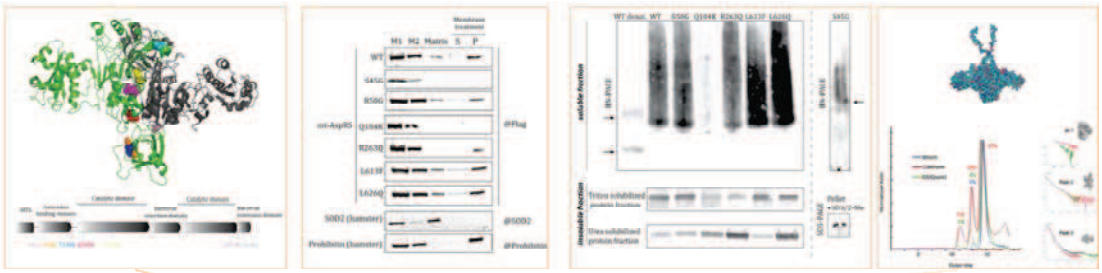
It is worth mentioning that two mutations (Q184K and R263Q) are found in patients in a compound heterozygous state with the R76SfsX5 mutation. As already recalled, the last mutation led to a decreased, but not zero, expression of the WT mt-AspRS. The engineered expression system, using MVA-mediated transfection, will allow us to simultaneously express both WT and mutated constructions, so that we can somehow mimic the natural situation found in the patients. It would thus be interesting to see if we can observed, under our experimental conditions, an influence of the presence of the two constructions on the general solubility of the proteins, or on the final sub-mitochondrial localization. It would also be interesting to investigate if the presence of the WT construction would modify the random association of WT and mutant monomers into either homodimers or heterodimers. It is likely that the strong tendency of R263Q mutant to aggregate (reducing its occurrence as either homodimer or heterodimer) indirectly favors the formation of the WT homodimers (with unaffected aminoacylation activity), which guarantee a basal translation activity.

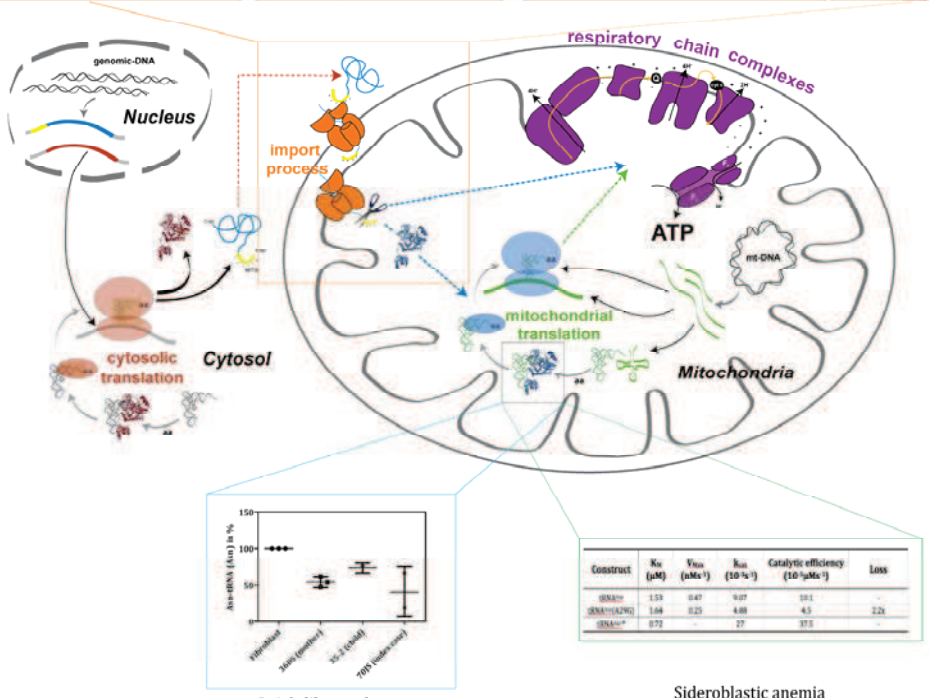
4 Achievements at a glance

- We investigate the influence of $tRNA^{Asp}$ mutation in patient with sideroblastic anemia
- We exclude that mutation in NARS2 found in patient with Leigh-like syndrome have a loss of aminoacylation function
- We showed that one amino acid change found in patients with LBSL can influence the solubility; import and sorting into mitochondria of mt-AspRS

5 Graphical Summary

LBSL: *In vivo and in vitro* investigation of solubility,oligomery and import of mt-AspRS mutants





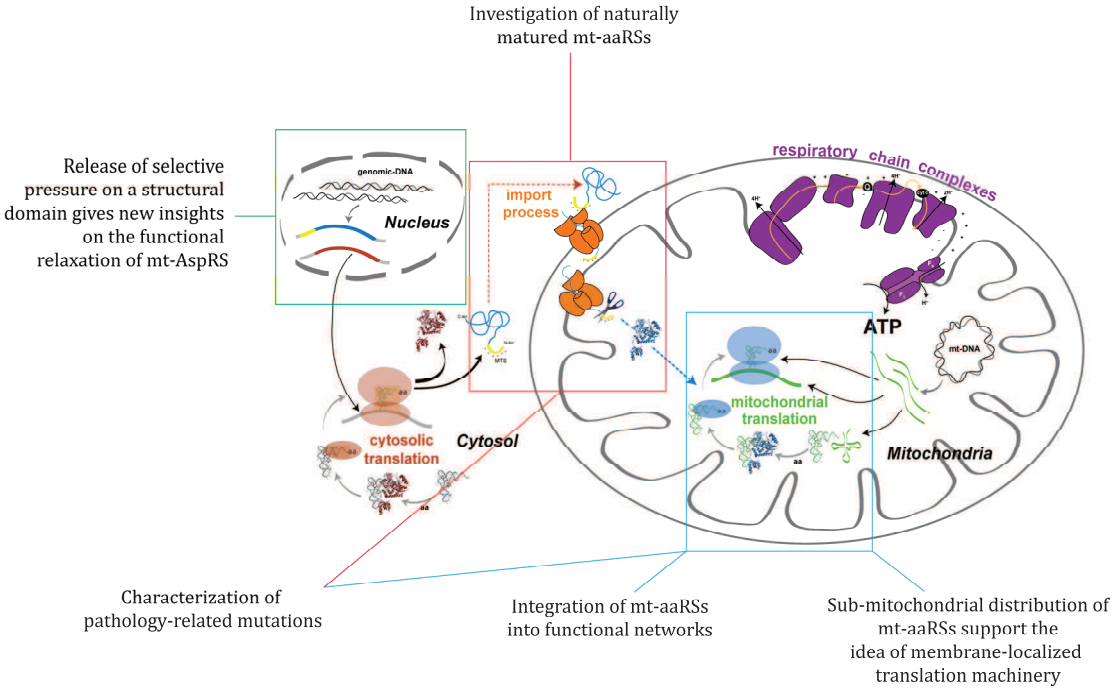
Leigh-like syndrom:
In vivo detection of aminoacyl-$tRNA^{Asn}$

Sideroblastic anemia
In vitro analysis of aminoacylation efficiency of $tRNA^{Asp}$

CONCLUSION

Conclusion

1 Summary at a glance



An important goal of our research group is to understand the fundamental mechanisms of the human mitochondrial translation machinery. In particular our interest focus on the mt-aaRSs and mt-tRNAs. In the past years, the fundamental properties of mt-aaRS have been intensively investigated, resulting in a better understanding of at least some of those. Among those, our lab contributed during the last years to establishment of the first crystallographic structure of a mt-aaRS (mt-TyrRS; (Bonnefond et al., 2007) or the biophysical and structural analyses of mt-AspRS (Neuenfeldt et al., 2012).

Beside these fundamental researches, our fields of investigation have now to turn to more groundbreaking aspects. While the organization of human cytosolic aaRSs is well explored and their involvement into alternate functions clearly established for some of them, the properties of the human mt-aaRSs remain scarcely known. However,

the discovery of connections of mt-aaRSs to diseases leads to an emerging demand for basic knowledge on functional properties and organization of mt-aaRSs. Our research group has been particularly focused on the functional properties, structure and role of mt-AspRS in mitochondrial diseases (e.g. Scheper et. al., 2007, Messmer et al., 2011, van Berge et al., 2012a,). It turns out that more knowledge about the mt-aaRSs is needed to understand the molecular background of some mitochondrial disease. To serve this, we continued some existing research lines and established new one concerning the mt-AspRS. In line with the objectives our group, my objectives were:

- To understand the role of the alternatively spliced transcript mt-AspRSΔExon13. The corresponding protein lack a peptide in the BID for which the function is so far unknown.
- To integrate the mt-AspRS into new cellular networks. These studies combine on one side the need to better understand the biogenesis of the mt-aaRSs and their integration into the mt-translation system, but also their integration into novel non-translational function.
- To evaluate the impact of some pathology-related mutations.

The outline of this conclusion is separated in three parts:

i) We evaluate the results concerning the objectives stated in the introduction
ii) We show the limitation(s) of the results and highlight perspectives
iii) We give concrete recommendations for following investigations

2 Evaluation of the obtained results

2.1 Released selective pressure on a structural domain gives new insights on the functional relaxation of mitochondrial aspartyl-tRNA synthetase

2.1.1 General implication

The goal of the investigation presented in ***chapter I*** was to analyze the functional role of the additional isoform of mt-AspRS mRNA, deleted of one exon. Our observations based on an alternatively spliced form of mt-AspRS impacting the bacterial insertion domain (BID) lead us to three possible hypotheses:

1. the isoform codes for a protein with a new alternative function
2. the isoform coexist in regulatory manner with full-length mRNA
3. the isoform represent non functional isoform ("low fitness form") free to evolve to functional isoform ("high fitness form")

Of course the first hypothesis is the most attractive one. It is known that alternative splicing is a mechanism to enlarge the proteomic variation of the genome (Graveley, 2001). For example, the mRNA coding for the cyt-TrpRS can undergo an alternative splicing event that generates an active angiostatic cytokine (Wakasugi et al., 2002). In the course of our investigations, several evidence show up that conflict with this hypothesis:

(i) We did not identify a possible protein.
(ii) We could not express the protein.
(iii) Knock-down experiments indicated no functional relevance of the putative protein for the viability of the cell and no regulatory role of this isoform in the expression of the full length protein.

(iv) *In silico* analysis revealed a high sequence divergence occurring only in this domain indicating a loss of selective pressure and thus a possible loss of function.

These results lead us to refine our hypothesis and to propose that the mt-AspRS undergoes a functional relaxation in comparison to the bacterial and plant homologs. The discussion of the arguments favoring this hypothesis can be found in the discussion part of article #1.

Among other outcomes, one important conclusion from this work is that the BID, likely already present in bacterial ancestor, may have lost its functional relevance in human mt-AspRS. This contrasts with high sequence conservation and high selective pressure on this domain in viridiplantea mt-AspRS and bacterial AspRS consistent with a possible functional role in these organisms (Reviewed in Giegé and Rees 2005), which is so far unclear. Clear interpretations about the role of this domain (in bacteria) and the consequences of the loss of function (in mitochondria) can only be done once the functional role in the bacterial ancestor is deciphered. However, gains and losses of functional domains appear to be common evolutionary mechanisms to increase the complexity of an organism or to drive the functional variability of the proteome (Guo et al., 2010; Mukherjee et al., 2013).

2.1.2 Value for a better understanding of mitochondrial diseases

The fact that the BID has no functional relevance in human mt-AspRS can be important for the evaluation of the impact of pathology-related mutation. It can be speculated that pathology-related mutations affect only functional relevant residue or domains. Thus, the absence of a function would decrease the impact of the mutation. Indeed, despite the fact that more than 25 pathology-related mutations were found in all parts of the mt-AspRS, there was no mutation found in the BID, so far. Of note, only two nonsense mutations, localized in the BID, lead to premature stop codon and may impact only the amount (the expression?) of protein (bookchapter #2).

2.2 Sub-mitochondrial distribution of mt-aaRSs support the idea of a membrane localized translation machinery

2.2.1 General implication

Mitochondrial aaRSs are synthesized in the cytosol, but must be translocated across the mitochondrial inner membrane, processed to removed the MTS sequence in order for the protein to be properly folded and integrated in the mitochondrial translation system. The goal of the investigation presented in ***Chapter II*** was to establish the sub-mitochondrial organization of the mt-aaRSs (sub-mitochondrial location and repartition of the full set of mt-aaRSs). Our study systematically investigated, for the first time, the sub-mitochondrial organization of the whole protein family involved in mt-translation. We showed that some proteins are double distributed in membrane and matrix, whilst others are only found either in matrix or in the membrane. Altogether, our results indicate that a major fraction of mt-aaRSs is located on the membrane. Our result is in line with the hypothesis that mitochondrial translation is occurring at the close vicinity of the mitochondrial membrane. The following arguments, in combination with your own observations, are further supporting this hypothesis:

1. A membrane-association was already shown for other translation factors like the mt-ribosome (Liu and Spremulli 2000) and the mt-EF-Tu (Suzuki et al., 2007)
2. All mt-DNA encoded proteins are hydrophobic and membrane-embedded proteins. They would thus need to be translated in a hydrophobic environment to guarantee a proper folding.

It would be interesting to see if mt-aaRSs interact directly with the mt-ribosome to promote the channeling of the aminoacylated-tRNA by the EF-Tu_{mt} to the ribosome. Indeed, it has been proposed that mt-tRNAs undergo a tight channeling process to protect them from nuclease degradation (Belostosky et al., 2012). Our preliminary experiments identified several mt-ribosomal proteins in the elution fraction obtained by co-immuno precipitation of mt-AspRS that are in line with this hypothesis. Of note, it is striking to see mt-aaRS in two different sub-fractions of the mitochondria. One can

speculate that localization of the mt-aaRS may modify the functionality, but this possibility must be tested. In addition, if this is true; it can then be hypothesized that the membrane serves as platform similar to the three scaffold proteins (p18, p38 and p43) in cytosolic MARS complex (Mirande et al., 1982). To recall, the MARS complex serves as a molecular reservoir ("depot hypothesis"), where anchored proteins perform aminoacylation, while released proteins are free to perform alternate functions (Ray et al., 2007).

When we further analyzed the membrane fraction of the mt-AspRS belonging to the double distributed proteins, it turned out that it is strongly anchored to the membrane. We exclude that the protein is only attached to the membrane through ionic or electrostatic interactions or by hydrophobic transmembrane motifs. Thus, we hypothesize that post-translational modification could play an important role for the anchorage of at least the mt-AspRS. Of note, it is striking to see that mt-aaRS is in two different sub-fractions of the mitochondria. One can speculate that this is maybe an evidence for functionally different proteins.

2.2.2 Value for a better understanding of mitochondrial diseases

The wild type picture of the sub-mitochondrial distribution can be compared to possible mutant situation and thus contributes to a better understanding of mt-diseases. And indeed, this knowledge helped to identify mutations with a potential impact on the sub-mitochondrial distribution of the mt-AspRS.

Moreover, we provide a model of the organization of aaRSs and translation machinery, which indicates the possibility of alternative functions in a different sub-mitochondrial location. This model can serve as starting point, which may lead to the discovery of non-translational function and thus to the integration of aaRSs in novel pathways including pathology related pathways.

2.3 Investigation of naturally matured mitochondrial proteins

2.3.1 General implication

Expression system for naturally maturated mammalian proteins imported into mitochondria

The goal of the investigation presented in ***chapter III*** was to establish and validate a unique and robust tool to enable the exploration of the impact of precursor sequences on the import process and ultimately to have access to naturally processed mitochondrial proteins. Our aim was to obtain further insight into the process of protein localization during translocation into the mitochondria and inside mitochondria.

The biosynthesis of the mt-aaRSs comprises three different compartments before they reach their final destination. The place of encoding is the nucleus, the place of synthesis is the cytosol, and the place of activity is the mitochondria. *In vitro* investigations require the knowledge of the maturation process of mt-aaRSs to access recombinant soluble and active proteins, both were shown to be dependent on the correct N-terminal sequence definition. A powerful tool to facilitate these investigations was developed and validated (article #2). Any sequence of interest can be considered and expressed, so that this method has unlimited applications (e.g. determination of N-terminal maturation site, mutational analysis, production and purification of mitochondrial processed proteins) to investigate nucleus-encoded human (mammalian) proteins of mitochondrial location.

The combination of MVA-mtr expression along with our novel incremental proteolytic based prediction of N-terminal peptide with mass spectrometry data analysis will undoubtedly be of great utility towards the systematic deciphering of N-terminal residues for mammalian mitochondrial proteins. The approach is extremely rapid and inexpensive when compared to standard techniques.

The MTS define the properties of the mt-aaRSs

Several predicted matured proteins corresponding to mt-AspRS were cloned and expressed (Gaudry et al., 2012). Significant differences in their expression efficiency, solubility, activity and crystallizabillity were observed (Gaudry et al., 2012.). Indeed, the inaccurate prediction of the MTS was previously shown also to be responsible for suboptimal expression of other human mt-aaRSs. Only the LeuRS variant deprived of its 39 N-terminal amino acids was sufficiently overexpressed in *Escherichia coli*, efficiently purified, and fully active, while the variant deprived of the predicted 21 amino acids remained insoluble (Bullard et al., 2000).

We experimentally determined the N-terminus of the mt-AspRS. The preparation of sufficient material of processed mt-AspRS by IP turned out to be inefficient for N-terminal sequencing. Thus, the established and validated expression system facilitated the determination of the N-terminal sequence of mt-AspRS. Strikingly, the defined N-terminal sequence corresponds to the recombinant protein with a reduced expression and solubility compared to other predicted forms (Gaudry et al., 2012). This suggests that additional (yet unknown) parameters, such as partner proteins or physiological conditions may play a role *in vivo* in stabilizing the proteins.

We theoretically analyzed the influence of the MTS on the pI of the protein. Strikingly, we observed three populations of aaRSs being predicted of different pI, above or below 7. Thus, how is it possible that proteins with these different physical-chemical properties exist in the same environment? We hypothesize that some additional (yet undeciphered) parameters exist which allow this observation. Those parameters can be different partners, post-translational modifications or different sub-mitochondrial localization of the aaRSs. Of note, this classification may also correlate with the origin of the aaRSs. It was already found that the pI from mt-aaRS are different from their homologs in other domains of life by a more alkaline pI (pI *E.coli* < *T.Thermophilis* < cytosolic < mitochondrial) (Sissler et al., 2008).

2.3.2 Value for a better understanding of mitochondrial diseases

Knowledge of the import process is important to evaluate the impact of pathology-related mutations. It will be of great value to decipher all signals important for the successful import into mitochondria. One of the widely used methods to characterize effect of pathology related mutations are the determination of aminoacylation kinetics *in vitro*. This requires access to recombinant soluble and active mt-aaRSs. Thus, a correct N-terminal sequence definition of the full set of mt-aaRS will be of enormous help.

Several pathology-related mutations in mt-aaRS were found in the vicinity of the MTS (summarized in bookchapter #2). And indeed, we showed that a pathology-related mutation S45G prevent the import of mt-AspRS into mitochondria (Messmer et al., 2011). Interestingly, this mutation is 12 amino acids upstream of the experimentally determined clivage site of the MTS sequence, but impairs the mitochondrial translocation suggesting that signals outside the classical MTS may play a role in the import process.

Our preliminary mutational analysis of MTS of mt-AspRS and SOD2 may go in the same line. Of note, we do not know, if the MTS^{SOD2}-AspRS construct is compatible with proper translation and folding. However, it is interesting to observe that despite both mutants MTS^{SOD2}-AspRS or MTS^{Asp}-SOD2 include full MTS sequences, only the SOD2 mutant was imported.

In general, the import of proteins in mitochondria can be affected in two ways. Either the mitochondrial import machinery is directly affected or mutations affect the MTS of the proteins, which hampers their import (reviewed in MacKenzie and Payne 2007). The mutation S45G in mt-AspRS of patients with LBSL goes in line with the latter group. As an other example, a pyruvate dehydrogenase deficiency leading to abnormalities in the central nervous system and partial arrest of manganese superoxide dismutase precursor protein within the inner mitochondrial membrane probably was associated with various diseases (Takakubo, et al 1995; Sutton et al. 2003).

A defect in mitochondrial import would likely result in a reduced level of active protein inside the mitochondria. In the case of mt-aaRS, this would affect the aminoacylation of the corresponding tRNA and thus lead to a reduced amount of

functional tRNA, which can be used for synthesis of the RC complex. Consequently, the mutation can potentially lead to impaired mitochondrial activity.

2.4 The integration of mt-aaRSs into functional networks remains open

2.4.1 General implication

The goal of the investigation presented ***Chapter IV*** was it to identify possible interacting components of the mt-AspRS and to integrate it into a functional interaction network. Our results were finally not sufficient to propose any possible partner proteins for the mt-AspRS with confidence. Thus, the integration of the mt-AspRS in functional network with other proteins remains still unclear. Nevertheless, the predicted association network and the list of putative partners obtained by co-IP will help to guide future experiments. In addition, we showed evidence that the mt-AspRS is involved in putative high molecular complex of ~1 000kDa. This is in line with another study showing that the mt-TyrRS is within a complex of similar size (Sasarman et al., 2012).

Our functional association network indicated that the mt-AspRS is co-expressed with ribosomal proteins. In addition, co-immuno precipitation experiments suggest that ribosomal proteins may interact with mt-AspRS. This observation brings both actors of the translation machinery into close proximity, which goes in line with the hypothesis that the mitochondrial translation machinery is located at the inner membrane. Additional weight to this hypothesis is brought by *in vitro* trapping experiment performed by Engelhard and co workers in 2011. They identified the mt-MetRS, mt-AlaRS and mt-GluRS as possible targets of Thioredoxin 2 a protein-thiol-reducing oxidoreductase. Thus, a redox regulation of mitochondrial translation was proposed (Engelhard et al., 2011). Of note, the inner membrane potential is a sensor for the redox / “health” status of the mitochondria which further supports the argument for a membrane location of the mt-translation machinery. It is hypothesized that mitochondria have to respond quickly to any environmental change and energy requirements, which may thus requires a physical association of genes and gene expression components with bioenergetics membranes to ensure a fast answer (Lane

and Martin 2010). For example, it is well known that the membrane potential can regulate the mitochondrial import machinery (reviewed in e.g. Kulawiak et al., 2013). Consequently, it would be interesting to find partner proteins that confirm this hypothesis. Of note, the redox-status also influences the pI of a protein. It thus has to be taken into account for the further investigation of mt-aaRSs peculiarities in regard to their net charge.

2.4.2 Value for a better understanding of mitochondrial diseases

Nowadays, only the interaction with the canonical partner molecule tRNA is described and pathology-related mutations in mt-aaRSs have been demonstrated to affect these interactions between both molecules (e.g. Riley et al., 2010 or Elo et al., 2012). In addition, several studies indicated that the "minimal" organization of mt-aaRSs in their oligomeric state could be disturbed by pathology-related mutations (Pierce et al., 2011; van Berge et al. 2012a). However, possible partner proteins are not identified yet, but nevertheless we have evidence that the mt-TyrRS and mt-AspRS are organized in a macromolecular, yet unknown and uncharacterized, complex. This/These complex(es) should be considered to reveal novel molecular mechanisms underlying a pathological mutation.

2.5 Characterization of pathology-related mutation

The goal of the investigation presented in ***chapter V*** was to initiate the *de novo* analysis of some newly discovered pathology-related mutations in tRNA, but also in mt-aaRS genes. In addition, we further analyzed known and characterized mutants of mt-AspRS in regard to possible physical-chemical defects and their impact on their sub-mitochondrial organization.

2.5.1 Sideroblastic anemia: Mutation in $tRNA^{Asp}$ does not impact aminoacylation

Our investigations showed that mutation A26G in the $tRNA^{Asp}$ do not impact the *in vitro* aminoacylation reaction. However, our experiment only investigated one step of the lifecycle of the mt-tRNAs and did not consider possible cumulative effects of mis-spliced, incorrectly post-transcriptionally modified or mis-folded tRNA. We hypothesize that several mild effects accumulate and lead to the phenotypic expression. This cumulative effect makes difficult to evaluate in regard the pathological impact of the mutation.

2.5.2 Leigh like syndrome: Mutations in $tRNA^{Cys}$, mto1 and AsnRS with unknown molecular link

Pathology-related mutations were also found in patients having the Leigh-like syndrome. In this case, the situation was even more difficult to analyze than in the previous case, since several potent-pathologic mutations were found in different genes. The potential-pathogenic mutations were found in a mitochondrial-encoded $tRNA^{Cys}$, the nucleus-encoded mt-AsnRS and the nucleus-encoded mto1. We tested the effects of the mutations in all three molecules and could not detect any obvious defect. The only effect we observed was a reduction in $tRNA^{Asn}$ and $tRNA^{Cys}$ levels. However, we could not assign this reduction of tRNA level to one mutation in particular. In the case of the mutations in mt-AsnRS, it would be now interesting to investigate other aspects than the mitochondrial aminoacylation properties, such as the impact of the mutation on the import, solubility, dimerization or the sub-mitochondrial localization of the protein. We will be able to answer these questions since we have now the tools and knowledge about the wild-type picture.

2.5.3 Leukoencephalopathy (LBSL): Mutations in mt-AspRS impact solubility, import and/or sub-mitochondrial distribution

We investigated seven mutations found in the mt-AspRS of patients with LBSL and could connect at least two of them with new pathological pathways. As mentioned above, we showed for the first time that mutation (S45G) could prevent the mitochondrial import of a mt-aaRS (Messmer et al., 2011). In the second example, we systematically investigated six mutations and clearly showed for the first time for aaRSs that a single amino acid exchange (Q184K) can lead to structural change and aggregation (article #4). Of note for the Q184K mutant, previous investigations indicate only a deficit in expression without a significant loss in aminoacylation (van Berge et al. 2012a). Thus, with our finding we could show an alternative explanation to the molecular link of the pathology-related mutation to the phenotypic manifestation.
In addition for another mutation (R263Q) we provide evidence that this mutant impact the sub-mitochondrial distributions of the AspRS.

3 Limitation and perspectives

In my thesis I offered a number of new technical methods and scientific insight on the biology of mt-aaRSs in humans. Nevertheless the investigations also disclose some limitations, which have not been solved so far and should be the task of future investigations.

3.1 Technical and scientific limitations

In ***chapter I***, we analyzed a putative alternative spliced transcript. So far, we were unable to detect or to express the corresponding protein. This leaves the question open if this protein would be functional. Preliminary experiments with poorly expressed, highly chaperone contaminated, recombinant mt-AspRSΔExon13 indicated a reduced activity.

The identification of the sub-mitochondrial distribution of aaRSs presented in ***chapter II*** was highly informative and will be undoubtedly of great help for further investigation. Nevertheless, the current model is a mean picture of the distribution of the mt-aaRSs. Further investigation should extend this picture to cell lines other than HEK293 or BHK21, and to consider different physiological conditions as different growth conditions or cellular stresses. In addition, the functional relevance of the double distribution of the mt-aaRSs within the matrix and along the membrane remains to be deciphered. Preliminary experiments evaluating the distribution of five tRNAs within the same fractions only indicate a clear accumulation in the membrane fraction for the $tRNA^{Asp}$. The analysis of the full set of tRNAs and the evaluation of the distribution of aminoacylated-tRNAs will shed further light on the functional relevance of mt-aaRS on both fractions. Of note, we hypothesized that the two sets of proteins may have different functions. It cannot however be excluded that one of the forms is a transit intermediate of import.

In this thesis two general approaches were proposed, which facilitate the investigation of the mitochondrial proteins. One of these approaches was the proteomic analysis of the sub-submitochondrial fraction. It turns out that the initial experiments were able to estimate the enrichment of proteins in each fraction and to analysis the enrichment of functional pathways in the fractions. The major drawback so far is our disability to distinguish between real cytosolic contamination (false positive) and real mitochondrial proteins (false negative), which were not annotated yet as mitochondrial. However, once a convenient pipeline is established overcoming this problem, the sub-mitochondrial fractionation combined with a systematic proteomic analysis will be a tool of enormous value to further decipher the sub-mitochondrial proteome and apply this for comparative analysis of wild type and mutant proteomes.

One of the major aims of this thesis was to push forward the understanding of molecular links of pathology-related mutations to the pathogenic manifestation. Investigating the impact of pathology related mutations confront us with the diversity of possible molecular effects. The distance between the birth of a mt-aaRS and its final role in ATP synthesis is very long, so that the gap between the comprehension of the molecular impact of mutations in macromolecules and a dysfunction of the respiratory chain is very large (bookchapter #2). Our studies investigated for the first time systematically other aspects of mt-aaRSs beside their aminoacylation activities but could only shed light on some points of the mechanisms in which the mt-aaRS are involved. In addition, our research is concentrated on the pathological impact of mutations on the mt-AspRS, it would be of great utility to extend our view to other mt-aaRS involved in mt-diseases.

Finally, I want to recall the major outcome of the comprehensive analysis of mutations in mt-aaRSs prepared during my thesis. Our analysis of all known mutations involved in mt-diseases suggests that pathogenic mutations do not just effect translation, but have a much broader effect on cellular/mitochondrial function. We showed that:

1. There is no common combination of affected steps that correlate to the various observed phenotypic expressions.

2. There is obviously no “favored” mt-aaRS gene.
3. There are also no “favored” affected parts of the protein
4. There is high tissue diverse manifestation of the pathology-related mutation

These facts complicate the prediction of the impact of pathology-related mutations and the interpretation of the results. So far mutations in mt-aaRSs were analyzed only case by case. Thus, it is necessary to further explore yet unknown cellular properties of aaRSs to be able in the future to develop general concepts, which may facilitate the prediction of the impact of mutation in mt-aaRS, not only restricted to the test tube in the laboratory but moreover in patient.

To date, our studies have been limited to human kidney cells as the most convenient model in terms of practicability and handling. In addition cell samples of patient are, in most cases, fibroblast or myoblast cells. One must note that most of the mt-diseases connected to aaRSs concern specifically the neuronal system. Thus, it remains open if the pathologic manifestation of the mutation is the same in fibroblast or myoblast cell culture. Thus, it will be of enormous help to develop animal models to further investigate the general mechanism of mutation in mt-aaRS. Fly and mouse models were already of great help to understand diseases in connection with mt-MetRS (Bayat et al., 2012) and cytosolic aaRSs (e.g. cyt-TyrRS, Storkebaum et al., 2009; cyt-AlaRS, Lee et al., 2006).

4 Recommendation for future short- and long-term research directions

The obtained results of this thesis open new direction for future research. They gave new starting points to investigate the emerging field of mitochondrial diseases being connected with nucleus-encoded mt-aaRSs. Some recommendations for further research are given in following text.

4.1 Deciphering the sub-mitochondrial organization

- **Characterization of the mode of anchorage of mt-AspRS**
  - 1st Hypothesis: Lipidation as C-terminal glycosyl phosphatidylinositol (GPI) anchor, N-terminal myristoylation, S-myristoylation, S-prenylation) are responsible for membrane anchorage
  - 2nd Hypothesis: C-terminal end of the mt-AspRS showed highly flexible structure which may serve as anchor to the membrane

- **Deciphering of the influence of tissue specificity, physiological parameters to organization of mt-aaRS**
  - Hypothesis: The current model is only a mean picture of the distribution of mt-aaRS and the distribution is may dependent from external conditions or cellular state (in regard to cell cycle)

- **Extended analysis of the distribution of mt-tRNAs in the fractions**
  - Hypothesis: Those fraction of aaRSs in which aminoacylated-tRNAs can be found indicate the fraction of translation active aaRSs.

- **Definition of the MTS of the whole set of mt-aaRSs**

- Hypothesis: The knowledge of the correct N-terminal sequence will improve the *in vitro* expression and will help to decipher the rules for the mitochondrial import and sorting of aaRSs

4.2 Deciphering new functional networks

- **Systematic cross-linking approach to further identify and confirm possible partner proteins of mt-AspRS**
  - Hypothesis: Evidence by cross-linking experiments indicates an involvement of mt-AspRS in a complex. The stabilization of interaction(s) by cross-linking will facilitate the purification the identification of protein partners.

- **Co-purification of mt-AspRS partner proteins of each sub-mitochondrial fraction**
  - Hypothesis: Identification of different partners can indicate different cellular functions

- **Identification of post-translational modifications of mt-AspRS and other aaRSs**
  - Hypothesis: Post-translational modifications such as phosphorylation have been shown to serve as trigger for functional switches

4.3 General suggestions for further research direction:

- Mt-aaRSs have to be considered as protein with a long way from the birth to the final activity in the mitochondrial translation. Mutations may impact not only the translational activity, but also the biogenesis, import or integration of the mt-aaRS in the mitochondrial translation machinery.

- Establishment of animal models to study genetic disease in connection with mt-aaRSs. Cell culture methods only show a restricted picture of phenotypic

expression and will not lead to a full understanding of tissue specificity. In addition future medical treatment need a well-understood model organism.

- Understanding concerning multi-tasking proteins is required to characterize overall protein function in fundamental and industrial research. A disease with apparently unrelated symptoms may result from defect in a single multifunctional protein. A given treatment may fail because it doesn't address all the functions of a missing protein, or a drug might have side effects because it interferes with protein activities distinct from the target functions

In summary, to our knowledge, this is the first investigation aiming to decipher and integrate the mt-aaRSs in new functional networks and non-translational function. In agreement, investigations towards alternative organization of mt-aaRSs are requested to understand mt-diseases and to develop potential therapeuthic strategies. Obviously, the present thesis sheds light on some new aspects of the peculiarities of mt-aaRSs. Also, it is now obvious that, in the long term, the influence of pathology-related mutations of mt-aaRSs in the context of their partners and organisation must be considered.

Materials and Methods

Materials and Methods

Materials

1 Chemical Products and Materials

Chemicals	
Agarose-A-beads	Sigma-Aldrich
Ampicilline	Carl Roth
Aspartate	Sigma-Aldrich
Bovine Serum Albumin	Sigma-Aldrich
Digitionin	Sigma-Aldrich
DMP	Sigma-Aldrich
DTE	Sigma-Aldrich
Fetal Calf servum	Eurobio
Imidazol	Sigma-Aldrich
IPTG	Euromedex
Kanamycin	Sigma-Aldrich
Mannitol	Roth
Percol	Sigma
PerfectHyb	Sigma-Aldrich
Phosphat-buffered Saline	Gibco
Rotiphorese Gel 30/40	Carl Roth
Sucrose	Sigm-Aldrich
TBE	Euromedex
TCA	Carl Roth
TEMED	Carl Roth
Tri-Reagent	Sigma-Aldrich
Tris-base	Sigma-Aldrich
Triton X-100	Carl Roth
Tween-20	Sigma-Aldrich

Cell Culture	
100 mm Cultur dishes	Corning
6-well/96-well plates	BD Falcon
Blasticydin	Invitrogen
Dioxycyclin	Sigma-Aldrich
Dulbecco's Modified Eagle's Medium	Gibco
Fetal Calf Serum	Eurobio
Glasgow-MEM	Gibco
Penecillin Streptomycin	Gibco

Puromycin	Invitrogen
SFMII 293 Medium	Invitrogen
T75/T175 Flask	BD Falcon
Trypsin EDTA	Gibco
Tryptose Phosphat Broth	Gibco
Uridine	Sigma-Aldrich

ECL-Solutions	
Lumi-Light^{Plus} Western Blot Substrat	Roche
Pierce ECL Western Blotting	Pierce Scientific
SuperSignal West Femto Maximum Sensity Substrat	Pierce Scientific
SuperSignal West Pico	Pierce Scientific

DNA Plasmid Preparation Kits	
NucleoBond Xtra Maxi Plus	Macherey-Nagel
NucleoSpin Gel and PCR Clean up	Macherey-Nagel
NucleoSpin Plasmid	Macherey-Nagel

2 Enzymes

BstN1	New England Biolabs
NcoI	New England Biolabs
NdeI	New England Biolabs
NsiI	New England Biolabs
DpnI	New England Biolabs
SAP (Shrimp Alkaline Phosphatase)	Fermentas
T4 PNK (Polynucleotide Kinase)	Fermentas
T4 DNA Ligase	Invitrogen
Phusion High Fidelity DNA Polymerase	Fermentas
Trypsin EDTA	Gibco

3 Biological Material

BHK21	(ATCC # CRL-12072)
HEK293T	Invitrogen, Carlsbad, CA. USA
HEK293F	Invitrogen, Carlsbad, CA. USA
Hepa1-2	gift from M. Frugier, Strasbourg, France
Myoblast	MYOSIX, France

4 Primers

Primers were purchased as lyophilized, desalted oligonucleotids from Sigma Aldrich. Upon arrival primers were resuspended to a final concentration of 100µM and stored as stock solution at -20°C. Working solution was prepared as 10µM concentration.

Cloning	
P1	CCAATCGATTTCCCTTCTTGGTTAAGTCAGC
P2	CCAATGCATATGAGCTCTTTCTGCTTTGGAG
P3	CCAATGCATCCAGAATTCAGTAGCTTTGTTGTC
P14	CAGAATTCAGTAGCTTTGTTGTCCGGACCAACACATG
P15	CAAAGCTACTGAATTCTGCATGGCGGGCCGC
P20	ATTCCAGAATTCAGTAGCTTTGTTG
P21	GAATCATGGCGGCCGCGGTG
P26	GAAGGAGGTTACCATATGCATCCCCTGCTCGGACTTC
P27	GTCATCTTTGTAGTCCCATGGTCAATGAGCTCTTTCTGCTTTG
P28	GTCCTCTCCATACTTCTC
P40	CAATGTTGATTCAACATATGGCTGCGTCCATGGC
P41	GGGATTAACATGCATTTAGCTTACAGCAGGCTGG
P42	CAATGTTGATTCAACATATGGCCCTGTATCAGAGG
P43	TATGGGATTAACCCATGGGAAAATTTCTTCGGCATTTGG
P44	CAATGTTGATTCAAATGCATCTGGGGGTCCGCTG
P45	TATGGGATTAACCCATGGTAAAAGGCATGAATGAGGAAAC
P46	CAATGTTGATTCAACATATGGAAGGGCTGCTGACAAG
P47	TATGGGATTAACATGCATTTAGACAGTCTGCACTGGG
P48	CAATGTTGATTCAACATATGGCGCTGCACTCAATG
P49	TATGGGATTAACCCATGGTAGAAAACCCACCAATTTCTTC
P50	CAATGTTGATTCAACATATGCCCTCTCCGCGTCC
P51	TATGGGATTAACCCATGGTTCCTCGATTGTCTCTTTTTTAC

Northern Blot Hybridization	
NP1	GTAAGATATATAGGATTTAG
NP2	GCTAGACCAATGGGACTTAAA
NP3	TCACTGTAAAGAGGTGTTGG
NP5	TGGAAGCCCCGGCAGGTTTG
NP6	TATTCTCGCACGGACTACAA
NP7	TGTTAAGAAGAGGAATTGAA

Mutagenesis PCR	
P6	GATGTTATAGCGATGAAGGTTCAAGACCAGACAGAC
P7	CCTTCATCGCTATAACATCGGGCAACCTGAAAATATC
P22	CCAATGCATCCCCTGCTCGGACTTCTTC
P23	CCATGGTCAAGACTCAGACAGTCGC
P24	GAAGGAGGTTACCATATGCATTACTTCCCTTCTTGGTTAAG
P25	GTCATCTTTGTAGTCCCATGGTCAAGACTCAGACAGTCGCTTC
MP1	CCAGAATTCGGTAGCTTTG
MP2	CAAAGCTACCGAATTCTGG
MP3	GAGAGTTGGGTTCGTCTCAC
MP4	GTGAGACGAACCCAACTCTC
MP5	CCAAATGAAGTATAACCTGC
MP6	GCAGGTTATACTTCATTTGG
MP7	GGTTGCCCAATGTTATCGAG
MP8	CTCGATAACATTGGGCAACC

RT-PCR detection	
P4	GTATCTGGGACAGTAATCTCC
P5	TTAAGAGCTCAAGTGTTTGGAAG
P8	TGTAATCTGCACGGGTTTGTG
P9	TGTCTAACCCTAAGGCAATTC
P10	CTAAGGCCAACCGTGAAAAG
P11	TCTCAGCTGTGGTGGTGAAG
P13	CTCAGATTGACATAGAGATGTC
P16	GTGGCTTCTGGAGAGACTGCAGATGTA
P17	GGGCACAGGAAGGTTGGCCATCTCTCT
P18	TATAGGCTAAATCCTATATAT
P19	AGGTCGCCTGGTTCTAG
P54	ATCTGCACGGGTTTGTGGAT
P55	CTGTGCATCATGGATTCGG

5 List of used antibodies

Antibodies were received as liquid solution and stored after recommendation of the supplier. If necessary, antibodies were stored in aliquots at -20°C.

Table 10: List of all used commercial antibodies.

Primary Antibodies								
Name	**Antibody**	**Antigene**	**Specificity**	**Clonality**	**Source**	**Conc. (mg/ml)**	**Dillution**	**Reference**
@AARS2	Anti mitochondrial Alanyl-tRNA synthetase	Center 373-402 aa	human	Polyclonal	Rabbit	0.25	1/100	Abgent-AP7559c
@CARS2	Anti mitochondrial Cysteinyl-tRNA synthetase	C-terminal 535-564	human	Polyclonal	Rabbit	0.25	1/100	Abgent-AP7846b
@CytoC	Anti mitochondrial Cytochrome C	nD	human mouse rat	Monoclonal	Rabbit		1/20000	Abcam-ab133504
@DARS2	Anti mitochondrial Aspartyl-tRNA synthetase	N-terminal 1-100 aa	human	Monoclonal	Mouse	0.05	1/500	Novus Biologicals H00055157-B01P
@DAR 2	Anti mitochondrial Aspartyl-tRNA synthetase	Full length	human	Polyclonal	Mouse	1	1/500	Abcam-ab69336
@FARS2	Anti mitochondrial Phenylalanyl-tRNA syntetase	Full length	human	Polyclonal	Mouse		1/1000	Abcam-ab68324
@FLAG	ANTI-FLAG® M2-Peroxidase (HRP)	FLAG epitope		Monoclonal	Mouse	1	1/800	Sigma-A8592
@Flag2	Anti-FLAG M2	FLAG epitope		Monoclonal	Mouse	2	1/1000	Agilent-200472
@GARS	Anti Glycyl-tRNA synthetase	C-terminal	human	Polyclonal	Rabbit	0.25	1/500	Abgent-AP7952b
@HARS2	Anti mitochondrial Histidyl-tRNA synthetase	C-terminal	human	Polyclonal	Rabbit	0.1	1/100	Santa Cruz Biotechnology sc-130586
@His	Anti-6x-His Epitope Tag- HRP conjugated	HHHHHH		Polyclonal	Rabbit		1/5000	Thermo Scientific PA1-30578
@IARS2	Anti mitochondrial Isoleucyl-tRNA synthetase	Center 607-636 aa	human	Polyclonal	Rabbit	0.25	1/100	Abgent-AP7840c
@KARS	Anti Lysyl-tRNA synthetase	Full length	human mouse	Polyclonal	Rabbit		1/8000	Abcam-ab31532
@LARS2	Anti mitochondrial Leucyl-tRNA synthetase	N-terminal 1-174 aa	human	Polyclonal	Rabbit	1	1/3000	Abcam-ab96221
@MARS2	Anti mitochondrial Methionyl-tRNA synthetase	C-terminal 564-593 aa	human	Polyclonal	Rabbit	0.25	1/100	Abgent-AP7871b
@MRLP39	Anti mitochondrial ribosomal protein L39	internal	human mouse	Polyclonal	Rabbit	1	1/500	Abcam-ab74107
@NARS2	Anti mitochondrial Asparaginyl-tRNA synthetase	C-terminal 448-477 aa	human	Polyclonal	Rabbit	0.25	1/100	Abgent-AP7585b
@PARS2	Anti mitochondrial Prolyl-tRNA synthetase	C-terminal 397-424	human	Polyclonal	Rabbit	0.25	1/100	Abgent-AP11090b
@Prohibitin	Anti mitochondrial Prohibitin	100-200 aa	human mouse rat cow dog	polyclonal	Rabbit		1/8000	Abcam-ab28172
@RARS2	Anti mitochondrial Arginyl-tRNA synthetase	C-terminal 377-406	human	Polyclonal	Rabbit	0.25	1/100	Abgent-AP8590b
@SARS2	Anti mitochondrial Seryl-tRNA Synthetase	Full length	human	Polyclonal	Mouse		1/1000	Abcam-ab69316
@TARS2	Anti mitochondrial Threonyl-tRNA synthetase	C-terminal 694-718	human	Polyclonal	Rabbit	0.25	1/100	Abgent-AP11450b
@VARS2	Anti mitochondrial Valyl-tRNA synthetase	N-terminal 46-270 aa	human	Polyclonal	Rabbit	0.96	1/5000	Abcam-ab155152
@VDAC	Anti mitochondrial voltage dependent anion channel, HRP conjugated	150-250aa	human mouse rat	Polyclonal	Rabbit	1	1/5000	Abcam-ab40747
@WARS2	Anti mitochondrial Tyrphanyl-tRNA synthetase	C-terminal 294-324	human	Polyclonal	Rabbit	0.25	1/500	Abgent-AP7586b
@YARS2	Anti mitochondrial Tyrosyl-tRNA synthetase	N-terminal	human	Polyclonal	Rabbit	0.25	1/100	Abcam-ab71419
Secondary Antibodies								
Name	**Antibody**	**Antigene**	**Specificity**	**Clonality**	**Source**	**Conc. (mg/ml)**	**Dillution**	**Reference**
@ms	Anti-Mouse IgG (whole molecule)–Peroxidase antibody	IgG mouse	mouse	Polyclonal	rabbit		1/20000	Sigma-A9044
@rab	Anti-Rabbit IgG (H+L)-HRP Conjugate	IgG (H+L) rabbit	rabbit	Polyclonal	goat		1/2000	Biorad-172-1019

Methods

1 Mammalian cell culture

1.1 Human Cell lines

1.1.1 HEK293T

Human Embryonic Kidney cells 293T (HEK293T, Invitrogen) were chosen because of their high content of mitochondria, their easy handling and ability to grow in large quantities. Cells were maintained in DMEM (Dulbecco's modified Eagle's medium, Invitrogen) supplemented with 10% FCS (fetal calf serum), 100 units/ml penicillin and 100µg/ml streptomycin at 37°C, in a 5% CO_2 humidified incubator. For cell splitting, cells were briefly washed with PBS (130mM NaCl, 2.6mM KCl, 10mM Na_2HPO_4 1.7 mM KH_2PO_4; pH=7.5) and trypsinized for two minutes. Cells were sub cultivated in a dilution of 1:4 to 1:8. For cryopreservation, cells were frozen in a dilution of 1:8 in growth medium containing 10% DMSO.

1.1.2 HEK293F

HEK293F (Invitrogen) is adapted for serum free media and growing conditions in suspension. This cell line was chosen for its fast growing and high-density growth properties. Cells are capable to grow in suspension or as adherent cells. For adherent culture, cells were cultivated in DMEM supplemented with 10% FCS, 100 units/ml penicillin and 100µg/ml streptomycin at 37°C, in a 5% CO_2 humidified incubator. For cell splitting, cells were resuspended in adequat amount of media and sub cultivated in a dilution of 1:10. For cryopreservation, cells were frozen in a dilution of 1:10 in growth medium with 10% DMSO.

For suspension culture, cells were grown in suspension under constant rotation (10 rpm) in a miniPERM bioreactor (Greiner Bio-one) using 293 SFM II medium (Gibeco®) with 1% Penicillin and Streptomycin at 37°C and 5% CO_2. Cells were harvested at a density of 10^7cells/ml. Cells were sub cultivated in a dilution of 1:4.

1.2 Other cell lines

1.2.1 Hepa 1-6

Hepa 1-6 cells (Hepatoma cells a gift from Magalie Frugier, IBMC, Strasbourg) are mouse cells derived from liver tissue. Cells were maintained in DMEM supplemented with 10% FCS, 100units/ml penicillin and 100µg/ml streptomycin at 37°C, in a 10% CO_2 humidified incubator. For cell splitting, cells were washed with PBS and trypsinized for 10min. Cells were sub cultivated in a dilution of 1:5.

1.2.2 BHK21 (ATCC # CRL-12072)

BHK-21 (Baby Hamster Kidney) cells are fibroblast deriving from kidney. This cell line was used for the viral infection with the modified Vaccinia Ankara Virus to over express recombinant human proteins (Jester et al., 2012). BHK-21 cells were cultivated in complete GMEM medium (Glasgow modification of minimum essential medium, Invitrogen) containing 10% FCS, 1.5g/ml bacto-tryptose phosphate, 100units/ml penicillin and 100µg/ml streptomycin at 37°C. For cell splitting, cells were washed with PBS and trypsinized for 10min. Cells were sub cultivated in a dilution of 1:10 for expression studies and 1:15 -1:20 for biomass production. For cryopreservation, cells were frozen in a dilution of 1:3-1:10 in growth medium containing 10% DMSO.

2 BHK21/Vaccinia Virus expression system

This method, based on the innate ability of the Modified Vaccinia Virus Ankara (MVA-T7) to co-internalize a vector into mammalian cells without using any further costly transfection reagents, was previously described, and named MVA-mediated transfer (MVA-mtr) (Jester et al., 2012). This virus is species-specific restricted to hamster cells (BHK21) and thus requires only safety level 1, facilitating the co-infection of high quantities of cells. The original vector pBCJ739.14 was prepared and described in Jester at al., 2012. The features and properties of this system is described in a method article (Jester et al. 2013, Chapter III). The sequences of interests are cloned downstream of the Bacteriophage T7-promoter, and the LacO binding site for IPTG controlled transcription. The appropriate T7-polymerase is constitutively expressed from the viral genome. Of note, the presence within the "shuttle" vectors of both bacterial and mammalian transcription/translation elements permits the easy shuttling between bacterial and eukaryotic expression systems. The preparation of different gene constructs, virus and protein production is described in detail.

2.1 Vector

The original vector, designed and prepared by Brian Jester, contains the ORF for the full-length mt-AspRS flanked by NsiI digestion sites and C-terminal Flag and Strep epitopes (**Figure 51**). The MVA-*E. coli* shuttle vector that places the Flag-Strep II affinity tag at the C-terminal end of expressed proteins (pBCJ735.20) was previously constructed (Jester et al., 2012). To construct the full-length mt-AspRS expression clone the *DARS2* gene was eluted as a Nsi I fragment from pFL and cloned into pBCJ735.20 resulting in pBCJ749.77. The SOD2 gene was amplified from cDNA that was constructed from RNA extracted from HEK 293 using Superscript II RT (Invitrogen, Carlsbad CA, USA) per manufacturer's protocol. Cloning of the SOD2 gene into the multi-cloning site of pBCJ735.20 resulted in the full-length affinity tagged constructs (pBCJ762.23). For fluorescent imaging studies the N-terminal peptides for the different

genes were fused to GFP. The different gene fragments were amplified from cDNA and ligated into the multi-cloning site of pBCJ739.14.

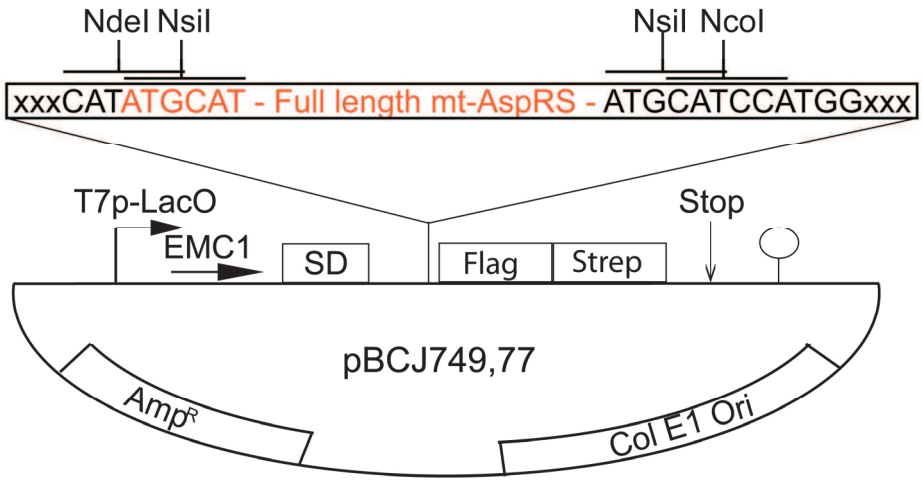


Figure 51: Representation of the expression vector construct and corresponding restriction sites NdeI, NsiI and NcoI flanking the ORF for mt-AspRS.

The N-terminal fragments cloned were: 168 aa of mt-AspRS (pBCJ753.2), and 89 aa of SOD2 (pBCJ747.2). To make a construct that expressed a known mitochondrial targeted protein we replaced the GFP in pBCJ747.2 with mCherry amplified from pRSET-mCherry (Invitrogen, Carlsbad CA, USA) resulting in pBCJ722.1.

2.2 Virus production

For production of virus, BHK21 cells were grown in a T175-flask to a confluence of 80-90%. Cells were washed with PBS and 1ml of "diluted" virus was added with 3ml GMEM supplemented with 1.5g/ml bacto-tryptose phosphate, 100units/ml penicillin, and 100µg/ml streptomycin at 37°C (GMEM without supplements). The cells were incubated for 30min at room temperature under agitation to keep cells moist. Afterwards, 50ml of complete GMEM medium was added and the cells were cultivated for 48h at standard conditions. The infected cells were harvested by scratching and the 50ml supernatant were centrifuge. For virus production ("diluted" virus stock), the cell pellet was resuspended in 50ml GMEM without supplements. For protein expression ("concentrated" virus stock), the cell pellet was resuspended in 2ml

GMEM w/o supplements. The virus was released from the cells by thaw and freeze cycles in a dry ice/ethanol bath and several strokes through a syringe. The lysate was centrifuged at 1000g to clear the virus solution from cell fragments. The supernatant containing the virus was stored at -20°C in GMEM w/o supplements.

2.3 Protein production

The expression of recombinant protein was performed in two formats: as a test expression or as preparative expression scale. In the test expression scale, one 10cm culture plate was used per tested construct. BHK21 cells were grown over night until reaching a confluence of 70%. Cells were washed with PBS. 1 ml of "concentrated" virus mixed with 10µg of plasmid DNA was added to the cells and incubated 30min at room temperature under agitation. Afterwards, 9 ml of complete GMEM medium was added containing 1mM IPTG. The cells were incubated 24-36h at standard growth condition and subsequently harvest. The pellet was washed with PBS. For preparative expression scale 10-15 x150mm culture plates were prepared with BHK21 cells at a confluence of 70%. Cells were washed with PBS and 3ml "concentrated" virus mixed with 30µg plasmid DNA were added. After incubation of 30min at room temperature, 17ml of complete GMEM medium was added containing 1mM IPTG. Cells were harvested after 36h of incubation under standard growth conditions. Cells were washed and pelleted for downstream investigations.

3 In vivo knock down of mRNA

3.1 Stable cell lines expressing shRNA against specific mRNAs

Cells containing stably expressed and inducible shRNA were constructed from Lucia Echevarria (Madrid, Spain). Briefly, HEK293T cells were stably transfected with pSUPERIOR vector facilitating the inducible shRNA expression. The vector has *E.coli* Tn10-encoded Tetracylcin resistance operon, which is under the control of the Tet

repressor (TetR). The TetR gene is constitutively expressed from pcDNA6/TR© regulatory vector, which is in addition stably introduced into HEK293T. Tetracyclin or Deoxycycline bind to the TetR and facilitate the expression of the gene of interest. Stably transfected HEK293T cells were maintained in DMEM medium supplemented with 10% FCS, 1% Penicillin-Streptomycin, 1.5µg/mL of Puromycin, 5µg/mL of Blasticydin, 50µg/mL of Uridine in a humidified atmosphere containing 5% CO_2 at 37°C. Expression of shRNA was induced with 2µg/ml Doxycycline.

3.2 Cell Proliferation assay

The effect of the knock down were quantified using the CellTiter96®AQueous One Solution Cell Proliferation assay (Promega). Briefly, the principle of this test is the bioreduction of a tetrazolium compound to a colored formazan product in cell culture. This reaction depends on NADPH or NADH existing in metabolic active cells. The color change can be quantified at 490nm and directly correlates with the number of living cells.

For each cell line, 2 x 10 cm culture plates were prepared and grown over night. The next day, one plate of each cell line was induced with 2µg/ml Doxycycline and incubated for 24h. 2000 cells in 100µl culture volume were seeded as quintuplet on 96-well plate format. For each time point, one separate 96-well plate was prepared. In addition, 100,000 cells were seeded on 6-well plates. The cells were grown for 24h before the first measurement was performed. For measurement, 20µl of CellTiter96®AQueous One Solution reagent was added and cells were incubated for 1h at 37°C. The absorbance was detected with a bench top spectrometer. The cell proliferation was analyzed up to 7 days after RNAi induction.

4 Sub-cellular fractionation

4.1 Preparation of mitochondria

Mitochondria were purified from adherent cells or cells cultured in suspension, using differential and gradient centrifugation steps. The purification and fractionation

steps have to be performed on a single day. Cells were harvested and washed two times in PBS. Afterwards, cells were resuspended in 10ml Buffer A (220mM Mannitol, 70mM Sucrose, 10mM Hepes-KOH; pH= 7.5, 1mM $MgCl_2$, 1mM EDTA and 1% BSA) and ruptured using a warring blender (5x10sec). The lysate was centrifuged for 10min at 800g to remove unbroken cells. The supernatant was centrifuged for 30min at 8600g. The obtained pellet was resuspended in Buffer A and centrifuged sequentially at 800g and 8600g. The final pellet was resuspended in gradient buffer (600mM Sucrose, 20mM Hepes-KOH; pH=7.4, 10mM EDTA and 2% ETOH;). Percol gradient was prepared by loading two phases on a centrifugation tube (Ultra Clear^{TM} tubes, 14x89mm, Beckman Coulter). The bottom phase contain 2ml 60% sucrose in gradient buffer and the top layer contains a mix of 50% Percol in gradient buffer. 500 µl of lysate were loaded on this gradient. The gradient was centrifuge 45min at 100,000xg in a swing out rotor. Purified mitochondria could be recovered in a concentrated band on the top 1/3rd of the gradient (**Figure 52**).

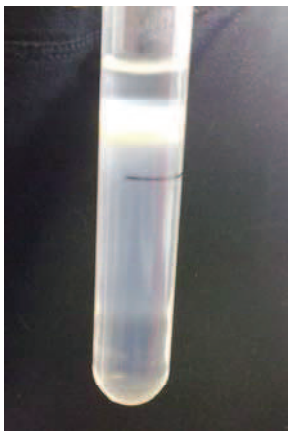


Figure 52: Percol-Sucrose gradient. Mitochondrial fraction appears as a yellow ring in top area of the gradient.

Recovered mitochondria were aliquoted in 2 ml tubes and diluted three times with 0.5x gradient buffer. Diluted mitochondria were pelleted at 8600g for 30min and three times washed in wash buffer (0.3M Mannitol, 10mM K_2HPO_4/KH_2PO_4 buffer; pH= 7.5, 1mM EDTA). Finally mitochondria were pelleted, weighted and ready for further applications.

4.1 Preparation of Mitoplast

Mitoplast were prepared following two different methods (osmotic shock and digitonin treatment). For the first method, purified mitochondria were resuspended in a volume of 100µl of 10mM Tris-HCl (pH=7,4 at 25°C) per 6mg of wet mitochondrial pellet and incubated for 20min on ice. Mitoplasts were recovered by centrifugation at 8600g for 10min and respuspended in wash buffer (0.3M Mannitol, 10mM K_2HPO_4/KH_2PO_4 pH=7.5, 1 mM EDTA).

For the second method, purified mitochondria were resuspended in wash buffer (0.3M Mannitol, 10mM K_2HPO_4/KH_2PO_4; pH=7.5, 1mM EDTA) and treated with 100µg digitonin per mg of mitochondrial protein. The final concentration of digitonin was adjusted to 0.05%. The Mixture was incubated for 20min at room temperature, 2-3 times diluted in wash buffer and the mitoplast was recovered by centrifugation at 8600g for 10min and washed twice with wash buffer. Freshly prepared mitoplasts were subjected to mitochondrial fractionation protocol.

4.2 Sub-mitochondrial fractionation

If not further specified, mitochondria or mitoplasts were fractionated as following. Mitochondria were resuspended in wash buffer (10µl/mg wet pellet) containing 1x Protease inhibitor cocktail (Protease Arrest, G-Bioscience). Mitochondria were ruptured by three freeze and thaw cycles using dry ice/ethanol bath. Subsequently, mitochondria were sonnicated three times for 10s (6W) and centrifuge at 8600g to remove unbroken mitochondria. The lysate was then centrifuge at 100,000g in 1.5ml test tubes to separate the matrix fraction (supernatant) from the total membrane fraction (pellet). This pellet fraction was further treated with 0.1M Na_2CO_3 (pH=11) under agitation at room temperature for 45min. The loosely anchored proteins (peripheral) in supernatant were separated by ultracentrifugation at 100,000g for 45min from the deeply anchored (integral) membrane protein fraction (pellet). If necessary the fraction (matrix and peripheral proteins) were concentrated to the same final volume (80µl) using the Centricon concentration device (exclusion size: 10kDa, Millipore).

5 Enrichment of proteins by immuno-precipitation

5.1 Production of antibodies in rabbits

Immuno-precipitation is a technique, which uses the antibody:antigen interaction to specifically pull down a protein of interest. We used this technique to purify either mt-AspRS alone or in complex with possible partners. To do so, an antibody specifically targeted against the full-length 44-645 amino acid long recombinant protein was produced in rabbits (IGBMC, M.Duval). Beforehand, the purity of the recombinant purified protein in PBS was controlled by SDS-PAGE and DLS analysis (performed from Bernard Lorber, IBMC, Strasbourg). Two injection of the Antigen (1mg/ml) were performed. The second injection was performed after 2 month of the first one. The collection of serum started after one month after the first injection. The serum was collected for eight weeks.

5.2 Preparation of covalently-linked antibodies:

Since the heavy chain of the antibody and the mt-AspRS have a similar MW and thus have the same migration on a SDS-PAGE, the antibodies were covalently linked to Agarose-A beads (Sigma-Aldrich). The corresponding protocol has been optimized by Marie Messmer (IBMC, Strasbourg, France). Briefly, 100mg beads were incubated in 3ml PBS for 15min at room temperature. The beads were washed twice with 3ml PBS and aliquoted into 2ml test tubes. Afterwards, 100µl of serum (diluted in 1.4ml PBS) were added to the beads and incubated for 2h at 4°C under continuous rotation. The excess of antibodies was removed by two washing steps with 0.2M sodium borate (pH=9). The cross-linking reaction was performed in 2ml sodium borate (0,2M) containing 5mg/ml DMP. The pH was adjusted to 9 with concentrated NaOH. After 30min of incubation at RT, the cross-linking buffer was renewed and the reaction was carried on for an additional 30min. The reaction was stopped by addition of 1ml of 0.2M ethanolamine and incubation for 2h. To remove non-cross linked antibodies, the beads were subsequently transferred to 10ml centrifuge columns (Thermo Scientific) and washed with 5 ml PBS and 5 ml Glycine (0.1 M; pH=3). The beads were equilibrated with 10 ml PBS and stored overnight at 4°C.

5.3 Preparation of protein extracts

5.3.1 Formaldehyde cross linking

Preliminary to the investigation of possible cellular partners of mt-AspRS, a formaldehyde cross-linking experiment was performed. Formaldehyde is a small compound (HCHO), which reacts with primary amino groups of proteins and cross-link two protein via a methylene-bridge $–CH_2-$. To identify potential macromolecular complexes containing mt-AspRS, HEK293T cells were spread out in a 6-well plate format and grown over night. Cells were washed twice with PBS and incubated with different concentrations of formaldehyde at room temperature for up to 1h. Afterwards, cells were washed twice with PBS and harvested. Cells were lysed in a method compatible with BN-Gel electrophoresis. Alternatively, purified mitochondria were resuspended in 1% formaldehyde and incubated for up to 1h at ice. Afterwards, mitochondria were washed twice with wash buffer (0.3M Mannitol, 10 mM K_2HPO_4/KH_2PO_4; pH=7.5, 1mM EDTA) and subjected to submitochondrial fractionation.

5.3.2 Preparation of lysate for mt-AspRS enrichment

Starting material, cells or mitochondria, were washed once in lysis buffer 1 (20mM Hepes; pH= 7.5, 200mM KAc, 0,2% Triton,). The samples were resuspended in lysis buffer 1 (0.1 ml/100 cm cell culture plate) and sonnicated three times for 15s (6W). Afterwards, two volumes of lysis buffer 2 (20mM Hepes; pH= 7.5, 200mM KAc) were added to dilute the Triton to a final dilution of 0.06%. The lysate was centrifuged for 10min at 800g and the supernatant was further subjected to an ultracentrifugation step (100,000g for 45min at 4°C). The supernatant was collected and ready for further applications.

5.3.3 Preparation of lysate under native conditions for co-immuno precipitation

Starting material, cells or mitochondria, were washed once in PBS and resuspended in Pierce IP-Buffer following the product instruction (Pierce Scientific). To keep possible interactions of proteins, all subsequent steps were performed on ice and at 4°C. After 10min of incubation the samples were subjected to an ultracentrifugation step (100,000xg for 45min at 4°C). The supernatant was collected and ready for further applications.

5.4 Immuno-precipitation

The lysate was pre-cleared by an incubation with 100mg of equilibrated Agarose-A beads under continuous rotation at 4°C for 1h. After removal of the Agarose-A beads, this pre-cleared lysate was incubated with 100mg of Agarose-A beads conjugated with unspecific antibodies of the pre-immune serum. After removal of the Agarose-A beads, this lysate was incubated with 200mg of Agarose-A beads conjugated with mt-AspRS antibodies for 2h at 4°C under continuous agitation. Afterwards, the beads were transferred to a 10ml centrifugation column and the lysate flew through the column. The beads with the bound mt-AspRS were incubated with 10ml of PBS for 30min at 4°C. The PBS was removed and the beads washed a second time with 5ml PBS. The beads were transferred in a 1.5ml test tube and incubated with 80-160µl elution buffer (250mM Tris, pH=7,5 at 25°C + 2% SDS) at 50°C for 5min. The supernatant was recovered and the elution step was repeated three times. In order to improve the elution, the beads were further washed with 0.1 M Glycine (pH=3) or directly incubated at 95°C with 80µl denaturating loading dye. All elutions were pooled and diluted so that the SDS concentration fall down to <1% SDS to ensure a full compatibility with the subsequent concentration step in Centricon devices (exclusion size: 10kDa).

6 DNA manipulation

6.1 Preparation of cDNA from total RNA

Total RNA was reverse transcribed with First Strand cDNA synthesis kit (Fermentas). To do so, 1-2µg of total RNA was pre-incubate with 0.1mM Oligo$(dT)_{18}$ at 65°C for 5min. Afterwards, 4µl of 5x reaction buffer, 1.25mM dNTPs and 1U M-MuLV Reverse Transcriptase and H_20 were added to final volume of 20µl. The reaction was performed at 37°C over 1h and stopped at 70°C for 15 min. The synthesized cDNA was stored at -20°C for downstream application.

6.2 Amplification of mt-aaRS sequences and cloning

Mt-aaRS sequences were amplified from the cDNA pool with gene specific primers. The 50µl reaction mixture contained 2µl cDNA, 0.5mM dNTPs, 1µM forward and reverse primer, 1xPhusion HF buffer and 1U Phusion High-Fidelity DNA Polymerase (Thermo Scientific). A standard PCR program was the following:

Denaturation	98°C	3 min
Denaturation	98°C	30 s
Annealing	T_m-2°C	30 s
Extension	72°C	4 min
Final extension	72°C	5 min

30X

PCR products were separated on 1% agarose gel and corresponding bands were cut out. The DNA was recovered from the gel using the NucleoSpin Gel purifcation kit (Macherey and Nagel) and eluted in 17µl H_20. The DNA concentration was determined and approximately 50-100ng of PCR product was ligated in 25ng PCR Blunt a vector (Invitrogen) in final volume of 10µl for 10 min at room temperature. 5 µl of ligation product was transferred into 50µl of chemical competent cells (OneShot TOP10 cells, Invitrogen) and incubated 30min at ice. After 45s of heat shock at 42°C, the transformation mixture was transferred into 250µl SOC medium and incubated for an

additional 60min at 37°C. The transformation mixture was spread out on LB-agar plates containing kanamycin (50µg/ml) as selection marker and incubated at 37°C over night.

6.3 Site directed mutagenesis

Plasmid (pBCJ749.77) containing the desired gene expression cassette was used as a template for mutagenesis PCR. Primers containing the nucleotide exchange of interest were designed so that the introduced mismatch was flanked by approximately 8-10 nucleotids. The PCR mixture contained 50-100ng template DNA, 1xPhusion HF buffer and 1U Phusion High-Fidelity DNA Polymerase, 1µM forward and reverse primer and 0.25 mM dNTPs(each) in a final volume of 50µl. The PCR was performed using following program:

Denaturation	98°C	10 min	
Denaturation	98°C	30 s	30X
Annealing	T_m-2°C	30 s	
Extension	72°C	1min 30s	
Final extension	72°C	10 min	

After completion of the PCR program, 1µl of DpN1 was added to the reaction mixture, followed by 2h incubation at 37°C. As a negative control the same amount of template DNA was digested. The reaction mixture was then used to transform chemical competent cells (OneShot TOP10 cells, Invitrogen) as described before. Positive colonies were selected through their Ampicillin resistance. Bacterial colonies were cultured and plasmids were extracted for DNA sequencing.

7 RNA manipulation

7.1 RNA Isolation from human cell lines

Overnight grown cells were washed with PBS and 3ml Tri-Reagent was added per T75 flask or 100 mm plate. Cells were incubated for 5min at room temperature and the whole lysate was splitted in 1ml aliquots and 0.2ml Chloroform was added. The mixture was incubated for 10min at room temperature. The organic phase was separated

from the aqueous phase by a centrifugation step at 13,000g over 10min. The total RNA was precipitated form the aqueous phase using 0.5ml isopropanol. After 10min incubation, the precipitate was centrifuged at 13,000g for 15min. The obtained pellet was washed with 80% EtOH and resuspended in an adequate amount of RNase free H_20. Concentration was determined using the NanoDrop measuring the optical density at 260nm.

7.2 Extraction of mRNA from purified polysomes

HEK293 cells were grown on 30x150 mm dishes. At 75% confluence, cells were harvested, washed three times with PBS, resuspended in buffer (10mM Hepes; pH= 7.5, 10mM KAc, 0.5mM $MgAc_2$, 5mM DTT), and lysed by 20 strokes with a syringe. After centrifugation at 13,000g, the cleared supernatant was concentrated in a Centricon concentration device (exclusion size: 10kDa). 100µl of concentrated sample were mixed with 7% Sucrose in a buffer containing 50mM KCl, 5mM $MgCl_2$, 25mM Tris-HCl; pH=7.5 at 25°C, 1mM DTE. The polysomal fractions were separated from 80S ribosomes and free RNA on a 7%–47% linear sucrose gradient (25mM Tris-HCl; pH 7.4 at 25°C, 50mM KCl, 5mM $MgCl_2$, 1mM DTE, 1mM ProteaseArrest) by centrifugation in a SW41 rotor (Beckman coulter) for 2.5h at 37,000 rpm at 4°C. The gradient was fractionated in 400µl aliquots and RNAs were extracted by phenol/chloroform. RNA was precipitated with EtOH, 0,3M Sodium acetate and incubated for 1h at -20°C. After centrifugation at 13000g, the pellet was washed with 70% EtOH. The dried pellet was resuspended in H_20 and RNA concentration was determined using the NanoDrop. The RNA was then ready for cDNA synthesis or other downstream experiments.

7.3 Northern Blot analysis

7.3.1 Gel electrophoresis and transferof RNAs on Zeta-Probe Blotting Membrane

For the detection of tRNAs, the total amount of extracted RNA was loaded on a 12% polyacrylamid gel containing 8M urea in a Tris-borate-EDTA buffer (TBE-buffer).

The samples were prepared in RNA loading dye (40% Formamid, 12.5mM EDTA, 0.02% xylene cyanol and 0.02% bromo phenol blue). The electrophoresis was performed in 1x TBE running buffer at 150V and room temperature. The migration was stopped when the Xylen-Cyanol blue reached the bottom of the gel. The transfer on the Zeta-Probe Blotting Membrane was performed using the Mini Trans-Blot® Electrophoretic Transfer Cell (BioRad). The transfer sandwich was prepared as following: The gel was placed in a sandwich cassette with the gray side of the cassette on the bottom. The sandwich contained, from bottom to top, following components: pre-equilibrated fiber pad, pre-equilibrated filter paper, gel, pre-equilibrated membrane, pre-equilibrated filter paper, pre-equilibrated fiber pad. The sandwich cassette was placed in the electrode module, which then was placed in the buffer tank containing 1xTBE buffer and a frozen cooling unit. The transfer was performed at constant 80V at 4°C under continuous stirring for 1h. After the transfer the RNA was UV-cross linked to the membrane with a UV-Stratalinker® 1800 (Stratagene) using the autocross linking program. The membrane was then stored between wet filter pads or directly used for further experiments.

7.3.2 Gel electrophoresis of RNA under acidic conditions

For the detection of aminacylated-tRNAs, the aboved-described method had to be modified. All steps were performed under strict acidic conditions and on ice. The RNAs were isolated under standard condition with the modification that NaOAc (pH=4.5) replaced all aqueous components. The samples were prepared in acidic loading dye (0.1M NaOAc;pH=4,5), 8mM Urea, 0.05% xylen-cyanol and 0.05% bromo-phenol blue). The electrophoresis was performed using a 40x30cm polyacrylamid gel (6.5%) containing 7M urea and 0.1M NaOAc (pH=4.5). A brief migration at 1000V for 30min was followed by a migration a 500V over 20h at 4°C in 0.1M NaOAc running buffer. The migration was stopped when the bromo-phenol migration front reached the bottom. The migration area of tRNAs was estimated using a co-migrated *in vitro* transcribed tRNA, which was detected by UV-shadowing. The corresponding area was sliced out using a scalpel. The gel was briefly equilibrated in TBE buffer before the transfer sandwich was prepared. The transfer was performed using the Mini Trans-Blot®

Electrophoretic Transfer Cell (BioRad) as described before.

7.3.3 Probe hybridization and detection

Oligonucleotides complementary to the 3'-end of the tRNA of interest were 5' labeled with [P^{32}]-nucleotided. The labeling reaction (10pmol oligonucleotides, 2µl T4-PNK, 5µl PNK-buffer, 50µCi ATP[^{32}P] in final volume of 50µl) was performed at 37°C for 45min and stopped at 60°C for 15min. The radio-labeled probe was purified on a Sephadex G25 spin column and directly used for the subsequent hybridization of the a membrane pre-hybridized with PerfectHyb^{TM} Plus Hybridization buffer (Sigma-Aldrich) at 1h at 42°C. The hybridization was carried out at 42°C over night in a final volume of 8ml PerfectHyb^{TM} Plus Hybridization buffer. The membrane was washed twice for 15min with 5xSSC (0.75M NaCl and 75mM Trisodium Citrate) containing 0.1% SDS at 37°C. Moist membranes was enclosed into plastic wrap and exposed for 5h to Phosphoimager plate (Fuji). The signal was read out on a Phosphoimager (Fuji Film FLA-5100) and analyzed using the Image Gauge software package.

8 Aminoacylation of *in vitro* transcribed tRNAs

8.1 *In vitro* transcription of tRNAs

Aminoacylation of tRNAs can be analysed using *in vitro* transcribed tRNAs and the corresponding purified recombinant aminoacyl-tRNA synthetase. The $tRNA^{Asp}$ gene was previously clone downstream of a T7-promotor and a hammerhead ribozyme. This construct has a high active T7-promotor site allowing for a maximum transcription rate. Transcription of this construct gave rise to a "transzyme" molecule, the autocatalytic activity of which liberates a 5'-OH tRNA transcript starting with the proper nucleotide (Fechter at al.,1998). The 3`end of the tRNA is generated by a run off transcription of the BstN1 digested DNA template. This construct was available as a plasmid in the laboratory transformed into *E. coli* DHα5 cells. The plasmid was purified using the Nucleobond Maxi Plasmid preparation kit from Macherey-Nagel.

In vitro transcription was carried out in 500µl reaction mixture containing 50µg BstN1-digested plasmid, 40mM Tris-HCl (pH= 8.1), 30mM $MgCl_2$, 1mM Spermidine, 5mM DTE, 0.01% Trition X-100, 8.3mM NTPs (each), 10.4mM GMP, and 7µl of T7-RNA

Polymerase (purified as described in Becker et al., 1996). The reaction was stopped after 4.5h at 37°C by Phenol extraction and the RNA was recovered by Ethanol precipitation. Of note, the T7-RNA polymerase was shown to add one or two non-templated 3'-nucleotides. Right size products had to be separated from unwanted products on a denaturating preparative gel (12% Acrylamid; 8M Urea; 1xTBE, 600V). The tRNA was recovered from the gel by electro-elution (2x 1 h at 150 V in 1xTBE) and a subsequent EtOH precipitation step. The tRNA was recovered in H_20 and the concentrations were determined.

8.2 Aminoacylation assay using in vitro transcribed tRNAs

8.2.1 Determination of tRNA charging level

A convenient method to compare two tRNAs is to establish their charging levels. The maximum charging level (plateau level) is determined under saturated concentration of aaRSs, and represents equilibrium between aminoacylated and uncharged tRNAs.

Before aminoacylation reaction, tRNAs were denaturated (1.5min at 60°C) and renaturated at room temperature. The aminoacylation assays were performed at 25°C with radiolabeled [^{3}H]-aspartate (31µM), 20pmol tRNA and 5µM mt-AspRS. The reaction buffer contained 50mM Hepes-KOH (pH 7.6), 25mM KCl, 12mM $MgCl_2$; 2.5mM ATP, 0.2mg/ml BSA and 1mM Spermine in a final volume of 50µl. Adding the enzyme started the reaction. Aliquots of 10µl were taken at 5, 10, 20 and 30min and the reaction was stopped by spotting the aliquots on pieces of whatmann papers saturated with 5% trichloroacetic acid (TCA). Uncharged and charged tRNAs precipitated on the whatmann paper under acidic conditions, while the free radio-labeled amino acids were removed by the three washing steps in 5% TCA The activity was measured in fluid scintillizer at 1min per sample.

8.2.2 Determination of kinetic parameters of tRNA charging:

The kinetic constants k_{cat} and K_M were derived from Lineweaver-Burk plots. The variation of the tRNA concentration allows the determination of the kinetic parameters

k_{cat}, K_M and k_{cat}/K_M of the aminoacylation of the tRNA. The charging level of the tRNAs were experimentally determined at several time points and blotted. The slope (reaction velocity V) was calculated for different tRNA concentrations [S] and using limiting aaRSs concentration. Based on the Lineweaver Burk plot:

$$\frac{1}{V} = \frac{K_M}{V_{Max}} \frac{1}{[S]} + \frac{1}{V_{Max}}$$

the maximum reaction velocity (V_{max}) and the Michaelis-Menten constant (K_M) were determined. The maximum reaction velocity and the initial enzyme concentration [E_0] allow the determination of the substrate turnover. Of note, k_{cat} stands for the catalytic constant.

$$k_{cat} = \frac{V_{\max}}{[E_0]}.$$

For this kinetic investigation, experimental conditions were similar to those described. The tRNA concentration varied from 20-400pmol tRNA and a final mt-AspRS concentration of 0.02 µM was used. The reaction was stopped after 20s, 40s, 60s and 80s.

9 Analysis of protein samples

9.1 Western Blot

9.1.1 "Home made" transfer

The protein samples were mixed with Laemmli-buffer, incubated 8min at 95°C, centrifuge for 5min (14,000g) and loaded on the SDS-PAGE. The run was stopped when the migration front reached the bottom. Proteins were transferred to a PVDF membrane (Immobilon P, 0.45µm, Millipore) with a semi-dry transfer system (Bio-rad). The membrane was briefly activated in EtOH, washed in H_20 and equilibrated in transfer buffer (20mM Tris-HCl; pH=7,5 at 25°C, 150mM Glycine, 0.02% SDS, 20% EtOH). The sandwich were constituted from the bottom to the top as following: two two

layers of whatman paper (9x7cm) presoaked in the transfer buffer, equilibrated PVDF membrane, gel and two additional layers of presoaked whatman paper. The transfer was performed for 1h at 10V and constant 400mA.

9.1.2 Commercial transfer kit

The systematic analysis of the localization of the full set of aaRSs was performed using the Trans-blot Turbo Transfer System from Biorad. Samples were prepared as described before and separated on a Mini-PROTEAN ® Precast Gels (Biorad) at 200V. The transfer to the membrane was performed as described in the manual.

9.1.3 Immuno detection:

At first the PVDF membrane was blocked with 5% milk in 1xTBST (50mM Tris-HCl; pH=7,5 at 25°C, 150mM NaCl, 1% Tween-20; pH=7.5) for 60min. Subsequently, the membrane was washed three times in 1xTBST for 10min. The incubation with the primary antibody (diluted in 1xTBST+ 1% milk) was performed at room temperature for 1h or at 4°C for over night. All used dilutions for antibodies are indicated in **Table 10**. After the incubation, the membrane was washed three times with 1xTBST. The incubation with the secondary antibody was performed for 1h at room temperature. Afterwards the membrane was washed three times in TBST and briefly dried on a Whatmann paper, and incubated for 5min with 2ml of enhanced chemiluminescence western blot substrate. The chemiluminescence signal was detected with Biorad Gel documentation system and analyzed with the QuantityOne Software package.

9.2 Native Gel electrophoresis

Native proteins and protein complexes can be separated using blue native gel electrophoresis (BN-PAGES; Schägger and von Jagow, 1991). In BN-PAGE, Coomassie G250 as a charge-shift molecule replaced SDS. In all experiments, ready-to-

use precasted 4-16% or 3-12% gels (NativePAGE^{TM}Novex from Invitrogen) were used. Samples were prepared as followed. Cells were harvested from a 10cm culture plate, washed with PBS buffer and centrifuged at 800g. To keep the possible interaction, the cells were lysed in 200µl Pierce-IP Lysis buffer (Pierce Scientific), carefully mixed by pipetting and incubated for 10min on ice. All subsequent steps were performed on ice and at 4°C in a cold room. The lysate was centrifuged at 14,000g for 15min. To prepare the samples, 10µl of supernatant from the centrifuge lysate were mixed with 4µl NativePAGE Sample buffer (4x), 0.5µl NativePAGE G-250 Sample Additive (5%) and 1.5µl deionized water. The samples were loaded on a precasted gel together with 7 µl of NativeMark^{TM} molecular weight marker and migrates at 150V at 4°C till the dye front reached 1/3rd of the gel. Afterwards, the Dark Blue Cathode Buffer (1% NativePAGE Cathode additive and 1x NativePAGE Running buffer) was replaced by the Light Blue Cathode buffer (0.1% NativePAGE Cathode additive and 1x NativePAGE Running buffer). The gel was further run till the dye front reached the bottom. Western blotting was performed as described before with minor modifications. Briefly, the native gel was equilibrated for 5min in transfer buffer and the transfer time was extended to 90min. After the transfer, the membrane was destained and shortly washed in TBS-T. Immuno detection was performed as described before (9.1.3).

9.3 2D-Gelelectrophoresis

For 2-D Gel electrophoresis, mitochondrial pellets were solubilized in 7M Urea, 2M Thiourea, 4% CHAPS and 20mM Tris-HCl pH=8.5 for 1h at 30°C. Samples were prepared by a reduction step with 50mM DTT for 10min at room temperature and alkylation step with 100mM Iodoacetamid for 15min at 37°C, followed by Sephadex G-25 purification step. First-dimension separation was performed on immobilized pH gradient (IPG) strip at pH 5–8 (BioRad) with 60,000Vh. Second Dimension separation was performed on a 10% SDS-PAGE.

9.4 Mass spectrometry analysis

Identification of full length mt-AspRS and mt-AspRSΔExon13 were performed by colloberation partners at “Laboratoire de Spectrométrie de Masse des Interactions et des Systèmes, Chimie de la Matière Complexe” (Strasbourg, France). A detailed description of the method can be found in the article # 1 and 2.

Bibliography

Bibliography

Adams, J., and Kauffman, M. (2004). Development of the proteasome inhibitor Velcade (Bortezomib). Cancer Invest. *22*, 304–311.

Agrawal, R.K., and Sharma, M.R. (2012). Structural aspects of mitochondrial translational apparatus. Current Opinion in Structural Biology *22*, 797–803.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). The Mitochondrion.

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.

Antonellis, A., Ellsworth, R.E., Sambuughin, N., Puls, I., Abel, A., Lee-Lin, S.-Q., Jordanova, A., Kremensky, I., Christodoulou, K., Middleton, L.T., et al. (2003). Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. Am. J. Hum. Genet. *72*, 1293–1299.

Arif, A., Jia, J., Mukhopadhyay, R., Willard, B., Kinter, M., and Fox, P.L. (2009). Two-site Phosphorylation of EPRS Coordinates Multimodal Regulation of Noncanonical Translational Control Activity. Mol Cell *35*, 164–180.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. *25*, 25–29.

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

Barra, D., Schinina, M.E., Simmaco, M., Bannister, J.V., Bannister, W.H., Rotilio, G., and Bossa, F. (1984). The primary structure of human liver manganese superoxide dismutase. J. Biol. Chem. *259*, 12595–12601.

Bayat, V., Thiffault, I., Jaiswal, M., Tétreault, M., Donti, T., Sasarman, F., Bernard, G., Demers-Lamarche, J., Dicaire, M.-J., Mathieu, J., et al. (2012). Mutations in the mitochondrial methionyl-tRNA synthetase cause a neurodegenerative phenotype in flies and a recessive ataxia (ARSAL) in humans. PLoS Biol. *10*, e1001288.

Becker, H.D., Giegé, R., and Kern, D. (1996). Identity of Prokaryotic and Eukaryotic tRNAAsp for Aminoacylation by Aspartyl-tRNA Synthetase from Thermus thermophilus†. Biochemistry *35*, 7447–7458.

Becker, T., Böttinger, L., and Pfanner, N. (2012). Mitochondrial protein import: from transport pathways to an integrated network. Trends in Biochemical Sciences *37*, 85–91.

Belostotsky, R., Ben-Shalom, E., Rinat, C., Becker-Cohen, R., Feinstein, S., Zeligson, S., Segel, R., Elpeleg, O., Nassar, S., and Frishberg, Y. (2011). Mutations in the mitochondrial seryl-tRNA synthetase cause hyperuricemia, pulmonary hypertension, renal failure in infancy and alkalosis, HUPRA syndrome. Am. J. Hum. Genet. *88*, 193–200.

Van Berge, L., Kevenaar, J., Polder, E., Gaudry, A., Florentz, C., Sissler, M., Van der Knaap, M., and Scheper, G. (2012a). Pathogenic mutations causing LBSL affect mitochondrial aspartyl-tRNA synthetase in diverse ways. Biochemical Journal.

Van Berge, L., Dooves, S., Van Berkel, C.G.M., Polder, E., Van der Knaap, M.S., and Scheper, G.C. (2012b). Leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation is associated with cell-type-dependent splicing of mtAspRS mRNA. Biochem. J. *441*, 955–962.

Berger, K.H., and Yaffe, M.P. (1998). Prohibitin Family Members Interact Genetically with Mitochondrial Inheritance Components in Saccharomyces cerevisiae. Molecular and Cellular Biology *18*, 4043 –4052.

Bonnefond, L., Fender, A., Rudinger-Thirion, J., Giegé, R., Florentz, C., and Sissler, M. (2005). Toward the full set of human mitochondrial aminoacyl-tRNA synthetases: characterization of AspRS and TyrRS. Biochemistry *44*, 4805–4816.

Bonnefond, L., Frugier, M., Touzé, E., Lorber, B., Florentz, C., Giegé, R., Sauter, C., and Rudinger-Thirion, J. (2007). Crystal Structure of Human Mitochondrial Tyrosyl-tRNA Synthetase Reveals Common and Idiosyncratic Features. Structure *15*, 1505–1516.

Brindefalk, B., Viklund, J., Larsson, D., Thollesson, M., and Andersson, S.G.E. (2007). Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases. Mol. Biol. Evol. *24*, 743–756.

Brown, S.P., Miller, W.C., and Eason, J.M. (2006). Exercise Physiology: Basis of Human Movement in Health and Disease (Lippincott Williams & Wilkins).

Bullard, J.M., Cai, Y.C., Demeler, B., and Spremulli, L.L. (1999). Expression and characterization of a human mitochondrial phenylalanyl-tRNA synthetase. J. Mol. Biol. *288*, 567–577.

Bullard, J.M., Cai, Y.C., and Spremulli, L.L. (2000). Expression and characterization of the human mitochondrial leucyl-tRNA synthetase. Biochim. Biophys. Acta *1490*, 245–258.

Cai, Y.-C., Bullard, J.M., Thompson, N.L., and Spremulli, L.L. (2000). Interaction of Mitochondrial Elongation Factor Tu with Aminoacyl-tRNA and Elongation Factor Ts. J. Biol. Chem. *275*, 20308–20314.

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

Castellana, S., Vicario, S., and Saccone, C. (2011). Evolutionary patterns of the mitochondrial genome in Metazoa: exploring the role of mutation and selection in mitochondrial protein coding genes. Genome Biol Evol.

Chacinska, A., Koehler, C.M., Milenkovic, D., Lithgow, T., and Pfanner, N. (2009). Importing mitochondrial proteins: machineries and mechanisms. Cell *138*, 628–644.

Chimnaronk, S., Gravers Jeppesen, M., Suzuki, T., Nyborg, J., and Watanabe, K. (2005). Dual-mode recognition of noncanonical tRNAsSer by seryl-tRNA synthetase in mammalian mitochondria. EMBO J *24*, 3369–3379.

Chinnery, P.F. (1993). Mitochondrial Disorders Overview. In GeneReviews™, R.A. Pagon, M.P. Adam, T.D. Bird, C.R. Dolan, C.-T. Fong, and K. Stephens, eds. (Seattle (WA): University of Washington, Seattle),.

Claros, M.G., and Vincens, P. (1996). Computational Method to Predict Mitochondrially Imported Proteins and their Targeting Sequences. European Journal of Biochemistry *241*, 779–786.

Cotter, D., Guda, P., Fahy, E., and Subramaniam, S. (2004). MitoProteome: mitochondrial protein sequence database and annotation system. Nucleic Acids Res *32*, D463–D467.

Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., and Leberman, R. (1990). A second class of synthetase structure revealed by X-ray analysis of Escherichia coli seryl-tRNA synthetase at 2.5 A. Nature *347*, 249–255.

Dias, J., Octobre, G., Kobbi, L., Comisso, M., Flisiak, S., and Mirande, M. (2012). Activation of human mitochondrial lysyl-tRNA synthetase upon maturation of its premitochondrial precursor. Biochemistry *51*, 909–916.

De Duve, C. (1988). Transfer RNAs: the second genetic code. Nature *333*, 117–118.

Edvardson, S., Shaag, A., Kolesnikova, O., Gomori, J.M., Tarassov, I., Einbinder, T., Saada, A., and Elpeleg, O. (2007). Deleterious Mutation in the Mitochondrial Arginyl-Transfer RNA Synthetase Gene Is Associated with Pontocerebellar Hypoplasia. The American Journal of Human Genetics *81*, 857–862.

Elo, J.M., Yadavalli, S.S., Euro, L., Isohanni, P., Götz, A., Carroll, C.J., Valanne, L., Alkuraya, F.S., Uusimaa, J., Paetau, A., et al. (2012). Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy. Hum. Mol. Genet.

Emanuelsson, O., Nielsen, H., Brunak, S., and Von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. *300*, 1005–1016.

Engelhard, J., Christian, B.E., Weingarten, L., Kuntz, G., Spremulli, L.L., and Dick, T.P. (2011). In situ kinetic trapping reveals a fingerprint of reversible protein thiol oxidation in the mitochondrial matrix. Free Radical Biology and Medicine *50*, 1234–1241.

Enriquez, J.A., and Attardi, G. (1996). Evidence for aminoacylation-induced conformational changes in human mitochondrial tRNAs. Proc Natl Acad Sci U S A *93*, 8300–8305.

Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990). Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. Nature *347*, 203–206.

Fechter, P., Rudinger, J., Giegé, R., and Théobald-Dietrich, A. (1998). Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: production and activity. FEBS Lett. *436*, 99–103.

Fender, A., Sauter, C., Messmer, M., Pütz, J., Giegé, R., Florentz, C., and Sissler, M. (2006). Loss of a primordial identity element for a mammalian mitochondrial aminoacylation system. J. Biol. Chem. *281*, 15980–15986.

Fender, A., Gaudry, A., Jühling, F., Sissler, M., and Florentz, C. (2012). Adaptation of aminoacylation identity rules to mammalian mitochondria. Biochimie *94*, 1090–1097.

Florentz, C., Sohm, B., Tryoen-Tóth, P., Pütz, J., and Sissler, M. (2003). Human mitochondrial tRNAs in health and disease. Cell. Mol. Life Sci. *60*, 1356–1375.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., Von Mering, C., et al. (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res *41*, D808–D815.

Fujiki, M., and Verner, K. (1993). Coupling of cytosolic protein synthesis and mitochondrial protein import in yeast. Evidence for cotranslational import in vivo. J. Biol. Chem. *268*, 1914–1920.

Fukui, H., Hanaoka, R., and Kawahara, A. (2009). Noncanonical Activity of Seryl-tRNA Synthetase Is Involved in Vascular Development. Circulation Research *104*, 1253–1259.

Gakh, O., Cavadini, P., and Isaya, G. (2002). Mitochondrial processing peptidases. Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research *1592*, 63–77.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. *31*, 3784–3788.

Gattermann, N. (2000). From sideroblastic anemia to the role of mitochondrial DNA mutations in myelodysplastic syndromes. Leuk. Res. *24*, 141–151.

Gaudry, A., Lorber, B., Neuenfeldt, A., Sauter, C., Florentz, C., and Sissler, M. (2012). Re-designed N-terminus enhances expression, solubility and crystallizability of mitochondrial protein. Protein Eng. Des. Sel. *25*, 473–481.

Graveley, B.R. (2001). Alternative splicing: increasing diversity in the proteomic world. Trends Genet. *17*, 100–107.

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

Gregersen, N., Hansen, J., and Palmfeldt, J. (2012). Mitochondrial proteomics--a tool for the study of metabolic disorders. J. Inherit. Metab. Dis. *35*, 715–726.

Gruschke, S., and Ott, M. (2010). The polypeptide tunnel exit of the mitochondrial ribosome is tailored to meet the specific requirements of the organelle. BioEssays *32*, 1050–1057.

Guo, M., and Schimmel, P. (2013). Essential nontranslational functions of tRNA synthetases. Nat. Chem. Biol. *9*, 145–153.

Guo, M., Yang, X.-L., and Schimmel, P. (2010). New functions of aminoacyl-tRNA synthetases beyond translation. Nat. Rev. Mol. Cell Biol. *11*, 668–674.

Han, J.M., Lee, M.J., Park, S.G., Lee, S.H., Razin, E., Choi, E.-C., and Kim, S. (2006). Hierarchical network between the components of the multi-tRNA synthetase complex: implications for complex formation. J. Biol. Chem. *281*, 38663–38667.

Han, J.M., Jeong, S.J., Park, M.C., Kim, G., Kwon, N.H., Kim, H.K., Ha, S.H., Ryu, S.H., and Kim, S. (2012). Leucyl-tRNA synthetase is an intracellular leucine sensor for the mTORC1-signaling pathway. Cell *149*, 410–424.

Hausmann, C.D., and Ibba, M. (2008). Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. FEMS Microbiology Reviews *32*, 705–721.

Helm, M., Giegé, R., and Florentz, C. (1999). A Watson-Crick base-pair-disrupting methyl group (m1A9) is sufficient for cloverleaf folding of human mitochondrial tRNALys. Biochemistry *38*, 13338–13346.

Helm, M., Brulé, H., Friede, D., Giegé, R., Pütz, D., and Florentz, C. (2000). Search for characteristic structural features of mammalian mitochondrial tRNAs. RNA *6*, 1356–1379.

Holt, I.J., and Reyes, A. (2012). Human Mitochondrial DNA Replication. Cold Spring Harb Perspect Biol *4*.

Hoogenraad, N.J., Ward, L.A., and Ryan, M.T. (2002). Import and assembly of proteins into mitochondria of mammalian cells. Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research *1592*, 97–105.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. *37*, 1–13.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44–57.

Ibba, M., Francklyn, C., and Cusack, S. (2005). The Aminoacyl-tRNA synthetases (Georgetown, TX: Landes Biosciences).

Igloi, G.L. (1988). Interaction of tRNAs and of phosphorothioate-substituted nucleic acids with an organomercurial. Probing the chemical environment of thiolated residues by affinity electrophoresis. Biochemistry *27*, 3842–3849.

Jester, B.C., Drillien, R., Ruff, M., and Florentz, C. (2011). Using Vaccinia's innate ability to introduce DNA into mammalian cells for production of recombinant proteins. J. Biotechnol. *156*, 211–213.

Jordanova, A., Irobi, J., Thomas, F.P., Van Dijck, P., Meerschaert, K., Dewil, M., Dierick, I., Jacobs, A., De Vriendt, E., Guergueltcheva, V., et al. (2006). Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy. Nat. Genet. *38*, 197–202.

Jorgensen, R., Søgaard, T.M., Rossing, A.B., Martensen, P.M., and Justesen, J. (2000). Identification and characterization of human mitochondrial tryptophanyl-tRNA synthetase. J. Biol. Chem. *275*, 16820–16826.

Kaminska, M., Havrylenko, S., Decottignies, P., Gillet, S., Le Maréchal, P., Negrutskii, B., and Mirande, M. (2009). Dissection of the structural organization of the aminoacyl-tRNA synthetase complex. J. Biol. Chem. *284*, 6053–6060.

Kappel, S., Matthess, Y., Kaufmann, M., and Strebhardt, K. (2007). Silencing of mammalian genes by tetracycline-inducible shRNA expression. Nat. Protocols *2*, 3257–3269.

Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M., and Stamm, S. (2013). Function of alternative splicing. Gene *514*, 1–30.

Kellems, R.E., Allison, V.F., and Butow, R.A. (1975). Cytoplasmic type 80S ribosomes associated with yeast mitochondria. IV. Attachment of ribosomes to the outer membrane of isolated mitochondria. J. Cell Biol. *65*, 1–14.

Kelley, S.O., Steinberg, S.V., and Schimmel, P. (2000). Functional defects of pathogenic human mitochondrial tRNAs related to structural fragility. Nat Struct Mol Biol *7*, 862–865.

Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. Nature Reviews Genetics.

Kim, D.G., Choi, J.W., Lee, J.Y., Kim, H., Oh, Y.S., Lee, J.W., Tak, Y.K., Song, J.M., Razin, E., Yun, S.-H., et al. (2012). Interaction of two translational components, lysyl-tRNA synthetase and p40/37LRP, in plasma membrane promotes laminin-dependent cell migration. FASEB J. *26*, 4142–4159.

Kim, S., You, S., and Hwang, D. (2011). Aminoacyl-tRNA synthetases and tumorigenesis: more than housekeeping. Nat Rev Cancer *11*, 708–718.

Kise, Y., Lee, S.W., Park, S.G., Fukai, S., Sengoku, T., Ishii, R., Yokoyama, S., Kim, S., and Nureki, O. (2004). A short peptide insertion crucial for angiostatic activity of human tryptophanyl-tRNA synthetase. Nat. Struct. Mol. Biol. *11*, 149–156.

Klipcan, L., Levin, I., Kessler, N., Moor, N., Finarov, I., and Safro, M. (2008). The tRNA-induced conformational activation of human mitochondrial phenylalanyl-tRNA synthetase. Structure *16*, 1095–1104.

Klipcan, L., Moor, N., Finarov, I., Kessler, N., Sukhanova, M., and Safro, M.G. (2012). Crystal Structure of Human Mitochondrial PheRS Complexed with tRNAPhe in the Active "Open" State. Journal of Molecular Biology *415*, 527–537.

Van der Knaap, M.S., Van der Voorn, P., Barkhof, F., Van Coster, R., Krägeloh-Mann, I., Feigenbaum, A., Blaser, S., Vles, J.S.H., Rieckmann, P., and Pouwels, P.J.W. (2003). A new leukoencephalopathy with brainstem and spinal cord involvement and high lactate. Ann. Neurol. *53*, 252–258.

Koc, E.C., and Spremulli, L.L. (2002). Identification of Mammalian Mitochondrial Translational Initiation Factor 3 and Examination of Its Role in Initiation Complex Formation with Natural mRNAs. J. Biol. Chem. *277*, 35541–35549.

Kulawiak, B., Höpker, J., Gebert, M., Guiard, B., Wiedemann, N., and Gebert, N. (2013). The mitochondrial protein import machinery has multiple connections to the respiratory chain. Biochimica Et Biophysica Acta (BBA) - Bioenergetics *1827*, 612–626.

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

Lee, J.W., Beebe, K., Nangle, L.A., Jang, J., Longo-Guess, C.M., Cook, S.A., Davisson, M.T., Sundberg, J.P., Schimmel, P., and Ackerman, S.L. (2006). Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. Nature *443*, 50–55.

Levin, I., Kessler, N., Moor, N., Klipcan, L., Koc, E., Templeton, P., Spremulli, L., and Safro, M. (2007). Purification, crystallization and preliminary X-ray characterization of a human mitochondrial phenylalanyl-tRNA synthetase. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *63*, 761–764.

Levinger, L., Mörl, M., and Florentz, C. (2004). Mitochondrial tRNA 3′ end metabolism and human disease. Nucleic Acids Res *32*, 5430–5441.

Liu, M., and Spremulli, L. (2000). Interaction of mammalian mitochondrial ribosomes with the inner membrane. J. Biol. Chem *275*, 29400–29406.

MacKenzie, J.A., and Payne, R.M. (2007). Mitochondrial protein import and human health and disease. Biochimica Et Biophysica Acta (BBA) - Molecular Basis of Disease *1772*, 509–523.

Marc, P., Margeot, A., Devaux, F., Blugeon, C., Corral-Debrinski, M., and Jacq, C. (2002). Genome-wide analysis of mRNAs targeted to yeast mitochondria. EMBO Rep. *3*, 159–164.

Matthews, D.E., Hessler, R.A., Denslow, N.D., Edwards, J.S., and O'Brien, T.W. (1982). Protein composition of the bovine mitochondrial ribosome. J. Biol. Chem. *257*, 8788–8794.

Meisinger, C., Sickmann, A., and Pfanner, N. (2008). The Mitochondrial Proteome: From Inventory to Function. Cell *134*, 22–24.

Messmer, M., Blais, S.P., Balg, C., Chênevert, R., Grenier, L., Lagüe, P., Sauter, C., Sissler, M., Giegé, R., Lapointe, J., et al. (2009). Peculiar inhibition of human mitochondrial aspartyl-tRNA synthetase by adenylate analogs. Biochimie *91*, 596–603.

Messmer, M., Florentz, C., Schwenzer, H., Scheper, G.C., Van der Knaap, M.S., Maréchal-Drouard, L., and Sissler, M. (2011). A human pathology-related mutation prevents import of an aminoacyl-tRNA synthetase into mitochondria. Biochem. J *433*, 441–446.

Mirande, M., Cirakoğlu, B., and Waller, J.P. (1982). Macromolecular complexes from sheep and rabbit containing seven aminoacyl-tRNA synthetases. III. Assignment of aminoacyl-tRNA synthetase activities to the polypeptide components of the complexes. J. Biol. Chem. *257*, 11056–11063.

Mogk, A., Schmidt, R., and Bukau, B. (2007). The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. Trends in Cell Biology *17*, 165–172.

Mossmann, D., Meisinger, C., and Vögtle, F.-N. (2012). Processing of mitochondrial presequences. Biochim. Biophys. Acta *1819*, 1098–1106.

Mudge, S.J., Williams, J.H., Eyre, H.J., Sutherland, G.R., Cowan, P.J., and Power, D.A. (1998). Complex organisation of the 5'-end of the human glycine tRNA synthetase gene. Gene *209*, 45–50.

Mukherjee, K., Campos, H., and Kolaczkowski, B. (2013). Evolution of Animal and Plant Dicers: Early Parallel Duplications and Recurrent Adaptation of Antiviral RNA Binding in Plants. Mol Biol Evol *30*, 627–641.

Nabholz, C.E., Horn, E.K., and Schneider, A. (1999). tRNAs and proteins are imported into mitochondria of Trypanosoma brucei by two distinct mechanisms. Mol. Biol. Cell *10*, 2547–2557.

Nagaike, T., Suzuki, T., and Ueda, T. (2008). Polyadenylation in mammalian mitochondria: insights from recent studies. Biochim. Biophys. Acta *1779*, 266–269.

Nagao, A., Suzuki, T., Katoh, T., Sakaguchi, Y., and Suzuki, T. (2009). Biogenesis of glutaminyl-mt tRNAGln in human mitochondria. Proc. Natl. Acad. Sci. U.S.A. *106*, 16209–16214.

Nett, J.H., and Trumpower, B.L. (1996). Dissociation of Import of the Rieske Iron-Sulfur Protein into Saccharomyces cerevisiae Mitochondria from Proteolytic Processing of the Presequence. J. Biol. Chem. *271*, 26713–26716.

Neuenfeldt, A., Lorber, B., Ennifar, E., Gaudry, A., Sauter, C., Sissler, M., and Florentz, C. (2013). Thermodynamic properties distinguish human mitochondrial aspartyl-tRNA synthetase from bacterial homolog with same 3D architecture. Nucleic Acids Res. *41*, 2698–2708.

Ofir-Birin, Y., Fang, P., Bennett, S.P., Zhang, H.-M., Wang, J., Rachmin, I., Shapiro, R., Song, J., Dagan, A., Pozo, J., et al. (2013). Structural Switch of Lysyl-tRNA Synthetase between Translation and Transcription. Molecular Cell *49*, 30–42.

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

Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.-E., Walford, G.A., Sugiana, C., Boneh, A., Chen, W.K., et al. (2008). A mitochondrial protein compendium elucidates complex I disease biology. Cell *134*, 112–123.

Park, M.C., Kang, T., Jin, D., Han, J.M., Kim, S.B., Park, Y.J., Cho, K., Park, Y.W., Guo, M., He, W., et al. (2012). Secreted human glycyl-tRNA synthetase implicated in defense against ERK-activated tumorigenesis. Proc Natl Acad Sci U S A *109*, E640–E647.

Park, S.G., Schimmel, P., and Kim, S. (2008). Aminoacyl tRNA synthetases and their connections to disease. Proc. Natl. Acad. Sci. U.S.A. *105*, 11043–11049.

Patzkó, A., and Shy, M.E. (2011). Update on Charcot-Marie-Tooth disease. Curr Neurol Neurosci Rep *11*, 78–88.

Perfettini, J.-L., Roumier, T., and Kroemer, G. (2005). Mitochondrial fusion and fission in the control of apoptosis. Trends in Cell Biology *15*, 179–183.

Pierce, S.B., Chisholm, K.M., Lynch, E.D., Lee, M.K., Walsh, T., Opitz, J.M., Li, W., Klevit, R.E., and King, M.-C. (2011). Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. Proc. Natl. Acad. Sci. U.S.A. *108*, 6543–6548.

Pierce, S.B., Gersak, K., Michaelson-Cohen, R., Walsh, T., Lee, M.K., Malach, D., Klevit, R.E., King, M.-C., and Levy-Lahad, E. (2013). Mutations in LARS2, encoding mitochondrial leucyl-tRNA

synthetase, lead to premature ovarian failure and hearing loss in Perrault syndrome. Am. J. Hum. Genet. *92*, 614–620.

Ray, P.S., Arif, A., and Fox, P.L. (2007). Macromolecular complexes as depots for releasable regulatory proteins. Trends Biochem. Sci *32*, 158–164.

Di Re, M., Sembongi, H., He, J., Reyes, A., Yasukawa, T., Martinsson, P., Bailey, L.J., Goffart, S., Boyd-Kirkup, J.D., Wong, T.S., et al. (2009). The accessory subunit of mitochondrial DNA polymerase gamma determines the DNA content of mitochondrial nucleoids in human cultured cells. Nucleic Acids Res. *37*, 5701–5713.

Riley, L.G., Cooper, S., Hickey, P., Rudinger-Thirion, J., McKenzie, M., Compton, A., Lim, S.C., Thorburn, D., Ryan, M.T., Giegé, R., et al. (2010). Mutation of the mitochondrial tyrosyl-tRNA synthetase gene, YARS2, causes myopathy, lactic acidosis, and sideroblastic anemia--MLASA syndrome. Am. J. Hum. Genet. *87*, 52–59.

Rodenburg, R.J.T. (2011). Biochemical diagnosis of mitochondrial disorders. J. Inherit. Metab. Dis. *34*, 283–292.

Rötig, A., Cormier, V., Blanche, S., Bonnefont, J.P., Ledeist, F., Romero, N., Schmitz, J., Rustin, P., Fischer, A., and Saudubray, J.M. (1990). Pearson's marrow-pancreas syndrome. A multisystem mitochondrial disorder in infancy. J. Clin. Invest. *86*, 1601–1608.

Sajish, M., Zhou, Q., Kishi, S., Valdez, D.M., Jr, Kapoor, M., Guo, M., Lee, S., Kim, S., Yang, X.-L., and Schimmel, P. (2012). Trp-tRNA synthetase bridges DNA-PKcs to PARP-1 to link IFN-γ and p53 signaling. Nat. Chem. Biol. *8*, 547–554.

Sampath, P., Mazumder, B., Seshadri, V., Gerber, C.A., Chavatte, L., Kinter, M., Ting, S.M., Dignam, J.D., Kim, S., Driscoll, D.M., et al. (2004). Noncanonical Function of Glutamyl-Prolyl-tRNA Synthetase: Gene-Specific Silencing of Translation. Cell *119*, 195–208.

Sasarman, F., Nishimura, T., Thiffault, I., and Shoubridge, E.A. (2012). A novel mutation in YARS2 causes myopathy with lactic acidosis and sideroblastic anemia. Human Mutation.

Schägger, H., and Von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Anal. Biochem. *199*, 223–231.

Schägger, H., Cramer, W.A., and Von Jagow, G. (1994). Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. Anal. Biochem. *217*, 220–230.

Schapira, A.H.V. (2002). Primary and secondary defects of the mitochondrial respiratory chain. J Inherit Metab Dis *25*, 207–214.

Scheffler, I.E. (2008). Mitochondria (Hoboken, New Jersey: Wiley).

Scheper, G.C., Van der Klok, T., Van Andel, R.J., Van Berkel, C.G.M., Sissler, M., Smet, J., Muravina, T.I., Serkov, S.V., Uziel, G., Bugiani, M., et al. (2007). Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. Nat Genet *39*, 534–539.

Schleiff, E., and Becker, T. (2011). Common ground for protein translocation: access control for mitochondria and chloroplasts. Nat Rev Mol Cell Biol *12*, 48–59.

Schmidt, O., Pfanner, N., and Meisinger, C. (2010). Mitochondrial protein import: from proteomics to functional mechanisms. Nat Rev Mol Cell Biol *11*, 655–667.

Schon, E.A., DiMauro, S., and Hirano, M. (2012). Human mitochondrial DNA: roles of inherited and somatic mutations. Nat Rev Genet *13*, 878–890.

Schwenzer, H., Zoll, J., Florentz, C., and Sissler, M. (2013). Pathogenic Implications of Human Mitochondrial Aminoacyl-tRNA Synthetases. Top Curr Chem.

Shalak, V., Kaminska, M., and Mirande, M. (2009). Translation Initiation from Two In-Frame AUGs Generates Mitochondrial and Cytoplasmic Forms of the p43 Component of the Multisynthetase Complex. Biochemistry *48*, 9959–9968.

Sharma, M.R., Koc, E.C., Datta, P.P., Booth, T.M., Spremulli, L.L., and Agrawal, R.K. (2003). Structure of the Mammalian Mitochondrial Ribosome Reveals an Expanded Functional Role for Its Component Proteins. Cell *115*, 97–108.

Shiba, K., Schimmel, P., Motegi, H., and Noda, T. (1994). Human glycyl-tRNA synthetase. Wide divergence of primary structure from bacterial counterpart and species-specific aminoacylation. J. Biol. Chem. *269*, 30049–30055.

Singer, S.J., and Nicolson, G.L. (1972). The fluid mosaic model of the structure of cell membranes. Science *175*, 720–731.

Sissler, M., Helm, M., Frugier, M., Giege, R., and Florentz, C. (2004). Aminoacylation properties of pathology-related human mitochondrial tRNA(Lys) variants. RNA *10*, 841–853.

Sissler, M., Lorber, B., Messmer, M., Schaller, A., Pütz, J., and Florentz, C. (2008). Handling mammalian mitochondrial tRNAs and aminoacyl-tRNA synthetases for functional and structural characterization. Methods *44*, 176–189.

Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics *4*, 1581–1590.

Sohm, B., Frugier, M., Brulé, H., Olszak, K., Przykorska, A., and Florentz, C. (2003). Towards understanding human mitochondrial leucine aminoacylation identity. J. Mol. Biol. *328*, 995–1010.

Speers, A.E., and Wu, C.C. (2007). Proteomics of integral membrane proteins--theory and application. Chem. Rev. *107*, 3687–3714.

Spelbrink, J.N. (2010). Functional organization of mammalian mitochondrial DNA in nucleoids: history, recent developments, and future challenges. IUBMB Life *62*, 19–32.

Storkebaum, E., Leitão-Gonçalves, R., Godenschwege, T., Nangle, L., Mejia, M., Bosmans, I., Ooms, T., Jacobs, A., Van Dijck, P., Yang, X.-L., et al. (2009). Dominant mutations in the tyrosyl-tRNA synthetase gene recapitulate in Drosophila features of human Charcot-Marie-Tooth neuropathy. Proc. Natl. Acad. Sci. U.S.A. *106*, 11782–11787.

Sutton, A., Khoury, H., Prip-Buus, C., Cepanec, C., Pessayre, D., and Degoul, F. (2003). The Ala16Val genetic dimorphism modulates the import of human manganese superoxide dismutase into rat liver mitochondria. Pharmacogenetics *13*, 145–157.

Suzuki, H., Ueda, T., Taguchi, H., and Takeuchi, N. (2007). Chaperone properties of mammalian mitochondrial translation elongation factor Tu. J. Biol. Chem *282*, 4076–4084.

Suzuki, T., Nagao, A., and Suzuki, T. (2011). Human mitochondrial diseases caused by lack of taurine modification in mitochondrial tRNAs. Wiley Interdiscip Rev RNA *2*, 376–386.

Takakubo, F., Cartwright, P., Hoogenraad, N., Thorburn, D.R., Collins, F., Lithgow, T., and Dahl, H.H. (1995). An amino acid substitution in the pyruvate dehydrogenase E1 alpha gene, affecting mitochondrial import of the precursor protein. Am. J. Hum. Genet. *57*, 772–780.

Taylor, S.W., Fahy, E., Zhang, B., Glenn, G.M., Warnock, D.E., Wiley, S., Murphy, A.N., Gaucher, S.P., Capaldi, R.A., Gibson, B.W., et al. (2003). Characterization of the human heart mitochondrial proteome. Nat Biotech *21*, 281–286.

Temperley, R., Richter, R., Dennerlein, S., Lightowlers, R.N., and Chrzanowska-Lightowlers, Z.M. (2010). Hungry codons promote frameshifting in human mitochondrial ribosomes. Science *327*, 301.

Timmer, J.C., Enoksson, M., Wildfang, E., Zhu, W., Igarashi, Y., Denault, J.-B., Ma, Y., Dummitt, B., Chang, Y.-H., Mast, A.E., et al. (2007). Profiling constitutive proteolytic events in vivo. Biochemical Journal *407*, 41.

Tolkunova, E., Park, H., Xia, J., King, M.P., and Davidson, E. (2000). The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual alternative splicing of the primary transcript. J. Biol. Chem. *275*, 35063–35069.

Umeda, N., Suzuki, T., Yukawa, M., Ohya, Y., Shindo, H., Watanabe, K., and Suzuki, T. (2005). Mitochondria-specific RNA-modifying enzymes responsible for the biosynthesis of the wobble base in mitochondrial tRNAs. Implications for the molecular pathogenesis of human mitochondrial diseases. J. Biol. Chem. *280*, 1613–1624.

Varshney, U., Lee, C.P., and RajBhandary, U.L. (1991). Direct analysis of aminoacylation levels of tRNAs in vivo. Application to studying recognition of Escherichia coli initiator tRNA mutants by glutaminyl-tRNA synthetase. J. Biol. Chem. *266*, 24712–24718.

Verner, K. (1993). Co-translational protein import into mitochondria: an alternative view. Trends Biochem. Sci. *18*, 366–371.

Vögtle, F.-N., Wortelkamp, S., Zahedi, R.P., Becker, D., Leidhold, C., Gevaert, K., Kellermann, J., Voos, W., Sickmann, A., Pfanner, N., et al. (2009). Global Analysis of the Mitochondrial N-Proteome Identifies a Processing Peptidase Critical for Protein Stability. Cell *139*, 428–439.

Voos, W. (2013). Chaperone–protease networks in mitochondrial protein homeostasis. Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research *1833*, 388–399.

Wakasugi, K., and Schimmel, P. (1999). Two distinct cytokines released from a human aminoacyl-tRNA synthetase. Science *284*, 147–151.

Wakasugi, K., Slike, B.M., Hood, J., Otani, A., Ewalt, K.L., Friedlander, M., Cheresh, D.A., and Schimmel, P. (2002). A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. Proc. Natl. Acad. Sci. U.S.A. *99*, 173–177.

Wang, X., Yan, Q., and Guan, M.-X. (2010). Combination of the loss of cmnm5U34 with the lack of s2U34 modifications of tRNALys, tRNAGlu and tRNAGln altered mitochondrial biogenesis and respiration. J Mol Biol *395*, 1038.

Watanabe, K. (2010). Unique features of animal mitochondrial translation systems. The non-universal genetic code, unusual features of the translational apparatus and their relevance to human mitochondrial diseases. Proc. Jpn. Acad., Ser. B, Phys. Biol. Sci. *86*, 11–39.

Wiedemann, N., Frazier, A.E., and Pfanner, N. (2004). The Protein Import Machinery of Mitochondria. Journal of Biological Chemistry *279*, 14473 –14476.

Wittenhagen, L.M., Roy, M.D., and Kelley, S.O. (2003). The pathogenic U3271C human mitochondrial tRNA(Leu(UUR)) mutation disrupts a fragile anticodon stem. Nucleic Acids Res. *31*, 596–601.

Xu, X., Shi, Y., Zhang, H.-M., Swindell, E.C., Marshall, A.G., Guo, M., Kishi, S., and Yang, X.-L. (2012). Unique domain appended to vertebrate tRNA synthetase is essential for vascular development. Nat Commun *3*, 681.

Yannay-Cohen, N., Carmi-Levy, I., Kay, G., Yang, C.M., Han, J.M., Kemeny, D.M., Kim, S., Nechushtan, H., and Razin, E. (2009). LysRS Serves as a Key Signaling Molecule in the Immune Response by Regulating Gene Expression. Molecular Cell *34*, 603–611.

Yokobori, S., Suzuki, T., and Watanabe, K. (2001). Genetic code variations in mitochondria: tRNA as a major determinant of genetic code plasticity. J. Mol. Evol. *53*, 314–326.

Yokogawa, T., Shimada, N., Takeuchi, N., Benkowski, L., Suzuki, T., Omori, A., Ueda, T., Nishikawa, K., Spremulli, L.L., and Watanabe, K. (2000). Characterization and tRNA recognition of mammalian mitochondrial seryl-tRNA synthetase. J. Biol. Chem. *275*, 19913–19920.

Annexe

French summary of the manuscript submitted at University of Strasbourg

INTRODUCTION

Les mitochondries ont un rôle essentiel dans le métabolisme énergétique cellulaire. Elles sont impliquées dans des fonctions majeures telles que la respiration, la production d'énergie, ou la biosynthèse de métabolites. En conséquence, tout dysfonctionnement mitochondrial (mt) est généralement délétère pour la cellule, et souvent associé à des pathologies humaines, désordres neurodégénératifs ou musculaires. Les mitochondries ont une machinerie de traduction propre dédiée, chez l'homme, à la synthèse de seulement 13 protéines, toutes des sous-unités des complexes de la chaine respiratoire. Alors que les 22 ARNt et les 2 ARNr requis sont également codés par le génome mitochondrial, toutes les autres macromolécules indispensables à la machinerie traductionnelle sont codées par le noyau, synthétisées dans le cytosol et importées dans la mitochondrie grâce à une pré-séquence d'adressage (MTS). Parmi les protéines importées, les aminoacyl-ARNt synthétases (aaRS) sont des enzymes clés dans le processus de traduction, responsables de l'attachement spécifique de chacun des 20 acides aminés sur l'ARNt correspondant. Les aaRS mt sont codés par un jeu de gènes nucléaires distincts de celui codant pour les aaRS à localisation cytosolique (à deux exceptions près). Les aaRS mt se distinguent donc par leur séquence (d'origine bactérienne pour la plupart) et la présence de la MTS, excisée après importation. Les études faites à ce jour révèlent des particularités (relâchements) fonctionnelles et structurales des aaRS mt comparées à leurs homologues cytosoliques ou bactériens, telles i) des activités enzymatiques réduites, ii) un nombre d'éléments d'identité (spécifient l'aminoacylation) restreints sur l'ARNt, iii) des propriétés biophysiques et de plasticité particulières (démontrées pour l'AspRS, enzyme spécifique de l'acide aspartique).

Récemment, le laboratoire d'accueil a contribué à corréler des mutations découvertes dans le gène nucléaire codant pour l'AspRS mt à certaines leucoencéphalopathies. Cette découverte a ouvert la voie vers un nouveau champ d'investigation puisqu'aujourd'hui, des mutations dans les gènes nucléaires de 9 autres aaRS mt ont été identifiées comme cause de pathologies. J'ai contribué à la rédaction de trois revues (présentées dans l'introduction du manuscrit), dans lesquelles j'ai répertorié l'implication de 65 mutations (combinées en 64 cas pathologiques) à ce jour décrites dans la littérature, et compilé les nombreuses routes décrites qui lient un effet moléculaire primaire d'une mutation à son expression phénotypique ou symptomatique. L'absence de défauts majeurs des propriétés canoniques des aaRS mt, dans certains des cas publiés, est un indice fort de l'existence de fonction(s) alternative(s) portée(s) par ces enzymes. Des fonctions alternatives ont déjà été décrites pour les aaRS à localisation cytosolique et leurs altérations expliquent, dans certains cas, les pathologies liées à ces enzymes. La bascule entre la fonction canonique et celle alternative est produite par *e.g.* un partenariat alternatif, une organisation macromoléculaire alternative, ou une relocalisation sous-cellulaire. Par opposition, les connaissances actuelles concernant de possibles fonctions alternatives ou simplement une organisation macromoléculaire pour les aaRS mt restent très succinctes. Le travail de cette thèse s'inscrit dans cette démarche par i) l'analyse de particularités évolutives, ii) la recherche de partenaires, iii) établissement de la localisation sous-mitochondriale d'aaRS(s) humaine et iv) l'étude de l'impact de mutations affectant des aaRS ou des ARNt et corrélées à des pathologies humaines.

RESULTATS/DISCUSSION

i. Produits d'épissage alternatif du gène de l'AspRS mt : vers la simplification d'un domaine ?

Lors du clonage de l'AspRS mt humaine, l'étape de PCR réalisée sur une banque d'ADNc a montré la présence de 2 produits de réaction. Les deux bandes ont été identifiées par séquençage : celle de haut poids moléculaire correspond à l'AspRS mt et l'autre correspond à l'AspRS mt épissé d'un exon [l'exon 13 code pour un peptide du domaine insertionnel de type bactérien (BID)]. Nous avons démontré i) que ce variant

co-existe avec l'ARNm de séquence complète dans tous les tissus analysés, ii) qu'il est poly-adénylé et pris en charge par les polysomes, mais iii) que la protéine correspondante n'est pas détectable sous forme stable. Des analyses de conservation de séquences et de conservation de topologies indiquent clairement un relâchement de pression de sélection au niveau du BID des AspRS mt, alors que des motifs de séquence et de structures sont maintenus dans les AspRS bactériennes. Aussi, la découverte dans d'autres organismes (souris, singe, sanglier, …) d'ARNm codant pour l'AspRS mt dont les délétions affectent des exons voisins (mais codant toujours pour des peptides du BID) favorise l'hypothèse de la perte de fonction de ce domaine pour les AspRS à localisation mitochondriale, et donc le relâchement de pression de sélection pour le maintenir intact. Cette découverte est un exemple supplémentaire de relâchement fonctionnelle de l'enzyme mt, par opposition à son homologue bactérien, et en dépit d'une origine évolutive commune. Ces travaux sont soumis pour publication.

ii) A la recherche de partenaire(s) cellulaire(s) d'aaRS mt

Une des manières les plus directes d'intégrer les aaRS dans de nouveaux réseaux fonctionnels ou activités cellulaires est d'identifier les partenaires cellulaires potentiels. Ceux-ci sont tout d'abord recherché pour l'AspRS mt. Des expériences de « cross-linking » in vitro m'ont conforté dans cette voie, puisque j'ai pu identifié l'AspRS mt dans un complexe macromoléculaire de très haut poids moléculaire. Ainsi, J'ai entrepris (après optimisation) des expériences de co-immunoprécipitation de l'AspRS mt (utilisant des anticorps polyclonaux anti-AspRS mt synthétisés à façon). L'AspRS mt est immuno-précipitée (à partir de mitochondries sont purifiées de cellules humaines HEK293T cultivées en suspension) et les protéines partenaires sont identifiées par LC-MS/MS. Après 3 séries d'expériences indépendantes, un premier jeu de partenaires possible a été identifié. Des analyses complémentaires (par couplage de l'AspRS mt recombinante à un réactif biotinylé) sont en cours au laboratoire.

iii) Organisation sous-mitochondriale des aaRS

Des expériences effectuées par d'autres ont démontré que le facteur d'élongation EF-Tu et les ribosomes mitochondriaux (2 autres acteurs clés de la traduction) sont localisés proches des membranes internes des mitochondries. Ces observations étayent

l'hypothèse selon laquelle la traduction mitochondriale puisse se faire à proximité des membranes internes. Il est donc essentiel d'établir le schéma de distribution des aaRS au sein de la mitochondrie. L'objectif est non seulement de définir le site de la traduction, mais aussi de corréler « localisation » à « fonction » et d'envisager la membrane interne des mitochondries comme possible plateforme d'ancrage aux aaRSs dédiées à l'aminoacylation.

Pour ce faire, J'ai entrepris 2 approches. Le principe de la première est de fractionner des mitochondries purifiées en extraits contenant des protéines matricielles, associées, ou incorporées aux membranes internes. La pureté des 3 sous-fractions est vérifiée par l'utilisation d'anticorps spécifiques de chacune de ces fractions [SOD2 pour la matrice, la protéine ribosomale MRPL39 pour les protéines associées à la membrane interne, et la prohibitine pour les protéines incorporées dans les membranes]. La seconde approche combine le fractionnement des mitochondries à l'expression de protéines recombinantes en utilisant le système d'expression BHK (cellule de Hamster) / Vaccine. Ce système permet l'expression modulée, la purification, la détection (par un anticorps unique contre une des étiquettes incorporées, généralement l'étiquette Flag), et la mutagenèse de tout protéine recombinante (y compris adressée à la mitochondrie) dans des cellules de mammifères.

J'ai tout d'abord validé (par une analyse comparative) l'utilisation de cellules de Hamster, et démontré que l'expression modulée de protéines recombinantes n'altère pas les propriétés d'adressage, d'internalisation, ou de localisation de protéines mitochondriales humaines. Aussi, le site de clivage naturel de l'AspRS mt a pu être défini, et identifié à distance du site prédit. Cette approche, en cours d'utilisation au laboratoire, aura de très nombreuses applications futures. Elle permettra en particulier d'analyser l'impact de mutations pathologiques répertoriées dans des protéines impliquées dans la biogenèse des mitochondries (article méthodologique en préparation).

Le protocole de fractionnement a ensuite été appliqué à des mitochondries ou des mitoplastes (mitochondries dépourvues de membranes externes) purifiées à partir de cellules en cultures, soit humaines (détection des aaRS par anticorps spécifiques), soit de hamster exprimant les aaRS humaines d'intérêt (détection des aaRS par un anticorps unique). Des expériences complémentaires de western blot, de northern blot ou de mass

spectrométrie ont été effectuées systématiquement et permettent de définir respectivement le contenu des fractions en aaRS, en ARNt, ou en protéines globales (protéome). Ainsi, le profil de répartition des aaRS a été établi (la plupart sont à la fois dans la matrice et accrochées au membranes), conforté par la co-localisation des ARNt correspondants. La double répartition observée pour certaines aaRS étaye l'hypothèse selon laquelle la membrane interne puisse servir de plateforme d'ancrage aux aaRS effectuant leur fonction canonique. Reste à démontrer que celles relarguées des membranes ont une fonction alternative.

iv) Impact de mutations affectant des aaRS ou des ARNt et corrélées à des pathologies humaines

Au cours de ma thèse, j'ai contribué à divers travaux visant à analyser l'impact de mutations corrélées à des pathologies (un article publié et un article en cours de préparation). J'ai contribué de manière plus marquée à l'étude suivante : V. Procaccio (UMR Pôle Santé d'Angers) nous a confié la découverte de 2 patients atteints du syndrome Leigh-like et dont l'analyse des exomes a identifié 2 mutations dans le gène nucléaire codant pour l'AsnRS mt (communication personnelle). Ne disposant pas de l'AsnRS mt recombinante, j'ai étudié l'effet des mutations pathologiques sur la réaction d'aminoacylation en analysant le taux de charge cellulaire de l'$ARNt^{Asn}$ mt par des expériences de northern. Les extraits cellulaires (ARN totaux extraits en condition acide) ont été obtenus à partir de lignées fibroblastes (fournies par V. Procaccio) des 2 patients, ainsi que de 2 contrôles sains. J'ai pu démontré que les mutations n'affectent pas le taux d'aminoacylation de l'$ARNt^{Asn}$ mt, mais induit une légère diminution de l'$ARNt^{Asn}$ lui même. J'ai analysé en outre l'impact d'une autre mutation identifiée chez ces mêmes patients, affectant une enzyme de modification post-transcriptionnelle d'ARNt (mto1). Il semblerait (à confirmer) que la pathogénèse survienne par addition d'effets aux conséquences individuelles modérées, mais cumulatifs.

CONCLUSION

Les travaux entrepris au cours de ma thèse ont conduit à une meilleure compréhension des aaRS mt. Ils ont permis d'illustrer d'avantage le relâchement fonctionnel de

l'AspRS mt, comparée à son homologue bactérien. L'analyse systématique de la répartition des aaRS à l'intérieure de la mitochondrie est inédite et permet i) non seulement de mieux définir la traduction mitochondriale, ii) mais aussi d'apporter une connaissance utile en vue d'identifier leurs possibles fonctions alternatives (pressenties par la communauté scientifique mais non encore élucidées), iii) mais surtout d'envisager une meilleurs compréhension de mutations corrélées à des pathologies humaines. Ces mutations n'étaient que au nombre de 15 en 2007, 65 sont répertoriées aujourd'hui. Il est vraisemblable que l'ensemble des aaRS mt soit impacté.

Les maladies mitochondriales : une nouvelle famille de protéines incriminée.

Hagen SCHWENZER et Marie SISSLER

Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire (IBMC), 15 rue René Descartes – 67084 Strasbourg cedex

M.Sissler@ibmc-cnrs.unistra.fr

Publié dans Regard sur Biochimi, 12/2012

Introduction

Les mitochondries occupent une place importante dans le métabolisme cellulaire énergétique. Elles interviennent dans des fonctions aussi primordiales que la respiration, la production d'énergie sous forme d'ATP, ou la biosynthèse de métabolites. Ainsi, un dysfonctionnement mitochondrial (mt) a généralement une répercussion importante sur le fonctionnement cellulaire, et peut être associé à diverses maladies humaines neurodégénératives et musculaires. Les mitochondries possèdent leur propre machinerie traductionnelle dédiée, chez l'homme, à la synthèse de seulement 13 protéines, toutes sous-unités des complexes de la chaîne respiratoire (figure 1). Alors que les 22 ARNt et 2 ARNr requis sont également codés par le génome mt, toutes les autres macromolécules nécessaires à la synthèse protéique sont codées par le génome nucléaire, synthétisées dans le cytosol et importées dans les mitochondries grâce à une séquence d'adressage spécifique. Parmi les protéines importées, les aminoacyl-ARNt synthétases (aaRS) ont un rôle clé dans le processus de traduction car elles sont responsables de l'attachement spécifique de chacun des acides aminés sur l'ARNt correspondant, qui les transfère à la chaine protéique naissante au niveau du ribosome. L'importance cruciale des ARNt et aaRS dans un processus cellulaire vital fait que leur implication possible dans des pathologies humaines n'a été suspectée que tardivement.

Ce n'est en effet que en 1989 qu'une première mutation dans un gène d'ARNt mt a été décrite et formellement corrélée à une maladie. Vingt ans plus tard, > 130 mutations touchant les 22 ARNt sont répertoriées dans la banque de référence Mitomap (http://www.mitomap.org/MITOMAP) et de nouvelles mutations sont fréquemment identifiées. Les phénotypes cliniques associés sont extrêmement variés, allant de désordres « légers » à apparition tardive, telles que des déficiences auditives neuro-

sensibles liées à l'âge, des myopathies oculaires, des diabètes, … à des désordres dévastateurs et souvent fatals dès l'enfance, tel que par exemple le syndrome de Leigh. Bien plus tardivement, en 2007, une première série de mutations a été découverte dans les gènes nucléaires codant pour l'aspartyl-ARNt synthétase (DARS2, enzyme spécifique de l'acide aspartique) mt de patients atteints d'un syndrome décrit sous l'acronyme LSBSL pour « Leukoencephalopathy with brain stem and spinal cord involvment and lactacte elevation » (Scheper *et al.*, 2007). Cette première description fut, à son tour, prémisse d'une liste qui risque de s'avérer longue, puisque en 5 ans seulement, des mutations dans 5 autres gènes d'aaRS mt ont été décrites comme causatives de maladies mitochondriales.

Les faits

Ainsi, les premières mutations décrites concernent le gène nucléaire de l'aaRS mt spécifique de l'acide aspartique (DARS2, Scheper *et al.*, 2007; Lin *et al.*, 2010), d'autres ont rapidement été découvertes, répertoriées et validées. Elles concernent les enzymes spécifiques de l'arginine (RARS2, Edvardson *et al.*, 2007), de la tyrosine (YARS2, Riley *et al.*, 2010), de la sérine (SARS2, Belostotsky *et al.*, 2011), de l'histidine (HARS2, Pierce *et al.*, 2011), et très récemment, de l'alanine (AARS2, Götz *et al.*, 2011) (tableau 1). Une pathologie ne se manifeste que en cas de double récessivité de mutations de type homozygotes ou hétérozygotes. Les mutations sont principalement de type faux-sens, mais des défauts d'épissage des introns sont également répertoriés. Ces défauts sont toujours partiels, permettant une production résiduelle de l'enzyme de séquence sauvage.

La fonction primordiale et connue des aaRS mt est d'aminoacyler les ARNt homologues afin de contribuer à la synthèse de 13 des sous-unités de la chaîne respiratoire lors de la traduction mitochondriale, et de ce fait à la production d'ATP. Les désordres de la chaîne respiratoire mitochondriale sont ceux parmi les plus ardus à diagnostiquer, guérir, ou prévenir. Il est toutefois anticipé que toute mutation produisant une pathologie affecte chacune des étapes listées ci-dessus, et en particulier les activités des complexes I et IV, qui contiennent le plus grand nombre de protéines codées par le génome mt (le complexe II en est totalement dépourvu, et d'origine exclusivement nucléaire). L'effet primaire attendu est une diminution de la production de l'énergie cellulaire (ATP), et parmi les effets secondaires possibles, la production de

dérivés réactifs de l'oxygène (ROS, pour « reactif oxygen species), chimiquement très réactifs, et causes de stress oxydatif, de vieillissement ou d'apoptose cellulaire. Cette cascade d'évènements est en effet généralement observée dans les cas décrits dans la littérature (tableau 1), avec des activités d'aminoacylation réduites (mais non abolies), des synthèses protéiques mitochondriales impactées, et des activités de complexes de la chaîne respiratoire restreintes. Pourtant, le décryptage des maladies mitochondriales liées à des défauts des aaRS mt reste complexe du fait d'éléments et particularités à ce jour encore incompris.

Les incompréhensions

Des activités d'aminoacylation réduites ont été démontrées pour un sous-groupe de DARS2 recombinantes portant des mutations faux-sens corrélées au syndrome LBSL, or, aucune perturbation significative de la traduction mt n'a été observée dans des fibroblastes ou lymphoblastes (*i.e.* des cellules non-neuronales) dérivés de patients. De même aucune différence notable des taux des complexes de la chaîne respiratoire n'est démontrée, et aucun défaut relié aux activités mt ne sont détectés dans ces mêmes cellules (tableau 1). Ces observations font exception à la cascade d'évènements attendus. A noter toutefois qu'il est à ce jour impossible d'exclure que des différences soient observables dans les cellules neuronales de patients.

De plus, une particularité additionnelle provient de l'observation de manifestations symptomatiques extrêmement variables chez les divers patients, ainsi que des expressions phénotypiques tissues-spécifiques propres à chacun des syndromes décrits, et donc propres à chacune des aaRS mt impactées (figure 2). Les premiers cas décrits proposaient une généralisation vers une altération du système nerveux central (mutations de DARS2 et de RARS2), rapidement démentie par la description des cas cliniques suivants, où des organes aussi diversifiés que le cœur, l'oreille interne, les poumons ou encore le rein sont affectés. Cette disparité symptomatique et tissulaire est d'autant plus surprenante compte tenu de l'ubiquité de l'expression et de la fonction des aaRS.

Enfin, une autre incompréhension résulte de l'absence totale de corrélation entre les phénotypes observés dans les cas de mutations dans les gènes nucléaires d'aaRS mt et de ceux observés dans les cas de mutations dans les gènes mt d'ARNt de spécificité correspondante. Les deux molécules sont pourtant étroitement impliquées dans le

même processus cellulaire et devrait donc produire la même expression phénotypique, à la condition que seules les fonctions primaires de ces molécules sont impactées. Or, une absence de corrélation a clairement été dénotée pour l'ensemble des systèmes décrits à ce jour. Par exemple, les phénotypes observés chez les patients atteints de LBSL (mutations de DARS2) n'incluent pas d'altération du nerf optique ou de dysfonctionnement musculaire, comme cela est observé suite à des mutations dans le gène de l'ARNt^{Asp} mt (Seneca *et al.*, 2005; Bosley *et al.*, 2008).

Les éventualités

Les aaRS mt sont des enzymes ubiquitaires, impliquées dans un processus cellulaire essentiel. Or l'expression phénotypique de leurs mutations varie en terme de tissue-spécificité et de présentation clinique. Un premier paramètre qu'il est utile de rappeler est la variabilité en besoin énergétique des différents tissus. Toutefois cette vue « réductionniste » désignerait les muscles, le cerveau, le cœur, ou le foie comme les organes les plus sensibles aux mutations des molécules de la traduction mitochondriale. Ils sont certes fréquemment affectés, mais pas systématiquement, et d'autres tissus le sont également.

En cas de composition génétique hétérozygote ou en cas de défaut d'épissage partiel des introns, il est envisagé qu'une proportion d'enzymes de séquence sauvage soit produite. Cette proportion, même faible serait visiblement suffisante pour expliquer qu'aucune diminution de la synthèse protéique mitochondriale n'ait été observée dans les cellules de patients atteints de mutations de DARS2, ou pour expliquer l'aminoacylation normale des ARNt^{Arg} résiduels produits dans les fibroblastes de patients atteints de mutations de RARS2.

Enfin, la variabilité tissue-spécifique d'éléments modulateurs, de facteurs d'épissage, ou de chaperons pourrait affecter le taux d'aaRS actives dans les différents tissus, rendant certains tissus plus vulnérables à une activité réduite, en deçà d'un seuil critique, alors que l'activité adéquate serait maintenue ailleurs. De même il est possible que la variabilité du niveau de base d'expression des gènes eux même puisse expliquer une sensibilité tissulaire différente. L'évaluation par PCR quantitative des ARNm correspondants aux aaRS mt (données personnelles non publiées) dans 20 tissus humains différents indique en effet une grande variabilité d'expression. De manière surprenante, les taux d'ARNm les plus bas sont observés dans les tissus aux besoins

énergétiques les plus grands. Ces données suggèrent que la plus grande sensibilité du cerveau, du cœur et des muscles aux mutations des ARNt ou aaRS mt pourrait résulter de dysfonctionnements même légers d'aaRS présentes en quantité intrinsèquement limitantes.

Les perspectives

Des observations similaires ont été décrites lors de l'étude des effets pathogéniques de mutations dans les gènes mt d'ARNt (revue par Florentz *et al.*, 2003). De fortes diminutions des taux de synthèses protéiques mitochondriales ont certes été observées dans la majorité des cas, mais de nombreuses exceptions ont été également répertoriées. Il a été démontré par ailleurs que, en fonction de la mutation analysée, chacune des étapes du cycle de vie de l'ARNt pouvait être affectée, démontrant la complexité de la relation génotype/phénotype. Dès lors, l'existence de facteurs additionnels ou la possibilité de mécanismes alternatifs (ou additifs) impliquant les ARNt ont été considérées. Leurs altérations par les mutations pourrait contribuer à l'étiologie de la maladie. Aujourd'hui, une hypothèse semblable d'implication des aaRS mt dans une ou plusieurs activités non-canonique(s) (différente de celle connue d'aminoacylation) est proposée. Un premier argument en faveur de cette hypothèse est l'absence dans certains cas d'une connexion claire entre l'expression moléculaire de la mutation et l'expression phénotypique de la maladie. Souvent, la relation génotype/phénotype ne peut s'expliquer par la seule possibilité d'une perte de fonction ou d'un gain d'une fonction toxique, mais éventuellement par l'altération d'une fonction secondaire, non canonique. Cette hypothèse est soutenue par les découvertes relativement récentes de nombreuses fonctions alternatives des aaRS humaines à localisation cytosolique (revue par Antonellis & Green, 2008; Park *et al.*, 2008). Ces fonctions incluent une large diversité d'activités, allant du contrôle transcriptionnel à la signalisation angiogénique. Ainsi, de telles fonctions alternatives doivent être recherchées pour les aaRS à localisation mitochondriale. On pourrait envisager qu'elles soient tissues-spécifiques. Ainsi et par exemple, l'identification de protéines qui s'associeraient spécifiquement à une ou plusieurs aaRS mt, de manière tissue-spécifique mais traduction-indépendante, suggèrerait très fortement la présence d'une fonction non-canonique importante.

Ainsi, l'incrimination des aaRS mt dans les pathologie dites « mitochondriales » est un domaine de recherche qui n'est aujourd'hui que à un stade naissant, mais les découvertes futures révèleront indéniablement une implication plus complexe et fascinante des aaRS dans les maladies génétiques humaines.

Légendes Tableau/Figures

	Type de mutation	Composition génétique	Activité (amino acylation)	Synthèse protéique mt	Impact(s) moléculaire(s)	Impact(s) sur les complexes de la CR	Manifestations pathologiques
DARS2 Aspartyl-ARNt synthetase (Scheper *et al.* 2007)	• non-sens / faux-sens • défaut d'épissage (partiel)[a]	hétérozygote	réduite	normale	activité, importation dans la mitochondrie [b]	non observée (dans fibroblastes et lymphoblastes)	**LBSL** "Leukoencephalopathy with Brain stem and Spinal cord involvement and Lactate elevation"
RARS2 Arginyl-ARNt synthetase (Edvardson *et al.* 2007)	faux-sens / défaut d'épissage (partiel)[a]	homozygote	normale (sur $ARNt^{Arg}$ résiduel)	impactée	enzyme tronquée, réduction du taux d'$ARNt^{Arg}$ mt [c]	I,III,IV	**PCH** "Severe Infantile encephalopathy associated with Pontocerebellar Hypoplasia"
YARS2 Tyrosyl-ARNt synthetase (Riley *et al.* 2010)	faux-sens	homozygote	réduite	impactée	expression de l'enzyme, reconnaissance de l'ARNt	I,III,IV,V (production de ROS, stress oxydatif)	**Syndrome MLASA** "Myopathy, Lactic Acidosis, and Sideroblastic Anemia"
SARS2 Seryl-ARNt synthetase (Belostotsky *et al.* 2011)	faux-sens	homozygote	réduite (seulement pour $ARNt^{Ser}$AGY)	ND	stabilité de l'ARNt, discrimination de l'ARNt	I,III,IV	**Syndrome HUPRA** "Hyperuricemia, Pulmonary Hypertension, Renal failure in Infancy and Alkalosis"
HARS2 Histidyl-ARNt synthetase (Pierce *et al.* 2011)	• faux-sens / défaut d'épissage • faux-sens	hétérozygote	réduite	impactée	activation de l'acide aminé, reconnaissance de l'adénylate	ND (production de ROS, initiation de l'apoptose)	**Syndrome de Perrault** "Ovarian dysgenesis and Sensorineural hearing loss"
AARS2 Alanyl-ARNt synthetase (Götz *et al.* 2011)	• faux-sens • faux-sens	homozygote ou hétérozygote	réduite	ND	activité, reconnaissance de l'ARNt	I,III,IV	**CMP** "Infantile Mitochondrial Cardiomyopathy"

Tableau 1. Aminoacyl-ARNt synthétases mitochondriales humaines et pathologies

ND, non déterminé(e). CR, chaîne respiratoire. [a], le défaut d'épissage des introns étant partiel, une faible quantité d'enzyme de séquence sauvage est produite. [b], le défaut d'importation a été démontré pour une mutation située dans la séquence d'adressage vers la mitochondrie (Messmer *et al.*, 2011). [c], indique une connexion entre les taux d'expression de RARS2 et de l'ARNt correspondant.

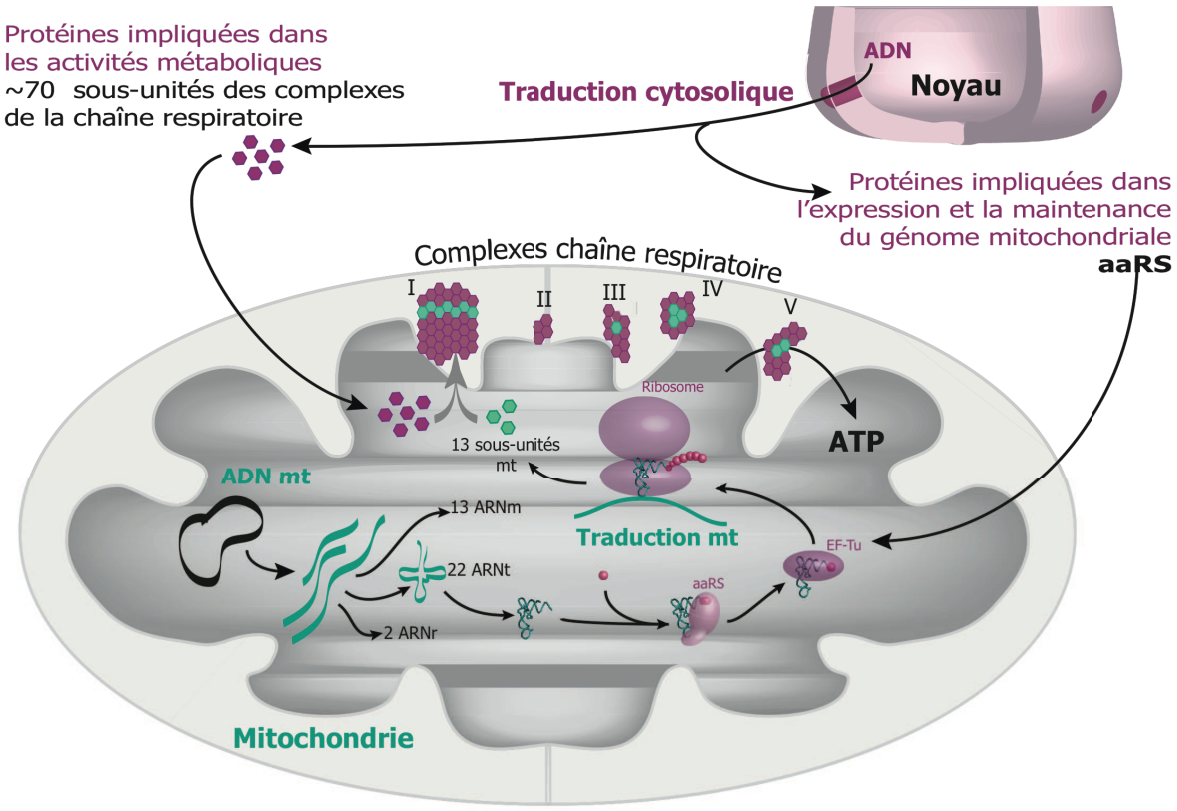


Figure 1. Origine double des molécules de la machinerie de traduction mitochondriale humaine. Le génome mt (ADN mt) code pour 22 ARNt, 2 ARNr et 13 ARNm (en vert), traduit en 13 des sous-unités des complexes de la chaîne respiratoire (hexagones verts). Toutes les autres molécules (en violet) requises à la fonctionnalité de la mitochondrie (impliquées dans les nombreuses voies métaboliques, le maintien et l'expression du génome mt) sont codées par le génome nucléaire, synthétisées dans le cytosol et importées dans la mitochondrie grâce à une séquence d'adressage spécifique. Figure adaptée de Florentz *et al.*, 2003 par Agnès Gaudry.

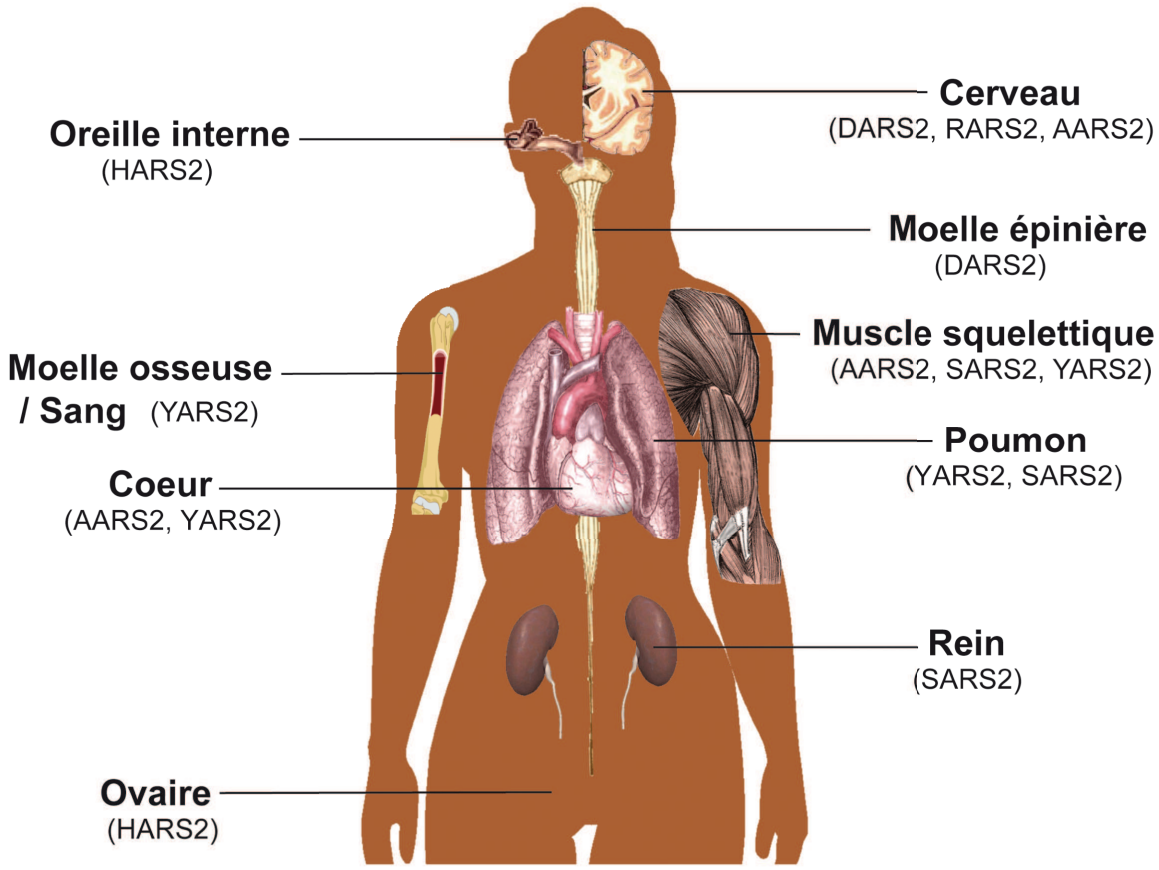


Figure 2. Manifestations symptomatiques variables et expressions phénotypiques tissues-spécifiques propres à chacune des aaRS mt humaines impactées

Références bibliographiques

Antonellis A & Green ED (**2008**) The role of aminoacyl-tRNA synthetases in genetic diseases. *Annu. Rev. Genom. Human Genet*. **9**: 87-107.

Belostotsky R *et al.* (**2011**) Mutations in the mitochondrial seryl-tRNA synthetase cause hyperuricemia, pulmonary hypertension, renal failure in infancy and alkalosis, HUPRA syndrome. *Am. J. Hum. Genet*. **88**: 193-200.

Bosler T *et al.* (**2008**) Sporadic bilateral optic neuropathy in children: the role of mitochondrial abnormalities. *Invest. Ophthal. and Visual Science* **49**: 5250-5256.

Edvardson S *et al.* (**2007**) Deleterious mutation in the mitochondrial arginyl-tRNA synthetase gene is associated with pontocerebellar hypoplasia. *Am. J. Hum. Genet.* **81**: 857-862.

Florentz C *et al.* (**2003**) Human mitochondrial tRNAs in health and disease. *Cell. Mol. Life Sci.* **60**: 1356-1375.

Götz A *et al.* (**2011**) Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. *Am. J. Hum. Genet.* **88**: 635-642.

Lin J *et al.* (**2010**) Leukoencephalopathy with Brain stem and Spinal Cord Involvement and Normal Lactate: a new mutation in the *DARS2* gene. *J. Child Neurol.* **25**: 1425-1428.

Messmer M *et al.* (**2011**) A human pathology-related mutation prevents import of an aminoacyl-tRNA synthetase into mitochondria. *Biochem. J.* **433**: 441-446.

Park SG *et al.* (**2008**) Aminoacyl-tRNA synthetases and their connections to disease. *Proc. Natl. Acad. Sci. USA* **105**: 11043-11049.

Pierce SB *et al.* (**2011**) Mutations in mitochondrial histidyl-tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. *Proc. Natl. Acad. Sci. USA* **108**: 6543-6548.

Riley LG *et al.* (**2010**) Mutation of the mitochondrial tyrosyl-tRNA synthetase gene, *YARS2*, causes myopathy, lactic acidosis, and sideroblastic anemia: MLASA syndrome. *Am. J. Hum. Genet.* **87**: 52-59.

Scheper GC *et al.* (**2007**) Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. *Nat. Genet.* **39**: 534-539.

Seneca S *et al.* (**2005**) A mitochondrial tRNA aspartate mutation causing isolated mitochondrial myopathy. *Am. J. Med. Genet. A*. **137**: 170-175.

Résumé

Les aminoacyl-ARNt synthetases (aaRS) sont des enzymes ancillaires, impliquées dans le mécanismes de la traduction. Elles catalysent l'aminoacylation spécifique de chaque ARNt pour l'acide aminé correspondant. Dans les cellules humaines, il existe deux jeux de gènes nucléaires codant pour les aaRS : un pour les aaRS cytosolique, le second pour les aaRS mitochondriales (mt). Les deux diffèrent par leurs séquences, et par l'absence ou la présence d'une séquence d'adressage mitochondriale (MTS). Les aaRS mt sont traduites dans le cytosole, adressées et importées dans la mitochondrie où elle sont maturées (clivage du MTS), structurées, avant d'atteindre une localisation sous-mitochondriale finale.

Les mitochondries possèdent une machinerie de traduction propre, et sont des acteurs clés dans le métabolisme énergétique cellulaire. Ainsi, tout disfonctionnement des mitochondries, ou de la traduction mitochondriale, peut–être délétère pour la cellule, et associé à des pathologies aux conséquences souvent sévères. Récemment, le laboratoire d'accueil a contribué à l'identification de mutations dans le gène nucléaire codant pour l'AspRS mt (enzyme spécifique de l'acide aspartique) comme cause de leucoencéphalopathies. Suite à cette première découverte, des mutations dans 8 autres gènes d'aaRS mt ont été démontrées comme responsables de pathologies. Certaines des mutations répertoriées induisent des pertes de l'activité de l'enzyme, d'autres n'affectent ni la propriété originelle d'aminoacylation, ni la traduction mitochondriale (revu dans les chapitres de livre 1 et 2). En conséquence, il a été proposé que certaines de ces mutations puissent affecter des propriétés alternatives, non encore décryptées, au delà de l'implication dans la traduction (hypothèse « d'un lien manquant »).

Des fonctions alternatives ont été décrites pour de nombreuses aaRS cytosoliques, et sont induites par e.g. une organisation cellulaire alternative, un partenariat différent, ou une re-localisation cellulaire. Alors que l'organisation des aaRS cytosolique est bien étudiée chez l'homme et que des implications dans des fonctions alternatives établies pour certaines d'entre elles, les connaissances quant aux aaRS mt restent parcimonieuses. L'objectif principal de ce manuscrit de thèse est triple: (i) révéler de nouvelles particularités de l'AspRS mt et intégrer cette enzyme dans un réseau fonctionnelle (organisation/localisation sous-mitochondriale) ; (ii) étendre l'analyse de l'organisation sous-mitochondriale (localisation/répartition) à l'ensemble des aaRS mt ; et (iii) contribuer à la compréhension de mécanismes moléculaires sous-jacents à certaines pathologies. Ce manuscrit est divisé en 5 chapitres.

Dans le 1er chapitre, nous montrons la perte de pression de sélection et l'existence d'un variant d'épissage qui affectent tout deux un unique domaine structurale de l'AspRS mt. Nous proposons un relâchement fonctionnel de ce domaine pour l'enzyme mitochondriale (Article 1).

Dans le 2^{ième} chapitre, nous étudions l'organisation sous-mitochondriale de l'ensemble des aaRS mt. La plupart ont une double distribution, dans la matrice et dans les membranes ; certaines ne se trouvent que dans une seule de ces fractions. De plus, nous analysons le mécanisme de répartition sous-mitochondriale et apportons des indices supplémentaires indiquant que la traduction mitochondriale se fait à proximité des membranes.

Dans le 3^{ième} chapitre, nos établissons et validons un outil robuste, qui permet d'explorer l'impact de séquences précurseurs sur le processus d'importation mitochondrial (Article 2). Cet outil a été utilisé pour identifier la séquence N-terminale de l'AspRS mt native, après maturation. Des comparaisons *in silico* des points isoélectriques des protéines précurseurs et maturées étayent l'hypothèse que les aaRS mt puissent avoir différentes localisations et/ou différents partenaires.

Dans le 4^{ième} chapitre, nous proposons une première liste de partenaires possibles de l'AspRS mt et des évidences quant à l'implication de cette enzyme dans un complexe de haut poids moléculaire.

Dans le 5^{ième} chapitre, nous initions l'analyse *de novo* de mutations dans les gènes d'ARNt ou d'aaRS mt, corrélées à des pathologies. De plus, nous apportons un regard complémentaire aux mutations de l'AspRS mt, déjà étudiées pour leur impact moléculaire, en termes d'effets sur les propriétés physico-chimiques et de localisation sous-mitochondriales. Nous montrons pour la première fois qu'une mutation peut empêcher l'import d'une aaRS mt (Article 3) et que le seul remplacement d'un acide aminé peut impacter la solubilité et les propriétés structure de l'enzyme entière (Article 4).

En substance, ces travaux contribuent à décrypter certains aspects de 'mitochondriology'. Ils ouvrent la porte vers d'autres investigations qui seront nécessaires à l'obtention de la vue globalisante de l'organisation des aaRS à l'intérieur de la mitochondrie. Ces contributions seront utiles à la

meilleure compréhension de mécanismes moléculaires sous-jacents à certaines pathologies mitochondriales.

Mots-clés: traduction mitochondriale, aminoacyl-ARNt synthétases, fonctions alternatives, organisation sous-mitochondriale, import, mutations reliées à des pathologies, évolution.

Abstract

Aminoacyl-tRNA synthetases (aaRSs) are housekeeping enzymes involved in translation. They catalyze the aminoacylation of tRNAs with their cognate amino acids. In human cells, two different sets of nuclear genes code for aaRSs. One codes for the aaRSs of cytosolic location, and the second one codes for aaRSs of mitochondrial (mt) location. Both differ in their sequences and mt-aaRS harbor in addition a mitochondrial targeting sequence (MTS). Mt-aaRSs are translated in the cytosol, targeted and imported into mitochondria where they are maturated by removal of the MTS and reach their final sub-mitochondrial localization.

Mitochondria, which possess own translation machinery, play an important role in cellular energetic metabolism. Any dysfunction of mitochondria or mt translation is usually deleterious for the cell, and has been associated with a growing number of disorders. Recently, the host laboratory contributed to the discovery that mutation(s) in nuclear gene coding for mt-aaRS (mt-AspRS, the enzyme specific for the aspartic acid) are the cause of some mt-diseases. This contribution led to the discovery of eight additional cases, where mutations in mt-aaRSs are responsible for diseases. Most of the reported mutations are leading to loss of enzyme activity, but some of them display no significant influence on the housekeeping aminoacylation activity and in consequence do not affect translation (reviewed in book chapters 1 and 2). Accordingly, it has been proposed within the scientific community that those mutations might affect alternative functions or properties of the proteins, beyond translation (the "missing link" hypothesis).

Alternate functions have already been described for several aaRSs of cytosolic location, and were shown to be due to alternate cellular organization, alternate partnerships or localization. While the organization of human cytosolic aaRSs is well explored and their involvement into alternate functions clearly established for some of them, the properties of the human mt-aaRSs remain scarcely known. The purpose of this thesis was double tracked. On one site, it should reveal new peculiarities of mt-AspRS and integrate it into new functional networks (sub-mitochondrial localization and partnership). On the other site, this thesis should expand the view of the sub-mitochondrial organization (location and repartition) to the full set of mt-aaRSs and should ultimately shed light into the molecular mechanisms underlying some of the pathologies. This thesis is divided in 5 chapters.

In the first chapter we showed that an extensive sequence divergence and alternative-splicing event are concentrated on a structural well-defined domain in mt-AspRS. We conclude that this domain undergo a functional relaxation (Article 1).

In the second chapter, we investigated the sub-mitochondrial organization of the full set of mt-aaRSs. We showed that some are double distributed in membrane and matrix while others were found only in one fraction. In addition, we analyzed the mechanism leading to sub-mitochondrial localization of mt- aaRSs and presented further proofs of membrane -associated mitochondrial translation machinery.

In the third chapter, we established and validated a unique and robust tool to enable exploration of the impact of precursor sequences on the mitochondrial import process (Article 2). This tool was used to decipher the N-terminal sequence of the mt-AspRS. *In silico* analysis of the pI of precursor proteins and matured proteins revealed peculiarities suggesting a different localization or different protein partners of the mt-aaRSs.

In the fourth chapter, we established a first list of potential partner proteins (for mt-AspRS, at first) and showed evidences for a possible macromolecular complex containing mt-AspRS.

In the last chapter, we initiate the *de novo* analysis of some discovered pathology-related mutations in tRNA and mt-aaRS genes. In addition, we re-analyze known and characterized mutants of mt-AspRS in regard to possible physical-chemical defects and their impact on their sub-mitochondrial organization. We showed the first time that a mutation can hamper the import of a mt-aaRSs (article 3) and that a single amino acid change can impact the solubility and structure of whole protein (article 4).

Altogether, these results contribute to decipher some aspects of "mitochondrial research". They also open the door for additional investigations to gain a complete view about the sub-mitochondrial organization of aaRSs. Those contributions will be of help for the understanding of molecular mechanisms underlying some mitochondrial disorders.

Keywords: mitochondrial translation machinery, aminoacyl-tRNA synthetase, non-translational functions, sub-mitochondrial organization, import, pathology-related mutations, evolution