



HAL
open science

Ingénierie d'un outil basé sur une GFP fragmentée pour l'étude des protéines multi-localisées chez les eucaryotes

Gaëtan Bader

► To cite this version:

Gaëtan Bader. Ingénierie d'un outil basé sur une GFP fragmentée pour l'étude des protéines multi-localisées chez les eucaryotes. Génomique, Transcriptomique et Protéomique [q-bio.GN]. Université de Strasbourg, 2017. Français. NNT: 2017STRAJ116 . tel-02003385

HAL Id: tel-02003385

<https://theses.hal.science/tel-02003385>

Submitted on 1 Feb 2019

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.

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ (ED414)

UMR7156 : Génétique Moléculaire Génomique Microbiologie

THÈSE présentée par :

Gaétan BADER

soutenue le : **15 Décembre 2017**

Pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline : Sciences du Vivant

Spécialité : Aspects moléculaire et cellulaires de la biologie

**Ingénierie d'un outil basé sur une GFP
fragmentée pour l'étude des protéines multi-
localisées chez les eucaryotes**

THÈSE dirigée par :

M. SENGER Bruno
M. BECKER Hubert D.

Docteur, Institut botanique (UMR7156)
Professeur, Institut botanique (UMR7156)

RAPPORTEURS :

M. MARTINOU Jean-Claude
Mme. BONNEFOY Nathalie

Professeur, Université de Genève
Docteur, Institut de Biologie Intégrative de la Cellule

AUTRES MEMBRES DU JURY :

M. LUQUE ROMERO Ignacio
Mme. DUCHENE Anne-Marie

Docteur, Instituto de Bioquímica Vegetal y Fotosíntesis
Docteur, Institut de biologie moléculaire des plantes

ABREVIATIONS	1
INTRODUCTION	9
I. Généralités	11
II. Le code génétique : théorie et évolution	11
III. La synthèse protéique	15
1. Le ribosome	17
2. L'initiation de la traduction	17
3. L'élongation de la traduction	20
4. La terminaison de la traduction.....	22
5. Le recyclage du ribosome.....	23
IV. Les ARN de transfert.....	25
1. La structure des ARNt	25
2. La maturation des pré-ARNt.....	27
3. L'export des pré-ARNt et l'épissage des introns	29
4. Les fonctions additionnelles des ARNt et des fragments d'ARNt	30
V. Les aminoacyl-ARNt synthétases	33
1. La fonction primaire des aaRS	33
2. Les classes d'aaRS	35
3. L'incorporation des 21 ^{ème} et 22 ^{ème} aa	36
4. La spécificité des aaRS.....	37
5. La mésincorporation des aa	38
6. Les voies indirectes de formation des aa-ARNt.....	40
7. Les fonctions additionnelles des aaRS	43
a. Les fonctions additionnelles des aaRS bactériennes.....	43
b. Les fonctions additionnelles des aaRS de champignons	43
c. Les fonctions additionnelles des aaRS de mammifères	44
8. La localisation subcellulaire des aaRS cytosoliques.....	45
VI. Les complexes multi-synthétasiques.....	48
1. Les domaines d'interaction retrouvés chez les aaRS.....	52
a. Les domaines d'interaction protéine/protéine	52
b. Les domaines d'interaction protéine/ARN	53
2. Les MSC chez les unicellulaires.....	54
a. Les MSC chez les bactéries	54
b. Les MSC chez les archées	54
c. Le MSC chez <i>Saccharomyces cerevisiae</i>	57
d. Le MSC chez <i>Toxoplasma gondii</i>	59

e. Le MSC chez <i>Trypanosoma brucei</i>	59
3. Les MSC des eucaryotes multicellulaires	61
a. Le MSC chez les mammifères	61
b. Le MSC chez le nématode	62
c. Le MSC chez les arthropodes.....	63
4. Les fonctions non canoniques des aaRS relâchées du MSC.....	63
a. Les fonctions additionnelles de la MetRS humaine	64
b. Les fonctions additionnelles de la GluProRS humaine.....	64
c. Les fonctions additionnelles des autres aaRS du MSC	67
5. Les fonctions non canoniques des AIMP	68
VII. Les mitochondries	70
1. La découverte et la structure des mitochondries	70
2. Les origines des mitochondries	71
3. Le génome et le contenu protéique des mitochondries.....	74
4. Les fonctions mitochondriales.....	76
5. Les mécanismes d'import des protéines mitochondriales	79
a. L'import des protéines contenant des préséquences en N-ter	79
b. L'import des protéines possédant des signaux d'adressage internes.....	83
6. Le contact entre la mitochondrie et les autres organelles	84
VIII. La Green Fluorescent Protein et ses applications.....	87
1. La découverte de la GFP	87
2. La structure et les caractéristiques de la GFP	89
3. La palette de variants de la GFP	93
a. Le super folder GFP.....	93
b. La <i>Cyan Fluorescent Protein</i>	94
c. La <i>Yellow Fluorescent Protein</i>	95
4. L'utilisation de la GFP et de ses variants.....	96
a. Le marquage de protéines.....	96
b. L'étude du photoblanchiment	97
c. L'étude de la dynamique cellulaire	98
d. Le marquage de cellules, tissus ou organes	98
e. Le marquage de l'ARN	99
f. L'étude de l'interaction protéine/protéine	103
g. L'étude de dynamique cellulaire par la Split-GFP.....	105
PROJET DE THESE.....	110
MATERIELS ET METHODES.....	116
I. Matériel biologique et milieux de culture	118

1. Souches bactériennes	118
2. Milieux de culture bactériens.....	118
3. Souches de levure	119
4. Milieux de culture levure	119
II. Techniques biochimiques utilisées.....	120
1. Préparations de bactéries chimio-compétentes	120
2. Transformations de bactéries.....	121
3. Préparation de plasmides	121
4. Restriction enzymatique	121
5. Préparation d'ADN génomique de levure	122
III. Amplification et clonage des gènes d'intérêts.....	122
1. Design des amorces	122
2. Amplification PCR et purification.....	123
3. Amplification PCR par le protocole <i>Phire Plant direct PCR</i>	123
4. Clonage des produits PCR	124
5. Sous-clonage classique.....	124
6. Sous-clonage par Gateway (kit invitrogen)	125
7. Sous-clonage par la technique de Gibson	125
IV. Utilisation de l'outil levure	126
1. Transformation de levures	126
2. Chasse du plasmide au 5FOA	127
3. Sporulation et dissections de tétrades	127
4. Tests en gouttes.....	127
V. Techniques de biologies cellulaires utilisées	128
1. Purifications de mitochondries.....	128
2. Fractionnements subcellulaires.....	128
3. Visualisations microscopiques.....	128
a. Cultures de levures	128
b. Etudes microscopiques et paramètres utilisés.....	128
c. Sondes fluorescentes utilisées	130
d. Marquage des compartiments.....	130
VI. Obtention et études des extraits protéiques	131
1. Préparations d'extraits protéiques totaux	131
2. Immuno-détections de protéines	131
RESULTATS ET DISCUSSIONS	134
PARTIE I : DYNAMIQUE DU COMPLEXE AME DANS LA REGULATION DE LA TRANSITION DIAUXIQUE CHEZ LA	
LEVURE <i>S. CEREVISIAE</i>	136

1. Contexte de la recherche et participation	138
2. Article 1 : Expression of nuclear and mitochondrial genes encoding ATP synthase is synchronized by disassembly of a multisynthetase complex	141
PARTIE II : LA SPLIT-GFP, UN OUTIL POUR L'ETUDE DES PROTEINES A LOCALISATIONS SUBCELLULAIRES	
MULTIPLES CHEZ <i>S. CEREVISIAE</i>.....	158
1. Problématique du projet.....	160
2. Ingénierie de l'outil de localisation basé sur la Split-GFP	161
3. Fonctionnalité des protéines étiquetées β 11	167
4. Autoassemblage des fragments β 1-10 et β 11	171
5. Autoassemblage mitochondrial des fragments β 1-10 et β 11.....	172
6. Reconstitution mitochondriale d'une Split-CFP et -YFP	178
7. Autoassemblage nucléaire des fragments β 1-10 et β 11	180
8. Reconstitution de la CFP à la surface de la vacuole	183
9. Reconstitution nucléaire et vacuolaire de la Split-GFP et -CFP.....	186
10. Localisation nucléaire et vacuolaire de NES- β 11.....	188
11. Expression du fragment β 1-10 à partir du génome mitochondrial.....	192
PARTIE III : APPLICATIONS DE L'OUTIL DE SPLIT-GFP POUR L'ETUDE DES PROTEINES A LOCALISATIONS SUBCELLULAIRES MULTIPLES CHEZ <i>S. CEREVISIAE</i>.....	202
1. Contexte du projet.....	204
2. Amplification et clonage des gènes d'aaRS.....	205
3. Etude de la relocalisation mitochondriale des aaRS cytosoliques, d'AIM10, la ProRS mitochondriale putative et de la protéine d'ancrage du complexe AME, Arc1	207
4. Etude des signaux d'import.....	216
a. Caractérisation du signal d'import de la GluRS	219
b. Caractérisation de la région contenant le signal d'import d'Arc1.....	221
c. Caractérisation de la région impliquée dans l'import de GatF	224
5. Etude de la relocalisation mitochondriale de protéines hétérologues	225
PARTIE IV : ETUDE DE LA GLUPRORS^{N1} HUMAINE ET DE SA FONCTION MITOCHONDRIALE	231
1. Contexte de la recherche	233
2. Résultats préliminaires.....	237
a. Etude de la GluRS humaine	237
b. Etude de la GluProRS ^{N1} dans des cellules Hela.....	239
3. Caractérisation de la fonction de la GluProRS ^{N1} chez la levure	241
4. Validation de la fonction mitochondriale de la GluProRS ^{N1} chez la levure	245
5. Relocalisation mitochondriale de la GluProRS ^{N1} avec l'outil Split-GFP	259
CONCLUSIONS ET PERSPECTIVES GENERALES	261

ANNEXE 1	268
ANNEXE 2	282
BIBLIOGRAPHIE	298

Abréviations

Symboles

β , α : bêta, alpha

3D : tridimensionnel

5FOA : *5-fluoro-orotique acid*

A

aa : acide aminé

aa-AMP : aminoacyl-adénylate

aa-ARNt^{aa} : aminoacyl-ARNt^{acide aminé}

aaRS : aminoacyl-ARNt synthétase

AD : *Activating domain*

ADN : acide désoxyribonucléique

ADP : adénosine diphosphate

AdT : amidotransférase

AIMP : *Aminoacyl tRNA synthetase complex Interacting Multifunctional Protein*

AME : **Arc1-MRS-ERS**

AMP : adénosine monophosphate

ARN : acide ribonucléique

ARNm, nc, t ou r : ARN messenger, non codant, de transfert ou ribosomique

ARNsn : petit ARN non codant

ASK1 : *Apoptosis Signal-regulating Kinase*

ATP : adénosine tri-phosphate

A, U, G, C et I : adénine, uracile, guanine, cytosine et inosine

B

BiFC : *Bimolecular Fluorescence Complementation*

BFP : *Blue Fluorescent Protein*

C

Ca²⁺ : Calcium

CCase : ARNt nucléotidyl transférase

Cdk5 : *cyclin-dependent kinase 5*

CFP : *Cyan Fluorescent Protein*

Cl⁻ : ion chlorure

CoIP : co-immunoprécipitation

COX : *Cyt c oxydase*

Cp : *ceruloplasmin*

CP1 : *Connective Peptide 1*

CR : chaîne respiratoire

C-ter : C-terminal

Cyt : cytochromes

D

DBD : *DNA binding Domain*

DAPK : *Death-associated protein kinase*

DHFR : Dihydrofolate Réductase

E

eCFP : *enhanced CFP*

eEF : facteur d'élongation eucaryotique

eGFP : *enhanced Green Fluorescent Protein*

eIF : facteur d'initiation eucaryotique

ELR : Glu-Leu-Arg

EMAPII : *Endothelial Monocyte-Activating Polypeptide*

eRF : facteur de terminaison eucaryotique

ERMES : *ER-Mitochondria Encounter Structure*

ERV: *Essential for Respiration and Vegetative growth*

eYFP : *enhanced YFP*

F

FATP1 : *Fatty Acid Transporter Protein 1*

FCCS : *Fluorescence Cross-Correlation Spectroscopy*

FCS : *Fluorescence Correlation Spectroscopy*

FITC : Fluorescein isothiocyanate

FLIP : *Fluorescence Loss In Photobleaching*

FRAP : *Fluorescence Recovery After Photobleaching*

FRET : *Förster resonance energy transfer*

frGFP : *folding reporter GFP*

G

GAIT : *IFN- γ -activated inhibitor of translation*

GAP : *GTPase-activating protein*

GAPDH : glyceraldehyde-3-phosphate deshydrogenase

GDP : guanosine di-phosphate

GEF : guanine exchange factor

GFP : *Green Fluorescent Protein*

GST-like : *Glutathion S-transférase-like*

GTP : guanosine tri-phosphate

H

H₂O : eau

H₂O₂ : peroxyde d'hydrogène

HOT : *Helper Of Tim*

Hsp90 : *Heat shock protein*

I

IFN γ : interféron γ

IL : Interleukine

IMP : *Inner Membrane Peptidase*

IMS : *InterMembrane Space*

ITC : *Isothermal Titration Calorimetry*

IPTG : Isopropyl β -D-1-thiogalacto pyranoside

K

K_D : constante de dissociation

kDa : kilodalton

L

LZ : *Leucine zipper*

M

MAPK : *Mitogen Activated Protein kinase*

MARS : multi aminoacyl-tRNA synthetase complex

MAVS : *Mitochondrial Antiviral Signaling Protein*

MCS : *Membranes Contact Sites*

MCP : *MARS Complex-associated Proteins*

MDH : malate déshydrogénase

MDM2 : *Mouse Double Minute 2 homolog*

Mdm : *Mitochondrial Distribution and Morphology*

Met-ARNt^{Met}_i : Méthionyl- ARNt^{méthionine} initiateur

_mHsp 70 : *Mitochondrial Heat Shock Protein*

Mia : *Mitochondrial Intermembrane space import and Assembly*

MIM : *Mitochondrial Import*

MITF : *Micropthalmia associated Transcription Factor*

MME : membrane mitochondriale externe

MMI : membrane mitochondriale interne

Mmm : *Mitochondrial Morphology Maintenance*

mnm5s2U34 : 5-méthylaminométhyl-2-thioUridine34

MPP : *Mitochondrial Processing Peptidase*

mrsp-38 : *MARS Scaffold Protein 38*

MSC : *MultiSynthetase Complex*

MTS : *Mitochondrial Targeting Signal*

N

NADH : *Nicotinamide Adénine DiNucléotide* ND : non discriminante

NES : *Nuclear Export Signal*

N-hélice : Hélice N-ter

NSAP1 : *NS1-associated protein 1*

N-ter : N-terminal

O

OB-fold :

O₂⁻ : ions super oxydes

P

PABP : *poly-A binding protein*

PAM : *Presequence translocase-Associated Motor*

PARP : *poly ADP ribose polymerase*
PCA : *Protein complementation Assays*
PIC : *Pre Initiation Complex*
PI3P et PI3,5P₂ : phosphatidylinositol 3 phosphate et 3,5 bi-phosphate
PS : phosphatidylsérine
PSTK : O-phosphoséryl-ARNt^{Sec} kinase
PYLIS : pyrroLysin inserted sequence
Pyr : pyrrolysine

R

RanGAP : *Ran GTPase-activating protein*
RBP : *RNA-binding protein*
RE : réticulum endoplasmique
RIG-I : *Retinoic acid-Inducible Gene I*
ROS : *Reactiv Oxygen Species*

S

S : Svedberg
SAM : *Sorting and assembly machinery*
SAXS : *small-angle X-ray scattering*
SCM : Site de Clonage Multiple
SECIS : selenoCystein inserted sequence
Sel : sélénocystéine
sfGFP : *super folder GFP*
SGD : *Saccharomyces genome database*
s-u : sous-unité(s)

T

TIM : *Translocase of Inner Membrane*
TNF : *Tumor Necrosis Factor*
TOM : *Translocase of Outer Membrane*
TORC1 : *Target Of Rapamycin Complex 1* TRAF2 : *TNF Receptor–Associated Factor*
Trbd111 : *tRNA-binding domain 111*
tRF : *tRNA Fragments*
tRK1 : ARNt^{Lysine1} cytosolique

U

UNE : UNique Element
UQ : ubiquinone
UQH₂ : ubiquinol
UV : Ultraviolet

V

VC : Venus C-ter
vCLAMP : *vacuole and mitochondria patch*
VDAC : *Voltage Dependant Anions Channel*

VEGF-A : *vascular endothelial growth factor A*

VIH : *Virus de l'Immunodéficience Humaine*

VN : *Venus N-ter*

Vps39 : *Vacuolar protein sorting*

W

WHEP : *TrpRS (W), HisRS (H), GluRS (E) ProRS (P)*

Y

YFP : *Yellow Fluorescent Protein*

Y2H : *Yeast Two hybrid*

Ypt : *Yeast Protein Two*

Z

ZIPK : *Zipper interacting protein kinase*

Introduction

I. Généralités

Chez les eucaryotes, l'information génétique de chaque cellule est stockée sous forme d'ADN dans le noyau et les mitochondries. Il existe néanmoins un 3^{ème} compartiment possédant de l'ADN, le chloroplaste, présent chez les plantes et algues. Pour tout être vivant, le décodage de l'information génétique est une étape essentielle pour la croissance, la pérennité et l'évolution de l'organisme. En effet, en cas de défaut de ce décodage, la viabilité cellulaire n'est plus assurée. Au sein du génome, il existe un mélange de régions codantes et non codantes permettant, toutes deux, la régulation de divers processus biologiques, la synthèse de protéines et d'ARN non codants (ARNnc), tels que les ARN de transfert (ARNt) et ARN ribosomiques (ARNr). L'expression de l'information génétique requiert une première étape de transcription en ARN (ARN messager (ARNm) et ARNnc) se déroulant dans le noyau. Chez les eucaryotes, cette molécule est ensuite exportée de ce compartiment par des mécanismes spécifiques pour finalement servir de matrice à la synthèse de protéines par le ribosome dans le cytosol. L'ARNm est constitué d'une suite de nucléotides dont une partie seulement est réellement codante. En effet, la séquence codante de l'ARNm, composée d'une succession de codons (triplets de bases), est bordée par des régions non traduites qui favorisent notamment le recrutement du ribosome. Ce dernier va alors se positionner au niveau du codon d'initiation AUG et se déplacer tout au long de la séquence pour traduire la succession de codons en une succession d'acides aminés (aa) et ainsi synthétiser la protéine. La succession de codons est dictée par la séquence de l'ARNm et l'information est traduite en protéines de par l'utilisation d'un code génétique qui est propre à l'organisme ou à l'organelle où se déroule la synthèse protéique.

II. Le code génétique : théorie et évolution

L'information génétique transcrite sous forme d'ARNm à partir de l'ADN est décryptée grâce au code génétique qui permet de traduire un alphabet nucléotidique constitué de combinaisons de quatre lettres en un alphabet protéique constitué de 20 lettres correspondant aux aa. Lors de sa découverte, ce code a été considéré comme universel (Crick, 1968). Cependant, de nombreuses variations de l'assignation des codons ont été mises en évidence dans les codes génétiques des différents organismes vivants étudiés à ce jour, suggérant qu'il n'était pas figé lors de son apparition, mais qu'il a évolué au cours du temps.

		Seconde lettre				
		U	C	A	G	
Première lettre	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
		Leu ● ●	Ser ●	Stop Gln Leu Ala	Stop Sec Ser Thr	A
		Leu ●	Ser	Stop Gln Pyl	Trp	G
	C	Leu Thr	Pro	His	Arg	U
		Leu Thr	Pro	His	Arg	C
		Leu Thr	Pro	Gln	Arg	A
		Leu ● Ser Thr	Pro	Gln ●	Arg	G
	A	Ile ●	Thr	Asn	Ser	U
		Ile ●	Thr	Asn	Ser	C
		Ile ● Met	Thr	Lys Asn	Arg ● Ser Gly	A
		Met ●	Thr ●	Lys	Arg ● Ser Gly	G
	G	Val	Ala	Asp	Gly	U
		Val	Ala	Asp	Gly	C
		Val	Ala	Glu	Gly	A
		Val ●	Ala	Glu	Gly	G

Figure 1 : Le code génétique standard

Le code génétique standard est défini par 61 triplets codants et trois codons stop. Les aa correspondant aux codons sont nommés selon le code à trois lettres. Les cases surlignées en rouge correspondent aux codons stop du code génétique standard, tandis que les cases jaunes correspondent aux codons qui ne sont pas utilisés chez certains organismes. Les exceptions d'assignation des codons par rapport au code génétique standard sont représentées par le code à trois lettres lorsque l'aa décodé par le codon est différent, par un rond orange lorsque le codon sert de codon d'initiation et par un rond rouge lorsqu'il correspond à un codon stop.

L'établissement du code génétique standard dans les années 60 a ouvert les portes à l'étude de la conservation et de l'évolution du code génétique à travers le temps et les espèces.

Le code génétique est constitué de 64 combinaisons possibles de triplets de bases qui se déclinent, dans le code génétique standard, en 61 codons codants et trois codons stop (**Figure 1**). Les 61 codons codants peuvent être subdivisés en 20 groupes contenant un, deux, trois, quatre ou six codons correspondant aux 20 aa standards et protéinogéniques auxquels ils ont été assignés. Lorsqu'un aa est codé par deux, trois ou quatre codons, la variabilité, ou dégénérescence dans la séquence des codons synonymes, se situe exclusivement sur le 3^{ème} nucléotide du triplet. Cette dégénérescence s'étend à l'un des deux autres ou aux deux autres nucléotides du codon lorsque six codons permettent le décodage d'un aa. L'assignation des codons est donc clairement ordonnée et n'est pas issue du hasard. La conséquence de celle-ci est la présence d'aa dont les propriétés physico-chimiques sont similaires dans des colonnes juxtaposées dans le tableau du code génétique. Ainsi, tous les aa appartenant à la même voie de biosynthèse métabolique sont décodés par un codon débutant par la même première base. Par exemple, les codons commençant par G codent pour des aa considérés comme prébiotiques. En revanche, la nature de la base en deuxième position du codon est reliée aux propriétés d'hydrophobicité ou d'hydrophilie des aa codés par ces codons. A titre d'exemple, tous les codons possédant un U en seconde position codent pour les aa les plus hydrophobes du répertoire du code génétique (isoleucine : Ile, leucine : Leu, valine : Val, phénylalanine : Phe et méthionine : Met). De cette manière, la structure de ce code génétique standard permet de limiter l'impact d'une mutation d'une des bases du codon. Une mutation dans la 3^{ème} position est silencieuse pour huit des 20 aa du code génétique (Leu, Val, sérine : Ser, proline : Pro, thréonine : Thr, alanine : Ala, arginine : Arg et glycine : Gly) et conduit au remplacement de l'aa initial en un aa de même nature physico-chimique pour sept couples d'aa du code (Leu-Phe, Ile-Met, histidine-glutamine (His-Gln), asparagine-lysine (Asn-Lys), aspartate-glutamate (Asp-Glu), cystéine-tryptophane (Cys-Trp) et Arg-Ser). Cependant, le code génétique standard est bien loin d'être le plus robuste. En effet, des études quantitatives montrent qu'il existe un nombre très important de variants de ce code génétique (10^{84}), dont un nombre considérable est plus robuste que le code génétique standard (Koonin and Novozhilov, 2009). L'évolution du code génétique n'est donc probablement pas finie comme en atteste la découverte plus récemment des 21^{ème} et 22^{ème} aa codés génétiquement que sont

la sélénocystéine (Sel) (codée par UGA) et la pyrrolysine (Pyr) (codée par UAG) (voir introduction § V.3). Pourtant, il est difficile de prédire cette évolution car de nombreux facteurs physico-chimiques et environnementaux entrent en jeu.

Selon Crick, la création du code génétique est un exemple de *frozen accident*. Le code génétique serait apparu par un mécanisme aléatoire et une combinaison d'événements très improbables (Crick, 1968). Néanmoins, cette hypothèse semble très peu probable du fait de la complexité et des variations du code que l'on connaît aujourd'hui. L'hypothèse de Crick a, dès le commencement de l'étude du code génétique, suscité de nombreuses controverses et de nombreux modèles ont été proposés pour expliquer l'évolution du code génétique. En effet, selon Wong, les codons ont co-évolué avec les voies de synthèse des aa (Wong, 1975). Une autre théorie suggère que les codons ont évolué pour augmenter drastiquement et de manière maximale la fidélité de décodage de l'information génétique, permettant ainsi de minimiser l'effet de mutations ponctuelles ou de minimiser les erreurs d'incorporation lors de la synthèse protéique (Sonneborn, 1965; Woese, 1965). De plus, il est maintenant admis que les propriétés physico-chimiques de l'ADN et des aa sont directement liées à l'évolution des codons et à l'apparition d'un code génétique quasiment universel (Pelc, 1965). Finalement, il semblerait que la structure actuelle du code génétique standard tienne de la combinaison des trois hypothèses mentionnées précédemment : (i) celle du *frozen accident* (Crick, 1968), (ii) celle de l'assignation des codons pour minimiser les erreurs (Sonneborn, 1965) et (iii) celle de la coévolution des codons avec les différentes voies de biosynthèse des aa (Wong, 1975).

Comme je l'ai dit précédemment, de nombreuses déviations par rapport au code génétique standard ont été observées dans plusieurs groupes d'organismes et plus particulièrement dans les génomes organellaires, de parasites ou de bactéries endosymbiotiques (Ling et al., 2015; Santos et al., 2004; Sengupta and Higgs, 2015). Trois scénarios sont possibles pour expliquer ces variations :

- (i) La réassignation d'un codon à l'un des 20 autres groupes de codons restants (19 groupes d'aa restants et le groupe des codons stop)
- (ii) La perte d'un codon codant (conversion en un codon stop ou non assignation)
- (iii) L'assignation d'un codon à des nouveaux aa n'existant pas dans le code génétique standard.

Il est à noter que sur les 21 variations du code génétique mitochondrial décrites à ce jour, six d'entre-elles correspondent à la réassignation d'un codon stop, trois correspondent à la conversion d'un codon codant en codon stop et douze correspondent à des codons codants réassignés à un aa différent ou non assignés (Sengupta and Higgs, 2015). Les exemples les plus fréquents de la variation de l'utilisation des codons concernent (i) le codon UGA des eucaryotes qui correspond à un codon stop dans les gènes nucléaires et Trp dans les gènes organellaires ; les (ii) 21^{ème} (Sel) et (iii) 22^{ème} (Pyl) aa codés par le code génétique de divers organismes qui sont décodés respectivement par les codons stop UGA et UAG (Hao et al., 2002; Soll, 1988; Srinivasan et al., 2002).

A ces variations dans l'assignation des codons s'ajoute un autre aspect du code génétique : l'usage des codons. En effet, la variabilité de ce code est également basée sur l'utilisation préférentielle de certains codons synonymes plutôt que d'autres et est différente en fonction des organismes. Cet usage préférentiel des codons est la conséquence, d'une part, du contenu en paires de bases G/C et A/T de l'ADN génomique de l'organisme considéré, et d'autre part, de l'expression plus ou moins importante de certains ARNt au sein d'une famille d'ARNt isoaccepteurs. Ainsi, un codon faiblement utilisé est décodé par un ARNt dit "mineur" (faiblement transcrit), tandis qu'un codon fréquemment utilisé est décodé par un ARNt dit "majeur".

III. La synthèse protéique

La dernière grande étape de l'expression des gènes est la synthèse protéique. Du fait de son importance, ce processus est fortement régulé et fait intervenir de nombreuses protéines additionnelles, seules ou sous forme de complexes. La synthèse protéique est divisée en quatre grandes étapes successives : (i) l'initiation, (ii) l'élongation, (iii) la terminaison et (iv) le recyclage du ribosome. C'est un mécanisme cyclique et dès lors que le ribosome est recyclé, l'initiation redémarre. Les facteurs protéiques participants à ces quatre étapes sont nommés en fonction de l'étape durant laquelle ils interviennent. Ainsi, chez les eucaryotes, il y a participation des facteurs d'initiation eucaryotiques (eIF), des facteurs d'élongation (eEF) et des facteurs de terminaison (eRF). De plus, les étapes ribosomiques de la traduction d'un ARNm étant GTP-dépendantes, des GTPases (GAP) et des facteurs d'échange de guanosine (GEF) interviennent également. Malgré les nombreuses études qui

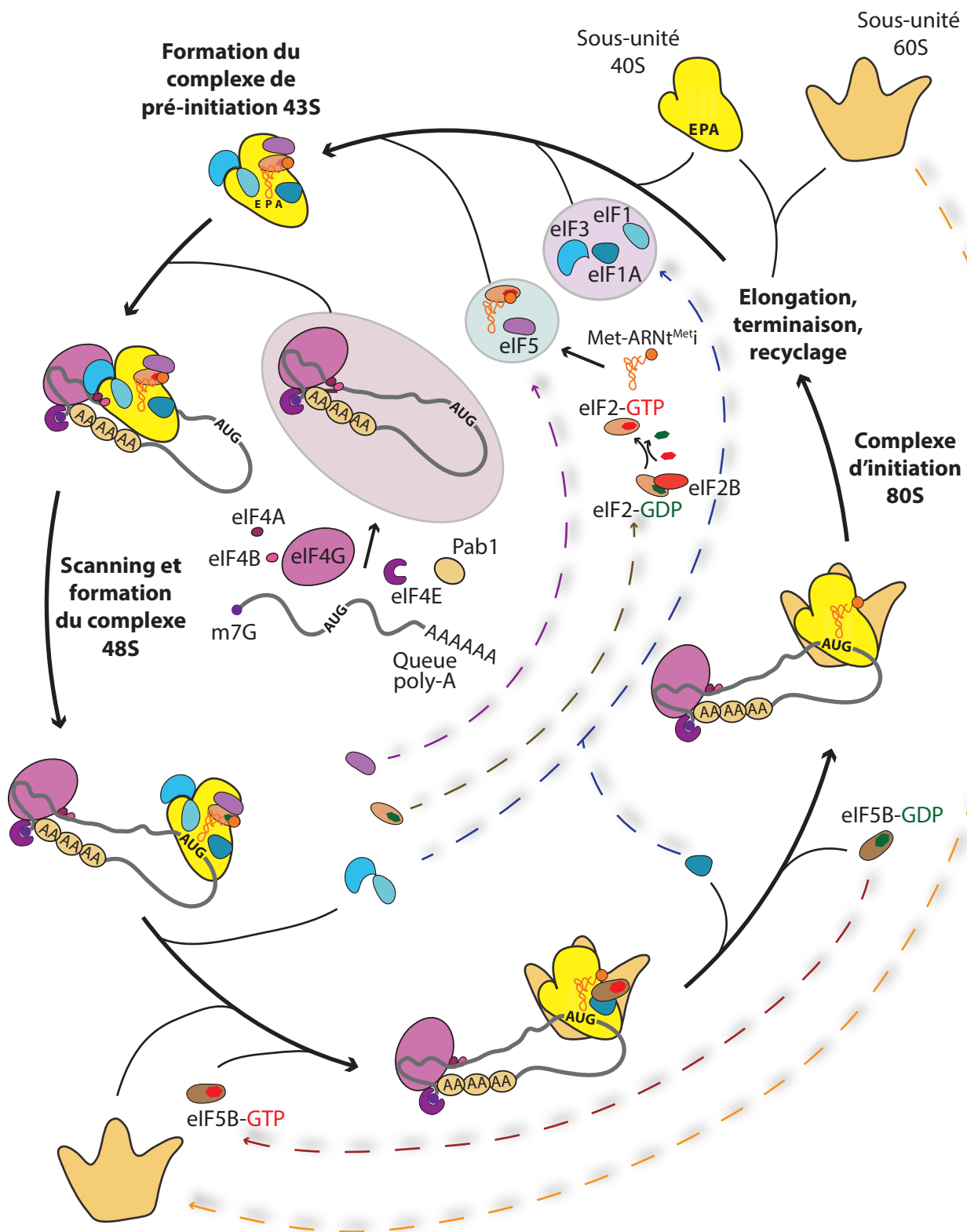


Figure 2 : L'initiation de la traduction chez les eucaryotes

La mise en place de la machinerie traductionnelle débute par la formation d'un PIC 43S qui va permettre le positionnement du Met-ARNt^{Met} sur le codon AUG initiateur par un mécanisme de *scanning*. Finalement, le ribosome 80S est assemblé et permet la synthèse de la chaîne polypeptidique lors de l'étape d'élongation.

ont permis de comprendre les étapes d'initiation, d'élongation et de terminaison, le mécanisme de recyclage du ribosome n'a, quant à lui, pas encore été totalement élucidé. Néanmoins, cette étape est cruciale pour la mise en place de la machinerie d'initiation afin de recommencer un cycle de synthèse protéique sur le même ARNm ou sur une autre molécule.

1. Le ribosome

La synthèse protéique fait intervenir un gros complexe ribonucléoprotéique appelé le ribosome. Ce complexe comprend trois sites principaux : le site A (*Aminoacyl-ARNt*) qui permet l'entrée de l'ARNt aminoacylé au sein du ribosome, le site P (*Peptidyl-ARNt*) qui contient la chaîne polypeptidique en cours de synthèse et le site E (*Exit*), contenant l'ARNt déacylé qui va être éjecté et recyclé. Le ribosome eucaryotique est formé de deux sous-unités (s-u), 40S et 60S, ayant un contenu protéique et ribonucléique différent. Chez la levure, le ribosome est composé de 79 protéines et de quatre ARNr. La petite s-u 40S, contenant 33 protéines différentes et l'ARNr 18S, est le centre de décodage du ribosome et le siège de l'interaction entre le codon présent dans l'ARNm et l'anticodon de l'ARNt aminoacylé correspondant. La grande s-u 60S, contenant 46 protéines et les ARNr 5S, 5,8S et 25S, catalyse, quant à elle, la formation de la liaison peptidique entre la protéine naissante et l'aa en cours d'incorporation. Le positionnement du ribosome sur l'AUG initiateur et sa capacité à se déplacer le long de l'ARNm, sont indispensables à la synthèse protéique. De plus, la participation de nombreux cofacteurs assure le bon fonctionnement de ce processus. La résolution des structures cristallographiques des s-u 40S et 60S, ainsi que du ribosome entier 80S, ont permis de décrypter, à l'échelle atomique, le mode de fonctionnement de ce complexe ribonucléoprotéique (Ben-Shem et al., 2011; Ben-Shem et al., 2010; Rabl et al., 2011; Wilson and Doudna Cate, 2012; Yusupov et al., 2001).

2. L'initiation de la traduction

L'initiation de la traduction eucaryotique est un mécanisme hautement régulé (**Figure 2**). En effet, il fait intervenir une douzaine de protéines différentes permettant l'association des deux s-u du ribosome et le positionnement du Met-ARNt^{Meti} au niveau du codon d'initiation AUG de l'ARNm. Chez les eucaryotes, la molécule d'ARNm subit des modifications post-transcriptionnelles, notamment l'ajout d'une coiffe (7-méthylguanosine) et d'une queue

poly-A, qui seront, toutes les deux reconnues par des eIF et la petite s-u du ribosome pour initier la synthèse protéique. La queue poly-A est ainsi recouverte de *poly-A binding protein* (PABP), capables d'interagir avec le complexe eIF4F et plus précisément avec la protéine de structure eIF4G, favorisant la circularisation de l'ARNm et la fixation du ribosome.

L'initiation de la traduction débute par la formation d'un complexe de pré-initiation 43S (PIC : *Pre Initiation Complex*) associant eIF2-GTP, impliqué dans le recrutement du Met-ARNt^{Met}_i, la s-u 40S du ribosome et les facteurs d'initiation eIF1, eIF1A, eIF3 et eIF5 (Algire et al., 2002; Asano et al., 2001; Kolupaeva et al., 2007; Majumdar et al., 2003). La fixation de ces facteurs à la s-u 40S induit une conformation ouverte favorisant ainsi l'interaction avec le complexe ternaire Met-ARNt^{Met}_i-eIF2-GTP. L'association des facteurs eIF1 et eIF1A induit un changement conformationnel de la s-u 40S qui conduit à l'ouverture du canal d'entrée où viendra se loger l'ARNm à traduire. Ces deux facteurs permettent également de stabiliser le complexe ternaire au sein du PIC 43S. Dans ce complexe 43S, la quantité d'eIF3 est importante (six s-u chez la levure et jusqu'à treize chez les autres espèces), ce facteur est donc quasiment aussi imposant que la petite s-u du ribosome et est positionné en périphérie de celle-ci. En effet, eIF3 se trouve, à la fois, près du site d'entrée et au niveau du canal de sortie de l'ARNm et il coordonne l'assemblage du PIC 43S.

En parallèle de la formation de ce PIC 43S, l'ARNm est pris en charge par plusieurs eIF. La coiffe 7-méthylguanosine est ainsi reconnue par le complexe multi protéique eIF4F. Celui-ci est composé de trois protéines : eIF4E, une *cap-binding protein* reconnaissant la coiffe et favorisant la circularisation de l'ARNm ; eIF4A, une hélicase permettant de déstructurer l'ARNm et eIF4G, une protéine de structure. La fixation du PIC 43S sur l'ARNm est médiée par deux étapes concomitantes. Dans un premier temps, l'activité hélicase de eIF4A, stimulée par son interaction avec eIF4G et le facteur additionnel eIF4B, génère une région simple brin dans la région 5'UTR (*UnTranslated Region*) de l'ARNm (au niveau de la coiffe). C'est alors l'interaction entre eIF4G et eIF3 qui initie la formation d'un pont entre le PIC 43S et l'ARNm, favorisant ainsi leur interaction. Ce PIC se déplace donc le long de la séquence nucléotidique de façon à positionner le Met-ARNt^{Met}_i, retrouvé dans le site P, en face du premier codon d'initiation de la traduction ; on parle de *scanning*. Lors de cette étape, l'AUG initiateur est reconnu car il se trouve dans un contexte nucléotidique optimal. En effet, le codon AUG est précédé d'une séquence consensus décrite pour la première fois par Kozak (Kozak, 1978).

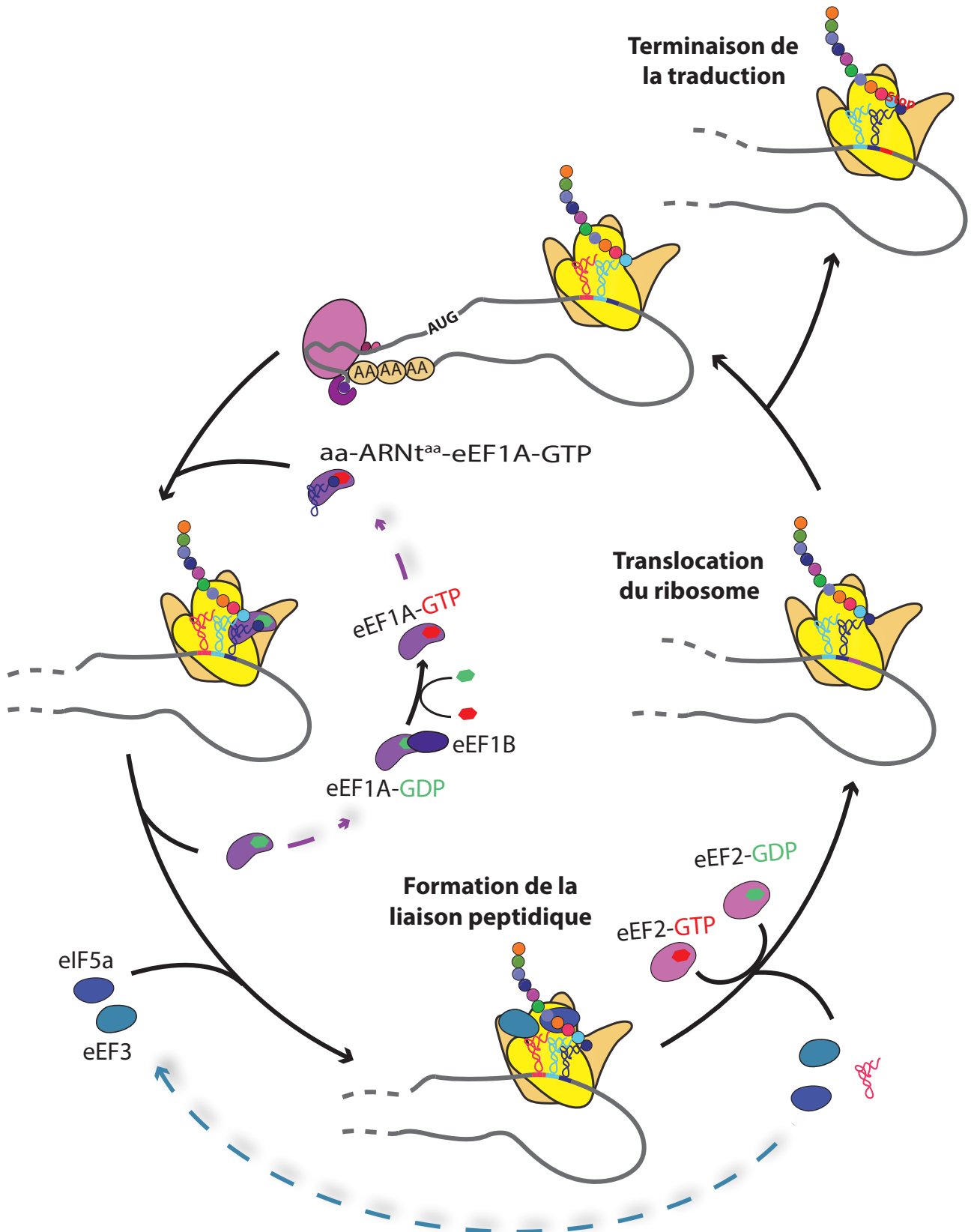


Figure 3 : L'élongation de la traduction chez les eucaryotes

L'étape d'élongation est cyclique et débute par le dépôt de l'aa-ARNT dans le site A du ribosome. Ensuite, l'étape de transpeptidation permet la formation d'une liaison peptidique entre l'aa-ARNT du site A et la chaîne polypeptidique naissante. Il y a alors translocation du ribosome, permettant le positionnement du codon suivant dans le site A, du peptidyl-ARNT dans le site P et de l'ARNT déacylé dans le site E. L'élongation se termine lorsqu'un codon stop entre dans le site A et induit le recrutement des facteurs de terminaison.

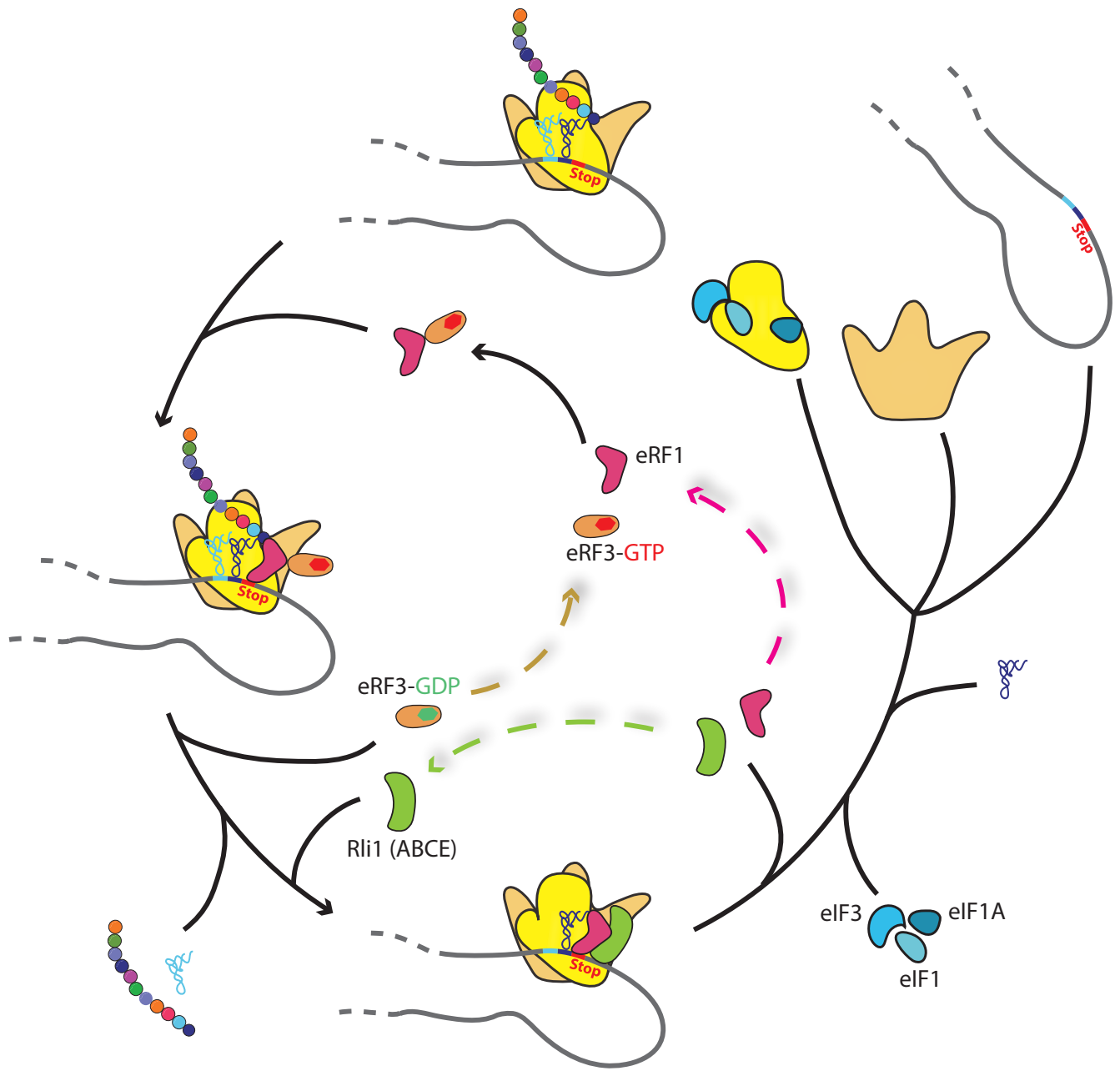
Celle-ci varie fortement en fonction des espèces (séquence chez la levure : AAAAAAAAAAUGTC (Cavener and Ray, 1991; Nakagawa et al., 2008). Cette séquence permet la fixation du Met-ARNt^{Met} sur le bon codon d'initiation et empêche l'initiation au niveau d'un autre codon AUG. Lorsque le codon d'initiation est reconnu, l'étape de *scanning* se termine ; il y a hydrolyse du GTP associé à eIF2 médiée par l'activité de la GAP eIF5 et formation du PIC 48S. L'affinité d'eIF2-GDP pour le Met-ARNt^{Met} diminue fortement et ce facteur se dissocie partiellement du complexe.

La finalisation de l'étape d'initiation et l'association de la grande s-u du ribosome est catalysée par l'action d'eIF4B-GTP. En effet, cette protéine induit le relargage d'eIF1, d'eIF3, d'eIF5 et d'eIF2-GDP résiduel. De plus, l'hydrolyse du GTP associé à eIF4B n'est pas essentielle pour l'assemblage du ribosome 80S mais stimule la dissociation de ce facteur. En parallèle, eIF5B interagit également avec eIF1A, favorisant leur dissociation du ribosome. L'élimination de la totalité des eIF est indispensable pour terminer l'étape d'initiation et permettre la fixation de la grande s-u 60S afin de commencer l'élongation de la chaîne polypeptidique. Le complexe 40S-60S-Met-ARNt^{Met}, aussi appelé complexe d'initiation 80S, est alors opérationnel pour entrer dans la phase d'élongation.

3. L'élongation de la traduction

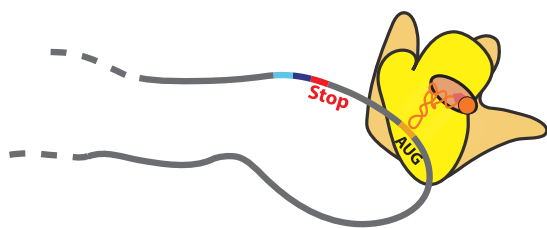
Contrairement à l'initiation, l'élongation de la traduction fait intervenir moins de facteurs. Cette étape permet la formation de la chaîne polypeptidique naissante par association des aa *via* la formation de liaisons peptidiques. Elle consiste en la répétition de trois sous-étapes jusqu'au positionnement d'un codon stop dans le site P du ribosome (**Figure 3**). Les structures du ribosome de *Saccharomyces cerevisiae* (*S. cerevisiae*) et de *Thermus thermophilus* ont révélé que ces deux complexes sont quasiment superposables, suggérant que la formation de la liaison peptidique est universellement conservée (Ben-Shem et al., 2011; Ben-Shem et al., 2010; Klinge et al., 2011).

L'initiation de la traduction se termine lorsque le ribosome 80S est positionné au niveau de l'ARNm avec l'anticodon du Met-ARNt^{Met} en interaction avec le codon d'initiation AUG dans le site P. Le codon suivant se trouve alors dans le site A du ribosome et est dans l'attente de l'association avec l'aminoacyl-ARNt^{aa} (aa-ARNt^{aa}) correspondant. Suite à la



Alternative au recyclage

Ré-initiation sur un AUG en aval



Ré-initiation sur l'AUG initial

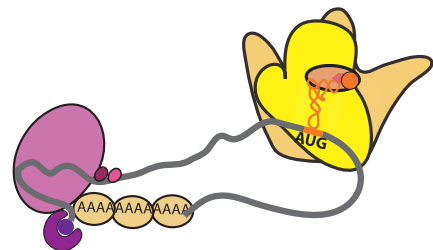


Figure 4 : La terminaison de la traduction et le recyclage du ribosome chez les eucaryotes

L'étape de terminaison et de recyclage du ribosome consiste à la mise en place des facteurs permettant la dissociation du ribosome afin de libérer la chaîne polypeptidique néosynthétisée et de réutiliser les deux sous-unités du ribosome pour un nouveau cycle de traduction. L'alternative au recyclage permet la ré-initiation, contexte spécifique, sur un AUG en aval ou sur l'AUG initial.

formation des différents aa-ARNt^{aa} par les aminoacyl-ARNt synthétases (aaRS) homologues (voir introduction § V.1), ceux-ci sont pris en charge par le facteur d'élongation eEF1A-GTP qui les délivre dans le site A du ribosome ; c'est la première étape de l'élongation. L'association de l'anticodon de l'aa-ARNt^{aa} avec le codon correspondant est médiée par l'hydrolyse du GTP qui permet l'accommodation de l'ARNt dans le site A et le relargage d'eEF1A-GDP. Ce dernier est alors recyclé par intervention du facteur eEF1B, un GEF, qui catalyse l'échange d'une molécule de GDP par un GTP. Le facteur eEF1A-GTP peut alors lier un autre aa-ARNt^{aa} afin de le déposer dans le site A du ribosome au niveau du codon correspondant. Le dépôt de l'aa-ARNt^{aa} dans le site A est suivi de l'étape de transpeptidation, médiée par l'ARNr28S, qui induit la formation d'une liaison peptidique entre le COOH du peptide en cours de synthèse et le NH₂ de l'aa-ARNt^{aa} présent dans le site A du ribosome. Le site P contient alors un ARNt déacylé et la protéine en cours de synthèse associée à l'ARNt (Peptidyl-ARNt) est retrouvée dans le site A. La dernière étape de l'élongation correspond à la translocation du ribosome, favorisée par l'hydrolyse d'eEF2-GTP, permettant ainsi de décaler le polypeptide en cours de synthèse dans le site P et l'ARNt déacylé dans le site E. Chez la levure, un autre facteur d'élongation est impliqué, eEF3. Ce facteur se retrouve dans la partie haute du ribosome et plus particulièrement à l'interface entre la protubérance de la s-u 60S et au niveau de la tête de la s-u 40S. Il est donc capable de se fixer au niveau du site E du ribosome et favorise l'éjection de l'ARNt déacylé (Andersen et al., 2006; Triana-Alonso et al., 1995). En plus de sa localisation au niveau du ribosome, eEF3 serait également capable de se fixer au sein de régions non traduites et assure potentiellement une fonction indépendante de celle qu'elle joue dans l'élongation (Anand et al., 2006; Anand et al., 2003).

Suite à l'étape de translocation, le site A est vide et disponible pour relancer un cycle d'élongation par fixation de l'aa-ARNt^{aa} correspondant au codon suivant. Ces trois sous-étapes d'élongation se suivent et permettent le décodage de l'ensemble de la molécule d'ARNm jusqu'à arriver à un des codons stop, induisant l'arrêt de la traduction (**Figure 3**).

4. La terminaison de la traduction

La terminaison de la traduction débute lorsque qu'un des codons stop (UAG, UGA ou UAA) est positionné dans le site A du ribosome (**Figure 4**). Chez les eucaryotes, la terminaison fait intervenir deux facteurs, eRF1 et eRF3, qui collaborent sous forme d'un complexe ternaire

eRF1-eRF3-GTP (Alkalaeva et al., 2006; Stansfield et al., 1995; Zhouravleva et al., 1995). Le premier, eRF1, est impliqué dans la reconnaissance spécifique du codon stop, ainsi que dans l'hydrolyse du peptidyl-ARNt. Ce facteur adopte une structure globale en forme d'ARNt permettant d'augmenter sa spécificité de reconnaissance du codon stop (Song et al., 2000). Le second facteur, eRF3-GTP, accélère la libération du polypeptide néosynthétisé et favorise la terminaison de la traduction grâce à l'hydrolyse du GTP en GDP (Alkalaeva et al., 2006; Frolova et al., 1996). En effet, eRF3 pourrait jouer un rôle similaire à eEF1A en emmenant eRF1, ayant une structure d'ARNt, au niveau du codon stop, favorisant finalement la terminaison de la traduction.

5. Le recyclage du ribosome

Une fois la chaîne polypeptidique éliminée, le complexe entre le ribosome, l'ARNm, l'ARNt déacylé et les eRF peut entrer dans l'étape de recyclage (**Figure 4**). Cette dernière permet la dissociation de ces composants et favorise ainsi la récupération de la s-u 40S pour une nouvelle initiation. Il existe deux alternatives à l'étape de recyclage : (i) le complexe peut se dissocier complètement pour que la s-u 40S soit disponible pour un nouveau cycle de traduction ou (ii) le complexe ne se dissocie pas en totalité (Pisarev et al., 2007; Rajkowitsch et al., 2004). Lorsque ce complexe ne se dissocie pas, le ribosome reste sur le même ARNm et peut initier la traduction sur le codon d'initiation utilisé lors de la première synthèse protéique du fait de la circularisation de l'ARNm. Il est également possible que l'initiation du nouveau cycle de traduction débute sur un codon AUG en aval du premier, on parle de ré-initiation (Hinnebusch, 1997; McCarthy, 1998; Sachs and Geballe, 2006). Cependant, l'étape de ré-initiation n'est possible que dans des conditions bien précises dépendant notamment de la vitesse des cycles d'élongation, de la longueur de la phase de lecture à traduire (elle doit être inférieure à 35 codons) ou du contexte nucléotidique au niveau du codon stop (Hinnebusch, 2005; Jackson et al., 2010).

Le recyclage chez les eucaryotes est très différent de ce que l'on retrouve chez les procaryotes. Certaines études suggèrent que le facteur d'initiation eIF3 participe activement à l'étape de recyclage du ribosome en s'associant à l'interface entre les deux s-u (Pisarev et al., 2007). En parallèle, les facteurs d'élongation eEF1 et eEF3, associés sous forme de complexe, peuvent également jouer un rôle dans la dissociation lente du ribosome

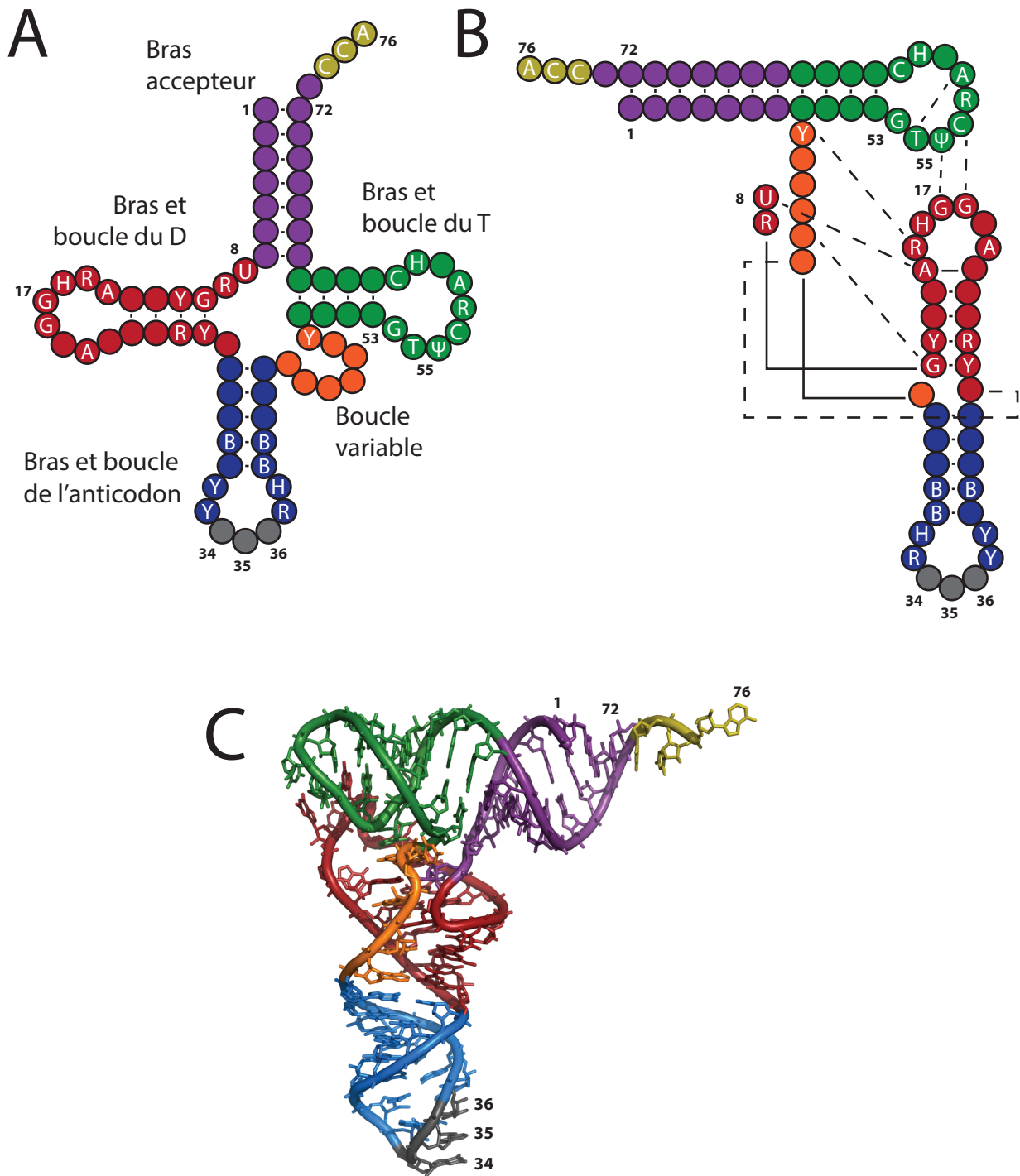


Figure 5 : Structure secondaire et repliement 3D de l'ARNt

(A) Structure secondaire en feuille de trèfle, les nucléotides sont numérotés (en noir) à partir de l'extrémité 5' phosphate. (B) Structure 3D en L renversé, les pointillés indiquent les interactions longues distances. (C) Structure cristallographique de l'ARNt^{Phe} de levure (accession pdb : 1TRA).

(Shoemaker et al., 2010). Quoi qu'il en soit, le recyclage chez les eucaryotes fait intervenir une protéine multifonctionnelle ABCE1 (Rli1 chez la levure) (Barthelme et al., 2011; Pisarev et al., 2010). Cette protéine utilise l'énergie produite par l'hydrolyse de l'ATP pour déstabiliser le complexe des s-u du ribosome et les séparer. De plus, ABCE1 augmente fortement la dissociation du complexe eEF1-eEF3-GDP (Khoshnevis et al., 2010; Shoemaker and Green, 2011). Les deux s-u du ribosome étant libres, la s-u 40S peut être prise en charge par les eIF pour permettre la formation du PIC 43S au niveau d'un nouvel ARNm.

IV. Les ARN de transfert

1. La structure des ARNt

Les ARNt sont des molécules ubiquitaires et essentielles à la synthèse protéique. Ce sont les ARN les plus représentés dans la cellule (environ 4 à 10 % des ARN cellulaires). Les ARNt, lorsqu'ils sont aminoacylés, agissent comme des adaptateurs lors de la synthèse protéique. En effet, leur fonction principale est d'amener l'aa au niveau du ribosome afin qu'il puisse être incorporé dans la chaîne polypeptidique en cours de synthèse. Une molécule d'ARNt a une taille comprise entre 73 et 90 nucléotides et se replie en une structure secondaire dite en feuille de trèfle arborant quatre boucles : la boucle D, la boucle T, la boucle (ou région) variable (sa taille varie de 3 à 21 bases) et la boucle de l'anticodon (**Figure 5A**). L'établissement d'interactions à longue distance au sein de la structure secondaire permet finalement à l'ARNt d'adopter une conformation tri-dimensionnelle (3D) en L renversé (**Figure 5B & C**).

Les ARNt sont synthétisés au sein du nucléole sous forme de longs transcrits possédant une séquence 5'-leader et une extension 3'-trailer, toutes deux éliminées lors de l'étape de maturation. En plus de ces extensions, certains gènes d'ARNt sont caractérisés par la présence d'introns (Heinemann et al., 2010). Néanmoins, le pourcentage de gènes d'ARNt possédant un intron varie fortement en fonction des espèces. Ainsi, ce pourcentage est quasiment nul chez les bactéries, d'environ 5 % chez l'Homme et la souris pour atteindre plus de 60 % chez certaines archées (Genomic tRNA Database, <http://lowelab.ucsc.edu/GtRNAdb>). A titre d'exemple, chez la levure *S. cerevisiae*, sur les 295 séquences de gènes d'ARNt répertoriées (à la date du 20/10/17), 59 possèdent un intron, soit environ 20 % (Genomic tRNA Database,

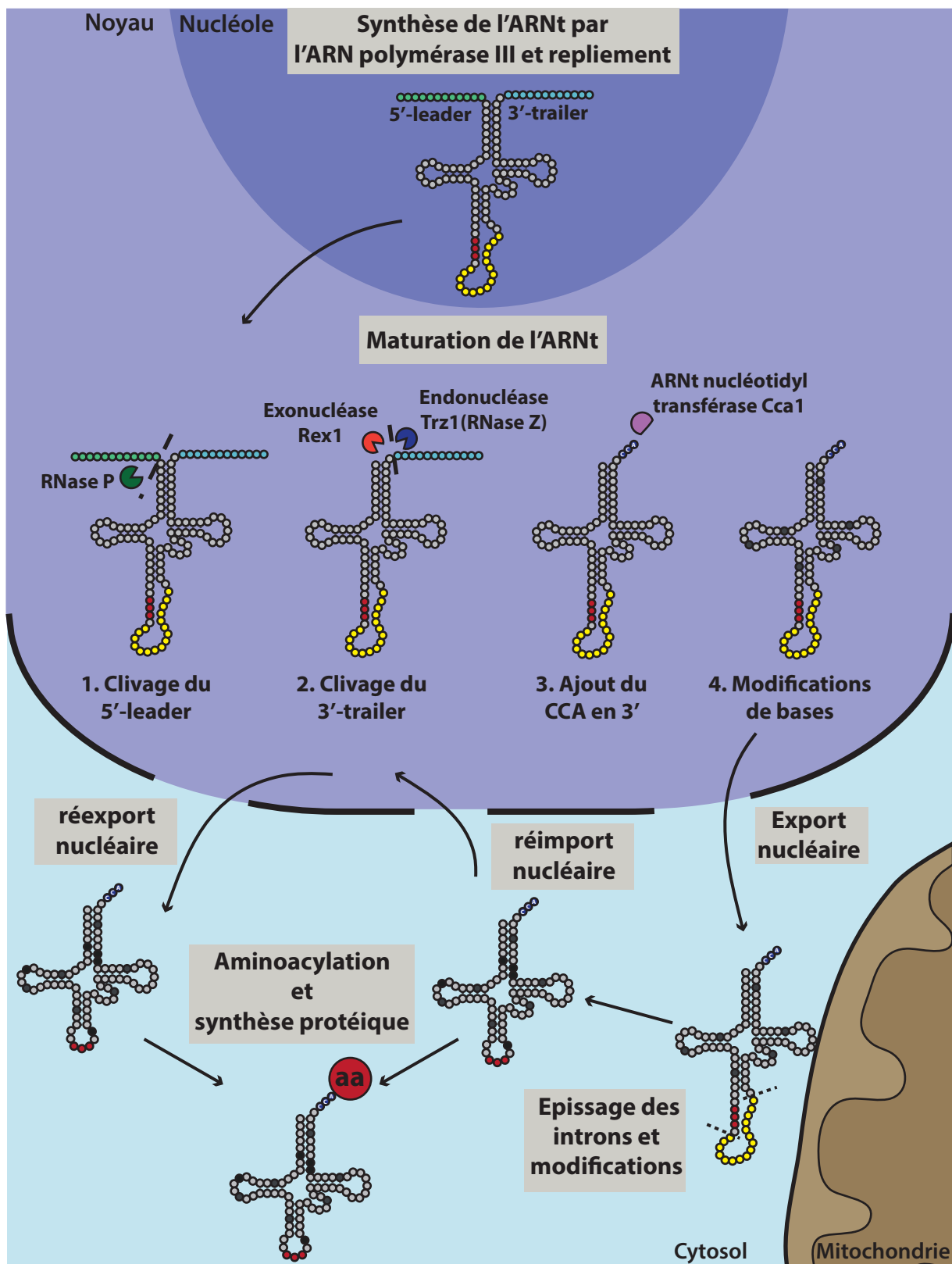


Figure 6 : Synthèse et maturation des ARNt

Les ARNt (avec ou sans introns) sont synthétisés dans le nucléole. Les extrémités 5'-leader et 3'-trailer sont alors éliminées et le CCA est ajouté. Après modifications de bases, les ARNt sont exportés du noyau. L'épissage des introns et d'autres modifications de bases ont alors lieu à la surface des mitochondries. Les ARNt sont finalement aminoacylés et utilisés lors de la synthèse protéique, ou réimportés dans le noyau pour subir un contrôle qualité. Ces ARNt sont finalement réexportés et servent à la synthèse protéique après leur aminoacylation.

<http://lowelab.ucsc.edu/GtRNADB>). Le nombre et la proportion des introns dans les gènes d'ARNt fluctuent fortement. Toutefois, la majorité des précurseurs contiennent un intron inséré entre les positions 37 et 38 de l'ARNt. L'élimination de cet intron est primordiale car il se retrouve juste après l'anticodon (positions 34, 35 et 36), qui permet le décodage d'un codon lors de la synthèse protéique.

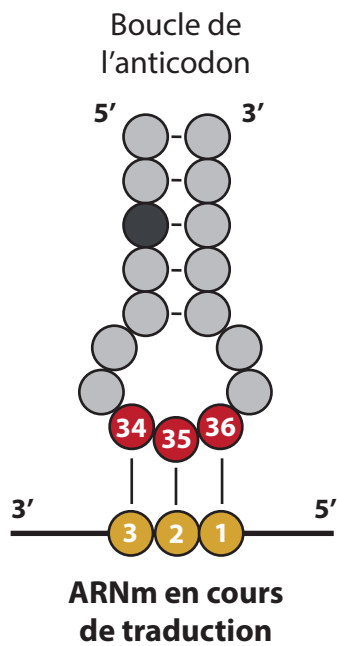
2. La maturation des pré-ARNt

Chez la levure, les extensions 5'-leader et 3'-trailer font environ douze nucléotides (O'Connor and Peebles, 1991). La maturation du précurseur d'ARNt débute généralement par l'élimination de l'extrémité 5'-leader par la RNase P (**Figure 6**). Chez la levure, ce complexe ribonucléoprotéique est composé d'un ARN catalytique et de neuf protéines essentielles (Chamberlain et al., 1998). Dans certains cas, l'extrémité 3' peut être éliminée en premier, comme dans le pré-ARNt^{Trp} par exemple (Kufel and Tollervey, 2003; O'Connor and Peebles, 1991). Pour l'ARNt^{His}, l'action de la RNase P n'est pas suffisante et la maturation de la région 5' nécessite l'intervention de la Guaninyl transférase Thg1. Celle-ci va permettre l'ajout d'une guanosine en 5' de l'ARNt^{His} (position -1 qui se lie à la base discriminatrice 73), essentielle pour l'aminocyclation de cet ARNt (Gu et al., 2005; Gu et al., 2003; Preston and Phizicky, 2010).

Chez la levure, l'élimination de l'extrémité 3'-trailer est plus compliquée. En effet, cette étape fait intervenir (i) l'exonucléase Rex1 (intervenant également dans la maturation de l'ARNr 5S et 5,8S, ainsi que des ARNsn (sn : *small nuclear*) (Copela et al., 2008; Ozanick et al., 2009; van Hoof et al., 2000) ou (ii) l'endonucléase Trz1, encore appelée RNase Z (Chen et al., 2005; Maraia and Lamichhane, 2011) (**Figure 6**). La régulation de ce clivage est possible par la fixation de la protéine Lhp1. Cette dernière empêche l'accès à l'exonucléase Rex1 et favorise donc l'élimination du 3' par l'endonucléase Trz1 (Yoo and Wolin, 1997).

Pour permettre son aminocyclation, un ARNt doit présenter en 3' un triplet ^{5'}CCA^{3'}. Chez la levure, celui-ci est rajouté par l'ARNt nucléotidyl transférase (ou CCCase) Cca1 (Aebi et al., 1990) (**Figure 6**). Néanmoins, l'ajout de ce triplet n'est pas nécessaire chez tous les organismes. En effet, chez *Escherichia coli* (*E. coli*) et d'autres espèces bactériennes, les ARNt sont directement synthétisés avec le CCA en 3'. L'ARNt nucléotidyl transférase est cependant présente chez ces organismes car elle catalyse l'ajout d'une étiquette CCA au niveau d'ARNt

A



B

Première base de l'anticodon	Troisième base du codon
34	3
G	U, C
C	G
A	U, C, G, (A)
U	U, (C), G, A
I	A, C, U
xm5s2U xm5U xm5Um Um	A, G
xo5	U, A, G

Figure 7 : Principe de l'appariement *wobble*

(A) Schéma de l'appariement codon/anticodon. (B) Tableau des interactions entre la base en position 34 de l'anticodon (boule rouge) et la 3^{ième} base du codon (boule jaune). Les interactions de type "*wobble*" sont en orange.

instables. Le motif $5'CCACCA^3'$ est alors reconnu par la machinerie de dégradation cellulaire (Wilusz et al., 2011).

Au cours de leur maturation, les ARNt vont également être modifiés, et on dénombre à l'heure actuelle une centaine de modifications de bases de l'ARNt chez les eucaryotes (<http://mods.rna.albany.edu/home>). Chez la levure, 25 modifications d'ARNt sont répertoriées et le nombre de modifications par ARNt peut varier de 7 à 17 (Phizicky and Hopper, 2010). Les modifications de bases des ARNt peuvent avoir lieu, soit dans le cytosol sur des ARNt épissés, soit dans le noyau sur des ARNt contenant encore les introns ou des ARNt épissés, réimportés dans le noyau par la voie rétrograde (Anderson et al., 1998; Hopper and Phizicky, 2003; Ohira and Suzuki, 2011; Tolerico et al., 1999). Les modifications post-transcriptionnelles des ARNt sont importantes pour différents mécanismes cellulaires essentiels et notamment dans le cadre de la reconnaissance codon/anticodon. Par exemple, l'adénosine (A) en position 34 peut être modifiée en inosine (I) qui permet le décodage de trois codons au lieu d'un seul. Une A34 pourra uniquement s'apparier avec une uridine (U), alors qu'une I34 pourra interagir avec une A, une cytosine (C) ou une U (**Figure 7**). Il est à noter que la modification de bases et plus particulièrement de la position wobble ne n'élargit pas uniquement la capacité décodante des ARNt. En effet, l'ARNt^{Glu} est modifié au niveau de la position 34 par un groupement 5-méthylaminométhyl-2-thio (mnm5s2) qui permet de moduler la vitesse de traduction. L'élimination du 5-méthylaminométhyl induit une augmentation de la vitesse de décodage des codons glutamate (deux à quatre fois), tandis que l'élimination du groupement 2-thio induit une diminution de cette vitesse (jusqu'à 20 %) par rapport à la capacité de décodage de l'ARNt^{Glu} sauvage (Kruger et al., 1998).

3. L'export des pré-ARNt et l'épissage des introns

Les ARNt en cours de maturation sont exportés du noyau vers le cytosol au travers des pores nucléaires par une voie à laquelle participe Gsp1 (l'homologue levure de la GTPase humaine Ran). L'association de l'ARNt avec les exportines est RanGTP-dépendante. Après passage des ARNt dans le cytosol, le GTP est hydrolysé en GDP par la RanGAP (*Ran GTPase-activating protein*) permettant le relargage de l'ARNt et le recyclage de l'exportine. Il existe deux voies principales d'export des ARNt chez la levure : la voie d'export des ARNt possédant un intron et la voie d'export des ARNt sans intron. L'export des ARNt possédant un intron peut

être accompli uniquement par l'exportine Los1 (Hellmuth et al., 1998), alors que les autres ARNt sont exportés, soit par une autre karyophérine, Msn5, soit par Los1. Cette dernière induit donc l'export des ARNt avec ou sans intron, mais également des ARNt ayant été épissés puis réimportés dans le noyau par la voie rétrograde. Cependant, la délétion de *LOS1* et *MSN5* n'est pas létale et il doit exister d'autres exportines des ARNt chez la levure. De manière intéressante, la double délétion de *LOS1* et d'*ARC1* est létale. Cette dernière dont je vais parler plus précisément lors de la partie suivante pourrait participer à l'export des ARNt. Cependant, aucune preuve biologique de son implication dans ce mécanisme n'a été mise en évidence.

Un fois exportés du noyau, les ARNt possédant des introns sont maturés par élimination de leurs introns en trois étapes à la surface de la mitochondrie (Yoshihisa et al., 2003). L'intron en aval de l'anticodon est tout d'abord éliminé par une endonucléase tétramérique (Sen2, Sen34, Sen15 et Sen54) (Paushkin et al., 2004; Trotta et al., 1997) et conduit à la formation de deux demi ARNt. Le fragment 5', encore appelé exon 1, se termine par un phosphate 2',3'-cyclique, alors que le fragment 3', ou exon 2, possède une extrémité hydroxyle (Knapp et al., 1979; Peebles et al., 1979). L'étape suivante consiste en la ligation des deux exons par l'ARNt ligase Trl1. Pour cela, le cycle 2',3' du phosphate de l'exon 1 est d'abord éliminé et le 5' de l'exon 2 est phosphorylé. L'extrémité 5' est alors activée par l'adénosine monophosphate (AMP) et l'ARNt ligase lie les deux exons par une liaison phosphodiester. L'épissage de l'ARNt se termine par l'intervention de l'ARNt 2'-phosphodiesterase Tpt1 qui excise le 2'-phosphate additionnel à la jonction des deux exons (**Figure 6**).

4. Les fonctions additionnelles des ARNt et des fragments d'ARNt

Le rôle principal d'un aa-ARNt se situe au niveau de la synthèse protéique *via* l'association codon/anticodon, indispensable pour l'incorporation de l'aa correspondant au niveau de la chaîne polypeptidique naissante. Cependant, les ARNt non chargés, mais également les aa-ARNt, peuvent être détournés de la synthèse protéique pour participer à d'autres mécanismes essentiels et assurer des fonctions dites additionnelles.

Chez les procaryotes, le Gly-ARNt^{Gly} est impliqué dans la biogénèse de la paroi (Navarre and Schneewind, 1999), les Lys-ARNt^{Lys} et Ala-ARNt^{Ala} peuvent être utilisés pour l'aminoacylation des lipides membranaires (Roy and Ibba, 2008). De plus, certains ARNt non

chargés sont également capables de réguler l'expression des gènes. C'est notamment le cas chez *Clostridium acetobutylicum*, où l'ARNt^{Asn} permet, *via* un mécanisme de *riboswitch*, de réguler l'expression d'un opéron comprenant les gènes codant l'aaRS correspondante et les trois s-u de l'amidotransférase (AdT) GatCAB (Saad et al., 2012). Comme je vais le décrire dans la partie suivante, les AdT sont responsables de la formation des Asn-ARNt^{Asn} et Gln-ARNt^{Gln} selon un mécanisme ARNt-dépendant. Ces deux ARNt jouent donc le rôle de cofacteurs pour la biogénèse d'aa (voir introduction § V.6). Cette fonction additionnelle n'est cependant pas restreinte aux procaryotes et les organelles des eucaryotes utilisent également l'activité d'une AdT pour la formation de certains aa-ARNt^{aa}.

Chez les eucaryotes, les ARNt assurent également des fonctions additionnelles. C'est notamment le cas de l'ARNt^{Glu} qui participe à la biosynthèse de l'hème et de la chlorophylle (Jahn, 1992) ou encore certains ARNt qui servent d'amorces pour la réplication des rétrovirus (Marquet et al., 1995). De plus, les avancées des techniques de séquençage à haut débit ont permis d'identifier une nouvelle classe de petits ARNnc. Ces nouveaux ARN correspondent à des fragments d'ARN cellulaires et peuvent être issus d'ARNr ou d'ARNt. Les fragments d'ARNt, encore appelés tRF (*tRNA Fragments*), ont des rôles biologiques essentiels et peuvent être synthétisés aussi bien à partir de précurseurs d'ARNt que d'ARNt matures chez tous les organismes référencés à ce jour (Lee et al., 2009). Les tRF sont classés en différents types qui correspondent à la région de l'ARNt impliquée. Actuellement, de nombreuses études suggèrent que ces fragments d'ARNt interviennent dans une multitude de processus biologiques. Chez l'Homme, ces tRF sont détectés dans diverses pathologies, notamment le cancer. Dans certains cas, ils peuvent même servir de marqueurs biologiques. Des études ont, en effet, montré que l'expression de tRF particuliers est augmentée dans les métastases de patients atteints de cancer de la prostate (Martens-Uzunova et al., 2012). En parallèle, l'expression des tRF peut être modulée par certaines hormones, et cela induit une augmentation de la prolifération cellulaire tissu-spécifiques chez les patients atteints de cancer de la prostate ou de cancer du sein (Dhahbi et al., 2014; Honda et al., 2015; Lee et al., 2009). L'expression d'un tRF issu de l'ARNt^{lys} est stimulée dans les cellules infectées par le VIH (Virus de l'Immunodéficience Humaine) et son absence pourrait favoriser la réplication virale (Yeung et al., 2009). Chez la levure *S. cerevisiae*, les tRF sont synthétisés dans différentes conditions de croissance mais leur fonction a très peu été étudiée. Toutefois, certains tRF sont

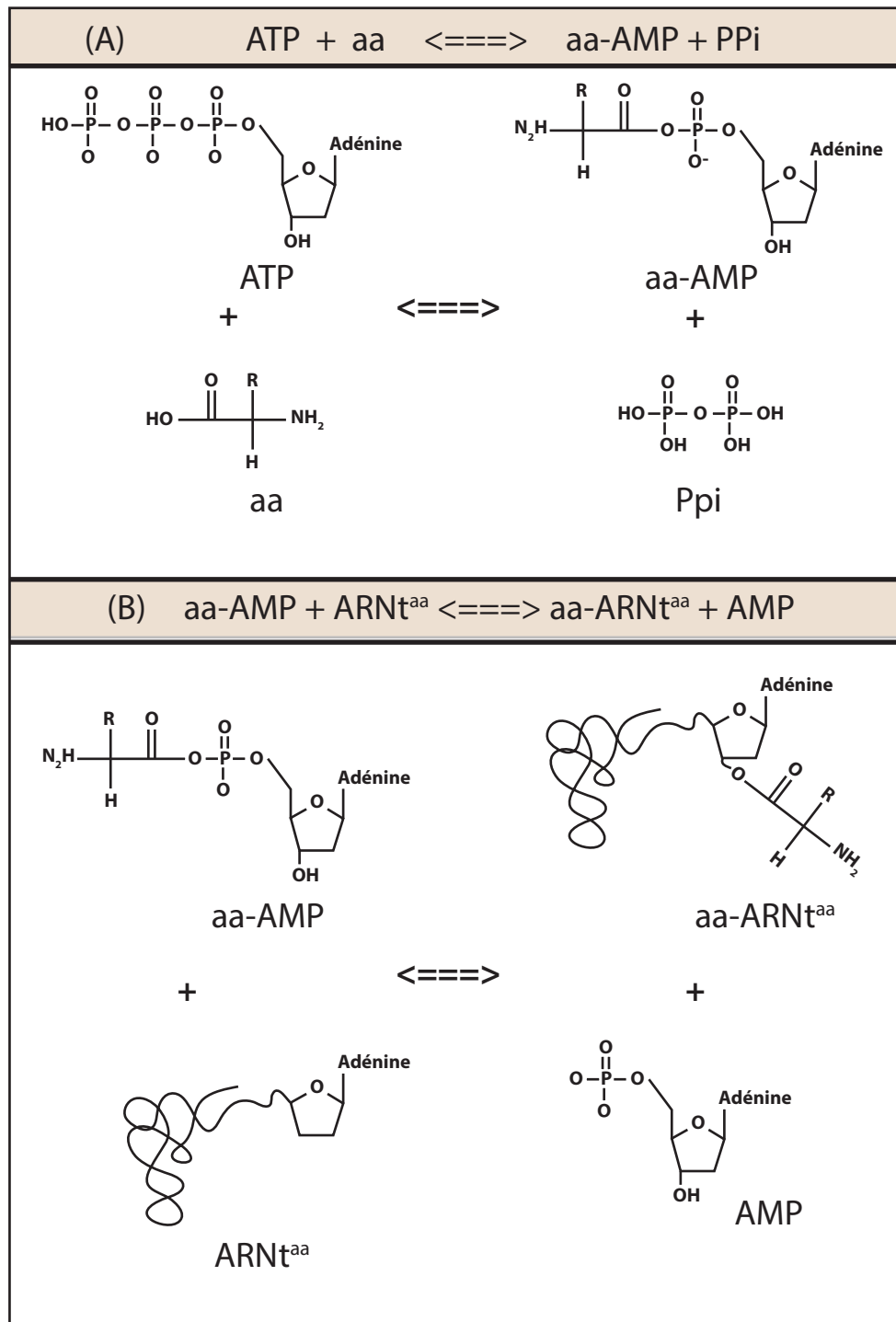


Figure 8 : La réaction d'aminocyclation

La réaction d'aminocyclation se décompose en deux étapes : (A) l'aa est activé par formation d'un aa~AMP en présence ou non de l'ARNt homologue ; (B) l'aa est alors transféré sur l'ARNt^{aa} correspondant et il y a libération d'AMP. Cet aa-ARNt^{aa} peut alors être pris en charge par les facteurs d'initiation (Met-ARNt^{Met}_i) ou d'élongation (les autres aa-ARNt^{aa}) afin que l'aa soit incorporé dans la chaîne polypeptidique naissante.

capables de s'associer au ribosome et favorisent l'inhibition de la synthèse protéique *in vitro* (Bakowska-Zywicka et al., 2016a; Bakowska-Zywicka et al., 2016b; Zywicki et al., 2012). L'étude de la fonction des tRF de levure en est donc à son commencement et la compréhension de ces fonctions chez cet unicellulaire pourrait permettre de mieux appréhender leurs fonctions chez l'Homme.

V. Les aminoacyl-ARNt synthétases

1. La fonction primaire des aaRS

Lors de la synthèse protéique, les aa-ARNt sont utilisés par le ribosome comme substrat pour l'incorporation des aa au sein de la chaîne polypeptidique naissante. Ces aa-ARNt sont synthétisés par les aminoacyl-ARNt synthétases *via* une réaction d'aminacylation qui permet la fixation d'un aa donné sur l'ensemble des ARNt isoaccepteurs. La réaction d'aminacylation se déroule en deux étapes : (i) l'activation de l'aa et (ii) le transfert de l'aa sur l'ARNt homologue (**Figure 8**).

Lors de l'activation de l'aa, le groupement carboxylique de celui-ci estérifie le groupement phosphate de l'AMP, formant ainsi un anhydride mixte nommé aminoacyl-adénylate (aa-AMP). La majorité des aaRS sont capables d'activer l'aa sans leur ARNt homologue. Cependant, l'activation par l'ArgRS, la GluRS et la GlnRS nécessite la présence de l'ARNt. Les aaRS sont en général très spécifiques lors de la réaction d'activation (Fersht and Dingwall, 1979). Néanmoins, des erreurs dans la sélection de l'aa sont possibles mais peuvent être corrigées par des domaines d'édition dont se sont dotées certaines aaRS (Mascarenhas and Martinis, 2008).

Suite à la formation de l'aa~AMP, l'aa est transféré sur l'ARNt homologue par formation d'une liaison ester entre le carboxyle qui était activé et le groupement 2' ou 3' OH du ribose de l'A76 du bras accepteur de l'ARNt. La reconnaissance de l'ARNt par une aaRS est primordiale pour la fiabilité de la synthèse protéique et cette spécificité s'est affinée au cours de l'évolution. En effet, au sein de chaque ARNt se trouvent des déterminants d'identités favorisant la reconnaissance d'un ARNt spécifique par son aaRS homologue (Sissler et al., 1998).

	Classe I	Classe II
Structure du domaine catalytique	Domaine de Rossmann 5 brins β parallèles	7 brins β antiparallèles entrecoupés d'hélice α
Séquences signatures du domaine catalytique	HIGH KMSKS	Motif 1 : gΦxxΦxxP$\Phi$$\Phi$ Motif 2 : fRxϵ-h/rxxxxFxxx(d/e) Motif 3 : gΦgΦgΦ(d/e)R$\Phi$$\Phi$$\Phi$$\Phi$
Résidu accepteur de l'aa~AMP	2'OH de l'Adénosine 76	3'OH de l'Adénosine 76
Aminoacyl-ARNt synthétases, sous-classes et oligomérisation	<p>Ia</p> Méthionyl-ARNt synthétase α Valyl-ARNt synthétase α Isoleucyl-ARNt synthétase α Leucyl-ARNt synthétase α Cystéinyl-ARNt synthétase α Arginyl-ARNt synthétase α	<p>Ila</p> Séryl-ARNt synthétase α_2 Thréonyl-ARNt synthétase α_2 Alanyl-ARNt synthétase α Glycyl-ARNt synthétase α_2 Prolyl-ARNt synthétase α_2 Histidyl-ARNt synthétase α_2
	<p>Ib</p> Glutamyl-ARNt synthétase α Glutaminyl-ARNt synthétase α Lysyl-ARNt synthétase α	<p>Ilb</p> Aspartyl-ARNt synthétase α_2 Asparaginyll-ARNt synthétase α_2 Lysyl-ARNt synthétase α_2
	<p>Iic</p> Tyrosyl-ARNt synthétase α_2 Tryptophanyl-ARNt synthétase α_2	<p>Iic</p> Phénylalanyl-ARNt synthétase $\alpha_2\beta_2$ Pyrrolysyl-ARNt synthétase α_2 O-phosphoryl-ARNt synthétase α_4

Figure 9 : Les différentes classes d'aaRS

Les aaRS sont séparées en deux classes basées sur la structure du domaine catalytique, les séquences signatures, ainsi que sur le groupement OH, accepteur de l'aa. Les structures oligomériques sont précisées en regard de l'aaRS.

2. Les classes d'aaRS

Le code génétique standard permet l'incorporation de 20 aa différents lors de la synthèse protéique. Il existe donc 20 aaRS qui catalysent l'aminocyclation de l'ensemble des aa sur leurs ARNt homologues. La classification des aaRS est basée sur la structure de leur site catalytique, leur structure 3D, ainsi que leurs caractéristiques catalytiques. Celles-ci sont séparées en deux classes, la classe I et la classe II (Cusack et al., 1990; Eriani et al., 1990; Ribas de Pouplana and Schimmel, 2001). Mise à part la lysyl-ARNt synthétase (LysRS), qui peut être de classe I ou II, il n'y a pas d'autre exception à l'appartenance d'une aaRS à l'une des classes (e.g. les glutamyl-ARNt synthétases (GluRS) sont toujours de classe I). Au sein de chaque classe, les aaRS sont séparées en sous-groupes (a-c) basés sur les caractéristiques du domaine catalytique ou de l'aa qu'elles reconnaissent, mais également sur la structure du domaine de liaison à l'anticodon (**Figure 9**).

Les aaRS de classe I sont généralement retrouvées sous forme de monomère, même si la TyrRS et la TrpRS sont actives uniquement sous forme de dimère (Iborra et al., 1973; Jones et al., 1985). Le site catalytique des aaRS de cette classe renferme un domaine de Rossmann caractérisé par la présence de deux séquences signatures : Lys-Met-Ser-Lys-Ser (KMSKS) et His-Ile-Gly-His (HIGH), toutes deux impliquées dans la stabilisation (i) de l'ATP pendant l'activation de l'aa et de (ii) l'ARNt au cours du transfert de l'aa. Finalement, les aaRS de classe I permettent l'association de l'aa sur le 2'-OH de l'A76.

Les aaRS de classe II sont, quant à elles, caractérisées par une organisation oligomérique (di- ou tétramérique) et possèdent un domaine catalytique formé de sept brins β antiparallèles encadrés par des hélices α . Au sein de cette structure, on retrouve trois motifs : 1 (g Φ xx Φ xxP Φ Φ), 2 (fRxe-h/rxxxFxxx(d/e)) et 3 (g Φ g Φ g Φ (d/e)R Φ Φ Φ Φ), où Φ correspond à un résidu hydrophobe, x n'importe quel aa et X un aa conservé. Ici, le code à une lettre est utilisé : P correspond à la proline dans le motif 1 par exemple. La majorité des aaRS de classe II aminoacylent sur le 3'OH de l'A76 de l'ARNt. Cependant, L'AsnRS et la PheRS fixent l'aa sur le 2'OH (Chinault et al., 1977; Sprinzl and Cramer, 1975). En plus des dix aaRS de cette classe, on retrouve deux autres enzymes de classe II, la O-phosphoséryl-ARNt synthétase (SepRS) et la pyrrolysyl-ARNt synthétase (PylRS) retrouvées majoritairement chez

les archées méthanogènes et qui sont impliquées dans la formation du Cys-ARNt^{Cys} et Pyl-ARNt^{Pyl} respectivement.

3. L'incorporation des 21^{ème} et 22^{ème} aa

Lors de la synthèse protéique, il y a incorporation des 20 aa définis par le code génétique. Cependant, il peut également y avoir incorporation de deux aa additionnels n'appartenant pas au code génétique standard, la Sec et la Pyl qui sont respectivement décodées par les codons stop UGA et UAG.

La Sec est un aa retrouvé dans les trois domaines du vivant (Commans and Bock, 1999), qui joue, notamment, un rôle de nucléophile dans les réactions d'oxydoréduction. Le Sec-ARNt^{Sec} est synthétisé *via* une voie indirecte faisant intervenir, dans un premier temps, la séryl-ARNt synthétase (SerRS) qui charge une Ser sur l'ARNt^{Sec}. Cet aa-ARNt mésacylé est ensuite pris en charge par des protéines spécifiques pour former le Sec-ARNt^{Sec}. Chez les eucaryotes et les archées, le Ser-ARNt^{Sec} est pris en charge par l'O-phosphoséryl-ARNt^{Sec} kinase (PSTK) qui phosphoryle la Ser pour former le Sep-ARNt^{Sec}. Ce dernier est finalement converti en Sec-ARNt^{Sec} par la Sep-ARNt:Sec-ARNt synthase (SepSecS) en présence de phosphate de pyridoxal et de sélénophosphate. Chez les bactéries, la réaction est plus directe ; c'est la sélénocystéine synthase SelA qui prend en charge le Ser-ARNt^{Sec} et permet la formation de Sec-ARNt^{Sec} en présence de sélénophosphate et de phosphate de pyridoxal. La voie d'aminacylation de la Pyl est, quant à elle, directe. Cet aa est retrouvé chez les archées méthanogènes et la formation du Pyl-ARNt^{Pyl} fait intervenir une PylRS qui active la Pyl et la fixe sur l'ARNt^{Pyl} (Polycarpo et al., 2004).

Pour permettre l'incorporation de ces deux aa au sein de la chaîne polypeptidique, il existe un signal structural sur l'ARNm pour signifier au ribosome que la traduction ne se termine pas au niveau de ces codons UAG ou UGA particuliers. Celui-ci prend la forme d'une tige-boucle en aval du codon stop, appelé "PYrroLysin Inserted Sequence" (PYLIS) ou "SElenoCystein Inserted Sequence" (SECIS). Cette structure empêche alors les facteurs de terminaison d'être recrutés sur les codons UGA et UAG et de mettre fin à la synthèse protéique. En parallèle, le Sec-ARNt^{Sec} est délivré dans le site A du ribosome par un facteur d'élongation spécifique, SelB chez les bactéries et EF-Sec chez les eucaryotes et les archées.

Le Pyl-ARNt^{Pyl} est, quant à lui, pris en charge par le facteur d'élongation traditionnel des eucaryotes, EF-1 α mais également par EF-Tu, spécifique des bactéries (Mukai et al., 2008; Theobald-Dietrich et al., 2004).

4. La spécificité des aaRS

La synthèse protéique est primordiale pour la prolifération cellulaire et la formation des 20 aa-ARNt est l'étape limitante pour l'incorporation des aa suivant l'ordre dicté par l'ARNm. La tâche est ardue car de nombreuses espèces d'ARNt, mais également d'aa, sont structurellement très proches. La capacité des aaRS à reconnaître les bons substrats (leur aa et leurs ARNt homologues) est donc cruciale. Pour cela, au cours de l'évolution, ces enzymes ont acquis des domaines spécifiques permettant d'augmenter la fidélité de reconnaissance de leurs partenaires.

Le repliement 3D des ARNt est conservé et ils adoptent tous une structure en "L renversé" ce qui ne facilite pas leur reconnaissance. Ainsi, les aaRS reconnaissent leurs ARNt homologues *via* des signaux de reconnaissance spécifiques, présents tout au long de l'ARNt, appelés les déterminants d'identité. Ceux-ci sont néanmoins retrouvés préférentiellement au niveau des extrémités distales de l'ARNt (au niveau du bras accepteur, de la boucle de l'anticodon ou de la région variable dans le cas des ARNt reconnus par les aaRS de classe II). Le nombre de déterminant d'identité n'est pas homogène au sein des ARNt. En effet, la reconnaissance de la paire de bases G3-U70 au niveau de l'ARNt^{Ala} est suffisante pour son aminoacylation par l'alanyl-ARNt synthétase (AlaRS) (Hou and Schimmel, 1988), alors que l'identité de l'ARNt^{Gln} fait intervenir dix déterminants différents répartis dans le bras accepteur et dans la tige-boucle de l'anticodon (Rogers and Soll, 1988). Finalement, au sein des ARNt, on retrouve également des signaux négatifs d'identité, appelés anti-déterminants, qui empêchent la reconnaissance d'un ARNt hétérologue.

La reconnaissance de l'aa est plus difficile que celle de l'ARNt, car les aaRS doivent faire la différence entre des aa qui ne peuvent pas forcément être discriminés en fonction de leur encombrement stérique. C'est pour cela que certaines aaRS ont acquis des modules d'édition des aa. Ceux-ci sont présents dans dix aaRS, aussi bien dans des enzymes de classe I que de classe II, et permettent de corriger les erreurs d'activation d'un aa non homologue. Le

mécanisme d'édition peut avoir lieu avant transfert de l'aa sur l'ARNt (édition pré-transfert) et l'aaRS va alors directement hydrolyser l'aa~AMP mésactivé en présence ou en absence de l'ARNt. L'aa~AMP mésactivé peut également être libéré en solution et subir une hydrolyse non enzymatique (Yadavalli and Ibba, 2012). Toutefois, le mécanisme d'édition post-transfert est le plus souvent utilisé (7 des 10 aaRS possédant des domaines d'édition). Pour cela, l'aa-ARNt mésacylé peut être directement hydrolysé par le domaine d'édition après un réarrangement conformationnel qui fait glisser le bras accepteur de l'aa-ARNt du site actif "synthétique" vers le site actif d'édition. De plus, il peut y avoir libération l'aa-ARNt mésacylé et clivage de la liaison ester par réassociation avec l'aaRS ou par action d'une protéine d'édition autonome (non fusionnée à l'aaRS correspondante). C'est notamment le cas des protéines AlaX qui sont des homologues des domaines d'édition de l'AlaRS et qui corrigent spécifiquement les Ser-ARNt^{Ala} ou Gly-ARNt^{Ala} (Ahel et al., 2003). L'édition des aa mésactivés a été l'objet d'une revue exhaustive par Perona et Gruic-Sovulj en 2014 (Perona and Gruic-Sovulj, 2014).

Il est à noter que les facteurs d'élongation ont également la capacité de discriminer les aa-ARNt mésacylés des autres et sont capables de s'associer spécifiquement aux ARNt associés à l'aa homologue (LaRiviere et al., 2001). Les espèces mésacylées ont donc une faible probabilité d'être déposées au niveau du site A du ribosome et de ce fait, les erreurs d'incorporation sont rares. Cependant, l'incorporation d'un aa hétérologue peut être utilisée par la cellule pour lutter contre certains stress. Les détails sont décrits dans la partie suivante.

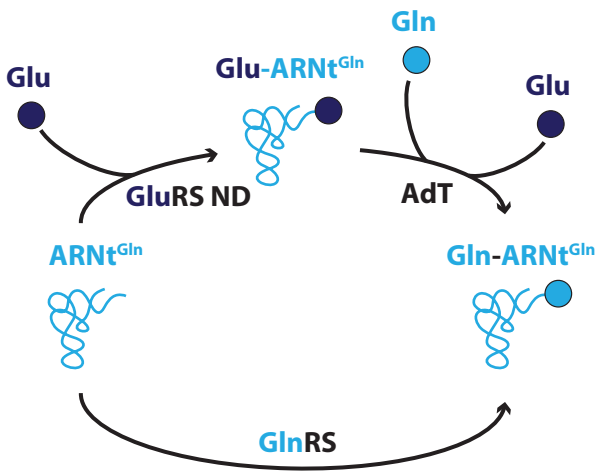
5. La mésincorporation des aa

Comme explicité précédemment, la fidélité de la traduction est primordiale pour la survie de la cellule et celle-ci dépend de l'association parfaite de l'aa à son ARNt homologue. Malgré la grande fidélité des domaines catalytique et d'édition des aaRS, l'aminoacylation d'un aa sur un ARNt hétérologue est possible. La mésincorporation d'aa lors de la synthèse protéique induit alors la formation de protéines mutées. Néanmoins, à faible dose, ces protéines de séquences dégénérées ne sont pas forcément néfastes pour l'organisme. L'incorporation de Met au niveau de codons autres que AUG, encore appelée misméthionylation, en est un exemple. Ce mécanisme est très important et conservé aussi bien chez les procaryotes que chez les eucaryotes. La MetRS d'*E. coli*, par exemple, est capable d'aminoacyler les ARNt^{Arg} et ARNt^{Thr} *in vitro*, alors que l'enzyme humaine mésacyle les ARNt^{Lys}.

Formation du Gln-ARNt^{Gln}

Voie indirecte

toutes les archées, la plupart des bactéries et les organelles



Voie directe

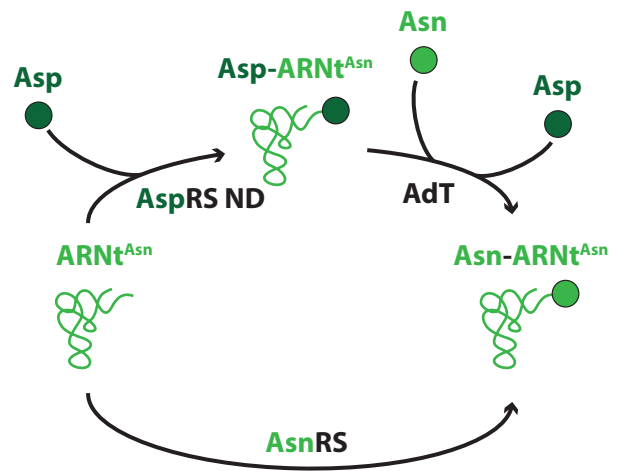
tous les eucaryotes et certaines bactéries

AdT : GatCAB, GatFAB, GatDE, GatAB

Formation de l'Asn-ARNt^{Asn}

Voie indirecte

toutes les archées, la plupart des bactéries



Voie directe

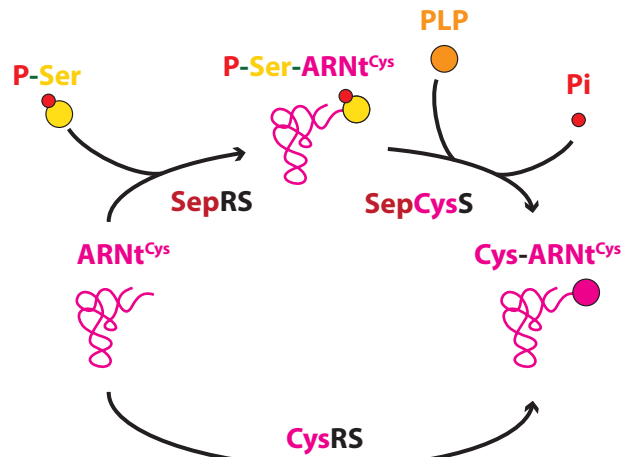
chez les eucaryotes, certaines bactéries et archées et les organelles

AdT : GatCAB, GatFAB, GatDE, GatAB

Formation du Cys-ARNt^{Cys}

Voie indirecte

les archées méthanogènes



Voie directe

chez les eucaryotes, bactéries, les organelles et la plupart des archées

Figure 10 : Les voies indirectes de formation des aa-ARNt^{aa}

La formation des aa-ARNt^{aa} *via* la voie indirecte fait intervenir une aaRS qui permet la formation d'un aa-ARNt^{aa} mésacylé qui est ensuite pris en charge par une seconde enzyme qui génère l'aa-ARNt^{aa} homologue. Pour l'Asp-ARNt^{Asn} et le Gln-ARNt^{Gln}, l'enzyme est une amidotransférase ARNT-dépendante (AdT) du type GatCAB, FAB, DE ou AB. La formation du Cys-ARNt^{Cys} fait, quant à elle, intervenir la O-phosphoséryl synthase (SepCysS) qui utilise le phosphate de pyridoxal (PLP) pour former l'aa-ARNt^{aa} homologue et relargue du phosphate inorganique (Pi).

Chez les mammifères, la méthylation est très fortement régulée et elle peut être bénéfique pour des cellules subissant un stress oxydatif (Netzer et al., 2009). Le taux de méthylation de Met est d'environ 1 % en condition basale, mais peut atteindre jusqu'à 10 % lors de l'activation du système immunitaire ou en réponse au stress oxydatif (Netzer et al., 2009). Des études suggèrent que les Met présentes à la surface ou à proximité du site actif peuvent être oxydées et protègent la protéine contre l'action délétère des ROS. La méthylation de Met à des endroits stratégiques permet d'augmenter cet effet de protection (Luo and Levine, 2009; Vogt, 1995). L'incorporation de cet aa au niveau de codons hétérologues doit donc être un mécanisme hautement régulé ; cependant, les détails moléculaires de celui-ci, ainsi que de la capacité du ribosome à incorporer les aa hétérologues, ne sont pas encore élucidés. En outre, la capacité des facteurs d'élongation à prendre en charge ces espèces méthyliées, spécifiquement en conditions de stress, reste à confirmer. Toutefois, il existe des évidences supportant cette hypothèse, notamment chez les champignons, où l'ARNt^{Ser} (CAG) peut être aminoacylé par la SerRS et minoritairement par la leucyl-ARNt synthétase (LeuRS). Les deux ARNt sont utilisés lors de la synthèse protéique et, de ce fait, pris en charge par le facteur d'élongation (Santos et al., 2011).

6. Les voies indirectes de formation des aa-ARNt

Pour assurer la synthèse protéique et l'incorporation de l'ensemble des aa dans les protéines néosynthétisées, il est nécessaire d'avoir 20 aaRS permettant la formation du jeu complet d'aa-ARNt et c'est le cas pour la majorité des organismes (Blattner et al., 1997; Fleischmann et al., 1995). Cependant certains organismes ne possèdent pas les gènes permettant l'expression des 20 aaRS. En effet, la majorité des bactéries, toutes les archées, ainsi que les organites des eucaryotes, sont dépourvus de une à quatre aaRS (Bult et al., 1996). Les enzymes catalysant la fixation de l'aa sur l'ARNt étant absentes, les aa-ARNt^{aa} correspondants sont formés par une voie alternative, encore appelée voie indirecte (**Figure 10**).

La synthèse de ces aa-ARNt passe par la formation d'intermédiaires méthyliés et nécessite l'intervention d'aaRS capables de reconnaître ces ARNt et de leur associer un aa hétérologue : on parle d'aaRS non discriminantes (aaRS ND). Quelle que soit l'aaRS manquante, l'aa hétérologue est toujours un précurseur métabolique de l'aa homologue.

C'est notamment le cas de l'Asp qui peut être modifié en Asn sur l'ARNt^{Asn} ou du Glu qui est converti en Gln sur l'ARNt^{Gln}. Pour former l'aa homologue, l'aa hétérologue est alors modifié par une enzyme de conversion ARNt-dépendante (une amidotransférase, dans le cas des Asp-ARNt^{Asn} et Glu-ARNt^{Gln}). Dans l'ensemble des espèces séquencées à ce jour, il existe quatre amidotransférase ARNt-dépendante (AdT) permettant la conversion des aa-ARNt mésacylés (Asp-ARNt^{Asn} et Glu-ARNt^{Gln}) : les complexes hétérotrimériques GatCAB et GatFAB, ainsi que les complexes hétérodimériques GatDE et GatAB.

Chez les eucaryotes, le Gln-ARNt^{Gln} cytosolique est formé par la glutaminyl-ARNt synthétase (GlnRS) qui aminoacyle l'ARNt^{Gln} avec de la Gln. Cependant, cette enzyme est absente chez la majorité des bactéries, toutes les archées et les organites des eucaryotes. Le Gln-ARNt^{Gln} est alors formé par une voie indirecte suivant deux étapes (Ibba and Soll, 2000; Wilcox and Nirenberg, 1968). Une GluRS ND aminoacyle tout d'abord l'ARNt^{Gln} avec du Glu, formant ainsi un intermédiaire mésacylé (Glu-ARNt^{Gln}). Le Glu présent sur l'ARNt^{Gln} est ensuite amidé en Gln par l'AdT GatCAB qui libère du Gln-ARNt^{Gln}. Au sein de la GatCAB, la s-u GatC est une protéine structurale qui maintient les deux autres, la s-u GatA possède une activité glutaminase qui hydrolyse de la Gln libre en Glu et NH₂. La s-u GatB, quant à elle, est une protéine bifonctionnelle possédant une activité kinase, permettant la phosphorylation du groupement γ -COOH du Glu, et une activité transamidase, responsable de l'amidation du Glu-ARNt^{Gln} en Gln-ARNt^{Gln}, grâce au NH₂ produit par GatA.

Il est à noter que chez la levure, la GlnRS mitochondriale étant absente, c'est également une voie indirecte qui catalyse la formation Gln-ARNt^{Gln} mitochondrial (Gln-mARNt^{Gln}). Cependant, l'AdT n'est plus la GatCAB car la s-u GatC est absente et remplacée par un homologue, GatF, possédant une extension N-terminale (N-ter) absente chez GatC. Cette extension joue probablement un rôle dans la modulation de la formation de l'AdT et participe à la formation du site de dé-amidation de la s-u GatA. C'est donc le complexe hétérotrimérique GatFAB qui participe à la voie indirecte de formation du Gln-mARNt^{Gln} de levure (Araiso et al., 2014; Frechin et al., 2009b). Chez l'Homme, la voie de formation du Gln-mARNt^{Gln} fait intervenir la GluRS mitochondriale qui mésacyle l'mARNt^{Gln} et la modification de cet aa-ARNt mésacylé est catalysée par l'AdT GatCAB pour générer le Gln-mARNt^{Gln} (Nagao et al., 2009). Chez les plantes, la situation est identique, l'absence de la GlnRS mitochondriale et chloroplastique nécessite l'import d'une GluRS ND dans ces deux compartiments. Que ce soit

au niveau de la mitochondrie ou du chloroplaste, l'aa-ARNt mésacylé est le substrat d'une AdT GatCAB doublement localisée, permettant de générer les deux Gln-ARNt^{Gln} organellaires (Pujol et al., 2008).

En parallèle, la quasi-totalité des archées et quelques bactéries sont également dépourvues de l'asparaginyl ARNt synthétase (AsnRS) et l'Asn-ARNt^{Asn} est synthétisé *via* une voie indirecte faisant intervenir une aspartyl-ARNt synthétase non discriminante (AspRS ND), formant de l'Asp-ARNt^{Asn} et la GatCAB, amidant cet aa-ARNt en Asn-ARNt^{Asn} (Becker and Kern, 1998; Becker et al., 2000; Curnow et al., 1996). De manière intéressante, les GatCAB bactériennes sont doubles spécifiques, elles sont donc capables de transamider, aussi bien l'Asp-ARNt^{Asn} que le Glu-ARNt^{Gln}, comme par exemple chez *Helicobacter pylori* qui est dépourvu à la fois de GlnRS et d'AsnRS (Sheppard et al., 2007). Cette double spécificité est d'ailleurs conservée chez les GatCAB des organismes dont une seule des deux enzymes est absente. Chez les archées, il existe également l'AdT GatDE, dont la s-u GatD possède l'activité glutaminase et GatE, homologue de GatB, possède les activités kinase et amidotransférase (Tumbula et al., 2000). Cependant, celle-ci est mono-spécifique et ne convertit que le Glu-ARNt^{Gln} en Gln-ARNt^{Gln}. La GatCAB archéenne est alors retrouvée uniquement chez les espèces dépourvues d'AsnRS et elle est mono-spécifique, ne convertissant que l'Asp-ARNt^{Asn} en Asn-ARNt^{Asn} (Namgoong et al., 2007; Roy et al., 2003). Récemment, la voie indirecte de formation du Gln-ARNt^{Gln} a été décrite chez *Plasmodium falciparum* et celle-ci fait intervenir une AdT apicoplastique et une GluRS ND (Mailu et al., 2015; Mailu et al., 2013). L'AdT est formée des s-u GatA et GatB, spécifiquement importées dans l'apicoplaste, et est impliquée dans la formation du Gln-ARNt^{Gln}, indispensable lors de l'étape érythrocytaire du parasite.

Finalement, une voie indirecte permet également la formation du Cys-ARNt^{Cys} chez un certain nombre d'archées méthanogènes dépourvues de cystéinyln-ARNt synthétase (CysRS) (Stathopoulos et al., 2001). Le Cys-ARNt^{Cys} est formé par l'intervention de la O-phosphoséryln-ARNt synthétase (SepRS) et la Sep-ARNt:Cys-ARNt synthase (SepCysS). Tout d'abord, la SepRS aminoacyle l'ARNt^{Cys} avec une O-phosphosérine qui est ensuite convertie par la SepCysS en Cys pour former le Cys-ARNt^{Cys}, en présence d'un donneur de sulfate et de phosphate de pyridoxal (Sauerwald et al., 2005) (**Figure 10**).

7. Les fonctions additionnelles des aaRS

La fonction primaire des aaRS consiste en la fixation d'un aa sur leurs ARNt homologues qui sont ensuite pris en charge par le facteur d'initiation eIF2 (pour le Met-ARNt^{Met}) ou par le facteur d'élongation eEF1A (pour les aa-ARNt élongateurs). Néanmoins, de plus en plus d'études soulignent la capacité de ces enzymes à remplir d'autres fonctions en dehors de la traduction, et ce, quel que soit l'organisme. Ces nouvelles fonctions couvrent un large spectre d'activités, comme l'expression de gènes, la régulation de la transcription, mais également la régulation de la synthèse protéique et la réponse à des stress.

a. Les fonctions additionnelles des aaRS bactériennes

Chez les bactéries, les fonctions additionnelles des aaRS concernent principalement la régulation de l'expression des gènes. L'AlaRS d'*E. coli* permet la régulation de la transcription de son propre gène (*AlaS*) par sa fixation à une séquence spécifique proche du codon d'initiation. Lorsque la concentration en Ala est élevée, l'AlaRS va se fixer à cette région pour réprimer la transcription d'*AlaS* (Putney and Schimmel, 1981). De manière comparable, la thréonyl-ARNt synthétase (ThrRS) d'*E. coli* est impliquée dans la régulation de la traduction de son propre gène. En effet, la région 5'UTR de son ARNm se replie en un mime structural de l'ARNt^{Thr}. La ThrRS se fixe alors à cette région et empêche la fixation du ribosome, bloquant ainsi la traduction (Brunel et al., 1993; Ehresmann et al., 1988; Moine et al., 1988; Romby et al., 1996). Finalement, un autre exemple de fonction non canonique a été mis en évidence chez les bactéries et plus particulièrement chez *Deinococcus radiodurans*. Chez cet organisme, suite à un dommage causé par des radiations, la tryptophanyl-ARNt synthétase II (TrpRS II) est exprimée et se lie à l'oxyde nitrique synthase. Ce dimère permet alors la nitration du Trp et la formation de 4-nitro-Trp, qui joue potentiellement un rôle lors de stress cellulaire (Buddha et al., 2004a; Buddha et al., 2004b).

b. Les fonctions additionnelles des aaRS de champignons

Les fonctions additionnelles des aaRS ne sont pas restreintes à l'enzyme cytosolique et certaines aaRS mitochondriales remplissent ce type de fonctions. C'est notamment le cas de la tyrosyl-ARNt synthétase (TyrRS) mitochondriale de *Neurospora crassa* et *Podospora anserina* ou encore la LeuRS mitochondriale de *S. cerevisiae* qui interviennent dans l'épissage

des introns de groupe I (Akins and Lambowitz, 1987; Herbert et al., 1988; Paukstelis and Lambowitz, 2008). Ces enzymes organellaires ont donc évolué pour favoriser l'émergence de ces nouvelles fonctions. En effet, des études sur l'épissage catalysé par la LeuRS mitochondriale ont montré que cette activité est médiée par le domaine d'édition CP1 (*Connective Peptide 1*) de cette enzyme (Herbert et al., 1988; Rho et al., 2002; Sarkar et al., 2012). Chez *S. cerevisiae* également, lors de la croissance à 37 °C, le précurseur de la LysRS mitochondriale permet, en association avec l'Enolase 2, l'import de l'ARNt^{Lys1} cytosolique (tRK1) dans la mitochondrie (Baleva et al., 2015; Gowher et al., 2013). Chez ce même organisme, les aaRS peuvent également intervenir dans certaines voies de signalisation déclenchées par un stress. C'est notamment le cas de la SerRS qui interagit avec la protéine Pex21, impliquée dans la formation du peroxysome (Godinic et al., 2007; Rocak et al., 2002). La formation de ce dimère induit (i) une augmentation de liaison de la SerRS avec l'ARN^{Ser}, (ii) une augmentation de l'activité d'aminacylation de la SerRS et (iii) potentiellement la coordination des voies de signalisation du stress (Godinic et al., 2007).

Le complexe TORC1 (*Target Of Rapamycin Complex 1*), structurellement et fonctionnellement conservé à travers l'échelle évolutive, joue un rôle primordial dans le contrôle de la prolifération cellulaire (synthèse protéique, autophagie...) en réponse aux hormones, facteurs de croissance ou aux quantités d'aa disponibles (Leu par exemple). Une étude récente a mis en évidence la capacité de la LeuRS à s'associer avec Gtr1, un activateur du complexe TORC1 (Bonfils et al., 2012). En présence de fortes concentrations en Leu, la LeuRS permet la formation du Leu-ARNt^{Leu} et s'associe à Gtr1 *via* son domaine d'édition CP1 et joue alors un rôle de senseurs d'aa pour le complexe TORC1.

D'autres fonctions additionnelles sont décrites pour des aaRS de champignons. Cependant, ces fonctions sont tributaires de la capacité des aaRS à s'assembler en complexe multi-synthétasique (MSC : *MultiSynthetase Complex*) ou d'en être relarguées. Ces fonctions additionnelles seront décrites dans le § VI.4 de l'introduction et complétées lors de la première partie de mes résultats.

c. Les fonctions additionnelles des aaRS de mammifères

Les fonctions additionnelles des aaRS qui ont été le mieux caractérisées sont celles des aaRS de mammifères dont la quasi-totalité possède une ou des fonctions non canoniques.

Uniquement la CysRS, la phénylalanine-ARNt synthétase (PheRS), l'isoleucyl-ARNt synthétase (IleRS) et l'arginyl-ARNt synthétase (ArgRS) n'ont pas encore de fonctions additionnelles identifiées à ce jour (Yakobov et al., in prep). Les aaRS des mammifères ont acquis au cours de l'évolution des domaines spécifiques permettant (i) leur assemblage sous forme de MSC (voir introduction § VI.1) et (ii) la catalyse de leurs fonctions additionnelles. C'est notamment le cas de la TrpRS, possédant un domaine WHEP qui est impliqué dans la régulation de l'angiogénèse (Kise et al., 2004; Otani et al., 2002; Wakasugi and Schimmel, 1999; Wakasugi et al., 2002). Deux domaines spécifiques sont présents chez la TyrRS des eucaryotes supérieurs et absents des TyrRS des autres eucaryotes, des archées ou des bactéries. Le premier, appelé tripeptide ELR (Glu-Leu-Arg), est situé au sein du domaine catalytique et le second, le domaine EMAPII (*Endothelial Monocyte-Activating Polypeptide*), est retrouvé dans la région C-terminale (C-ter). Ces deux domaines sont impliqués dans la fonction angiogénique de la TyrRS. En effet, au sein de la structure 3D, le domaine EMAPII bloque l'accessibilité du motif ELR et permet ainsi d'inhiber l'activité de la TyrRS au niveau de l'angiogénèse.

Comme pour les fonctions additionnelles des aaRS de champignons, de nombreuses autres fonctions non canoniques des aaRS ont été répertoriées, et celles-ci sont directement liées à la capacité de ces enzymes à former des MSC. Ces complexes, en plus de leur rôle dans la synthèse protéique, influent donc également fortement sur les activités additionnelles des aaRS de par la modulation de leur libération en réponse à différents stimuli. Ces fonctions sont décrites dans le § VI.4.

8. La localisation subcellulaire des aaRS cytosoliques

Au sein de chaque organisme, un jeu complet d'aaRS doit être présent dans chaque compartiment traductionnellement actif. Chez les procaryotes, ceci ne pose pas de problème, puisque le cytosol est le seul compartiment où se déroule la traduction et un seul jeu d'aaRS est suffisant. Cependant, comme je l'ai décrit dans la partie traitant des voies indirectes d'aminocyclation, certains gènes codant des aaRS sont absents du génome et c'est une aaRS ND qui catalyse la formation de l'aa-ARNt^{aa} orphelin *via* une voie indirecte. Chez les eucaryotes, la compartimentalisation induit la nécessité de produire plusieurs jeux d'aaRS, un pour le cytosol et un pour les organites traductionnellement actifs. Les génomes organellaires (celui de la mitochondrie, du chloroplaste ou de l'apicoplaste) sont des génomes minimaux

qui ne codent que pour quelques protéines, mais pas pour les aaRS retrouvées dans leurs compartiments ; celles-ci sont donc toutes codées par le génome nucléaire. Ainsi, ce dernier permet la synthèse de deux ou même trois jeux d'aaRS nécessaires à la production des aa-ARNt^{aa}, utilisés par l'ensemble des machineries traductionnelles de la cellule.

Malgré la présence de plusieurs jeux d'aaRS, il n'existe, dans certains cas, qu'un seul gène codant pour une aaRS donnée. Le gène présent code alors pour les deux ou trois versions de l'enzyme (cytosolique et organellaire). Dans ce cas, la seule différence entre les deux variants est la présence d'une séquence d'import au niveau de la version organellaire. Chez la levure, c'est le cas de l'AlaRS, l'histidyl-ARNt synthétase (HisRS) et la valyl-ARNt synthétase (ValRS) (Chatton et al., 1988; Huang et al., 2006; Natsoulis et al., 1986). De plus, la glycyL-ARNt synthétase (GlyRS) est codée par deux gènes différents (*GRS1* et *GRS2*). Néanmoins, la suppression du gène de la *GRS2* n'est létale, ni en fermentation, ni en respiration. En parallèle, deux isoformes sont produits à partir du gène *GRS1* dont l'un est plus long de 23 aa ; ces résidus correspondant probablement au signal d'adressage mitochondrial (Chang and Wang, 2004; Turner et al., 2000).

Chez l'Homme, il y a uniquement deux gènes, *GARS* et *KARS*, codant respectivement pour la GlyRS et la LysRS, dont la version mitochondriale est codée par le même gène (Alexandrova et al., 2015; Chien et al., 2014; Tolkunova et al., 2000). La version cytosolique est alors produite par épissage alternatif (LysRS) ou par un codon d'initiation de la traduction alternatif (GlyRS), permettant l'élimination du signal d'adressage dans la version non importée (cytosolique). Chez *Trypanosoma brucei*, quatre aaRS cytosoliques et organellaires sont également codées par le même gène (l'IleRS, la GluRS, la GlnRS et la prolyl-ARNt synthétase ou ProRS) et la synthèse de l'aaRS cytosolique est catalysée par un mécanisme d'épissage en trans (Cestari et al., 2013; Rettig et al., 2012; Rinehart et al., 2004). La relocalisation des aaRS cytosoliques a également été étudiée chez le parasite intracellulaire *Plasmodium falciparum*. Chez cet organisme, il y a uniquement seize aaRS cytosoliques, quinze sont retrouvées dans l'apicoplaste (relique du chloroplaste) et quatre sont partagées entre l'apicoplaste et le cytosol. On retrouve ainsi l'AlaRS, la GlyRS et la ThrRS dont leurs versions cytosoliques sont produites par un mécanisme d'initiation alternative de la traduction, ou encore la CysRS, où les deux enzymes sont générées par épissage alternatif (Jackson et al., 2012; Pham et al., 2014).

Chez les plantes, il semblerait que les gènes codant les aaRS organellaires et nucléaire soient différents et la double localisation des aaRS mitochondriales et chloroplastiques est probablement un mécanisme conservé. En effet, chez *Arabidopsis thaliana*, sur les 24 aaRS organellaires identifiées, 15 (probablement 17) sont présentes dans les deux organelles. De plus, cinq aaRS cytosoliques (la LeuRS, l'AlaRS, la ThrRS, la GlyRS et la ValRS) sont doublement localisées et importées dans la mitochondrie. En parallèle, il n'y en a aucune qui est uniquement mitochondriale et deux qui semblent être uniquement chloroplastiques. Finalement, il n'y a qu'un seul gène codant la GlnRS suggérant que cette enzyme assure une fonction d'aminocyclation dans les trois compartiments (Duchene et al., 2005; Duchene et al., 2001; Mireau et al., 1996; Souciet et al., 1999).

En parallèle des gènes communs permettant la formation de la version cytosolique et organellaire des aaRS, on retrouve également la relocalisation d'aaRS pour compenser l'absence d'une aaRS organellaire. L'exemple le plus commun est la compensation de l'absence de la GlnRS organellaire. Du fait de l'absence de cette enzyme, la formation du Gln-ARNT^{Gln} dans ces compartiments est catalysée par une GluRS ND et l'intervention d'une AdT spécifique (Frechin et al., 2009a). Chez la levure, la formation de cet aa-ARNT est médiée par l'import de la GluRS cytosolique dans la mitochondrie. Celle-ci va alors mésacyler l'ARNT^{Gln} avec le Glu et le Glu-ARNT^{Gln} va finalement être modifié en Gln-ARNT^{Gln} *via* une voie indirecte faisant intervenir l'AdT GatFAB (Frechin et al., 2009b). Cependant, l'import de la GluRS cytosolique humaine ne semble pas conservé car l'enzyme responsable de la formation du Glu-mARNT^{Gln} est la GluRS mitochondriale (Nagao et al., 2009). Chez la levure, il reste néanmoins des aa-ARNT^{aa} mitochondriaux dont la voie de biogénèse n'est pas encore élucidée. La formation du Pro-mARNT^{Pro} n'a pas été confirmée, mais il existe toutefois un gène putatif codant pour cette enzyme, identifié par alignement de séquence. Cependant, personne n'a confirmé, que ce soit *in vitro* ou *in vivo*, l'activité de cette enzyme. De plus, la CysRS mitochondriale est absente, la synthèse du Cys-mARNT^{Cys} pourrait donc être catalysée par la CysRS cytosolique qui serait potentiellement importée dans la mitochondrie ; mais elle pourrait également faire intervenir une voie indirecte comme c'est le cas chez plusieurs organismes. Ce mécanisme doit cependant être étudié dans les détails pour comprendre la voie de formation de cet aa-ARNT.

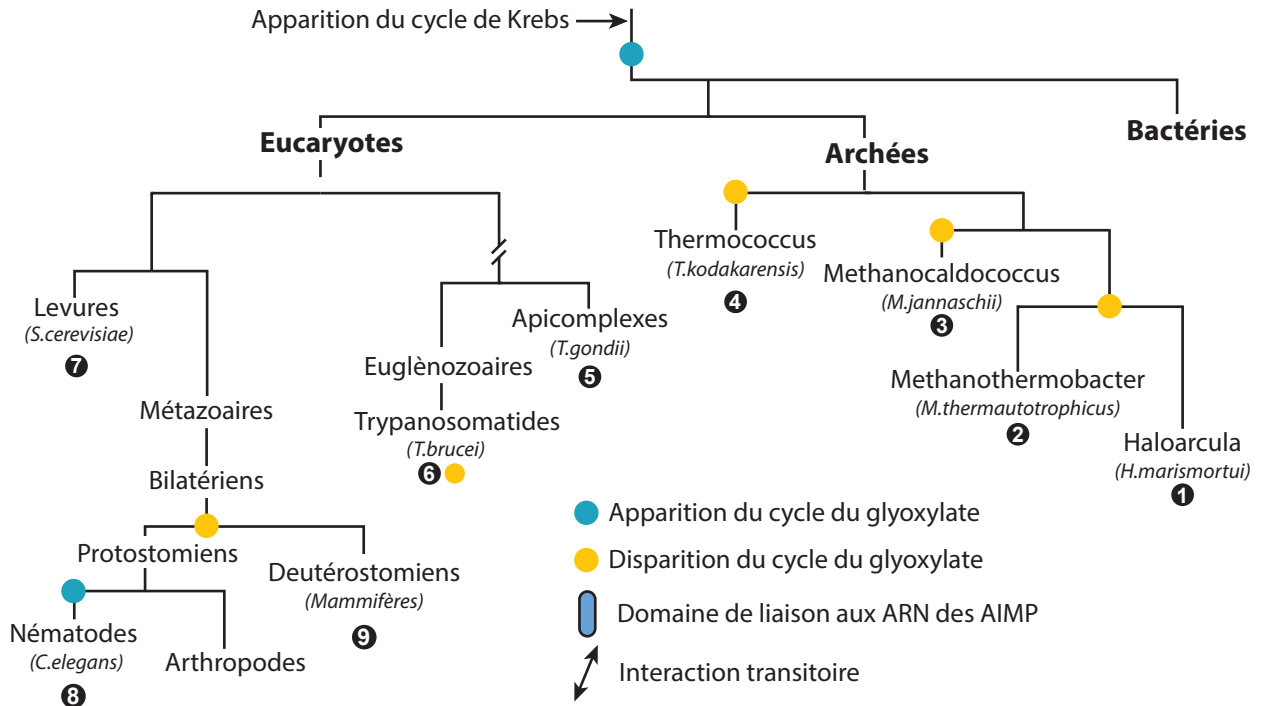
En plus de leurs localisations mitochondriales et chloroplastiques, les aaRS cytosoliques peuvent également se relocaliser dans d'autres compartiments pour y assurer l'aminocyclation d'ARNt ou des fonctions additionnelles. Le noyau est un lieu de relocalisation privilégié pour ces enzymes. Chez l'Homme, on retrouve dans ce compartiment la MetRS, la TyrRS, la TrpRS, la LysRS et la PheRS. Toutes ces enzymes se relocalisent suite à des stimuli spécifiques et sont, soit dissociées du complexe MARS en réponse au stress (MetRS) ou suite à une modification post traductionnelle (phosphorylation pour la LysRS), soit ce sont des aaRS libres qui se relocalisent en réponse aux stress (TrpRS et TyrRS) (Fu et al., 2012; Gunasekera et al., 2004; Ko et al., 2000; Mirande et al., 1985; Paley et al., 1991; Popenko et al., 1993; Sajish et al., 2012; Wei et al., 2014). Chez *S. cerevisiae*, deux aaRS se relocalisent dans le noyau, la MetRS et la TyrRS. La MetRS, retrouvée au sein du MSC présent chez cet organisme, se relocalise suite à la libération du complexe (Galani et al., 2001). La TyrRS, quant à elle, n'est pas retrouvée dans un MSC et est libre dans le cytosol. La relocalisation de cette enzyme pourrait jouer un rôle dans l'export des ARNt aminoacylés (Azad et al., 2001).

De plus, une étude bio-informatique sur les signaux d'adressage nucléaires des aaRS de levure suggère que la quasi-totalité de ces enzymes est capable de se relocaliser dans ce compartiment (Bader et al., 2017; Schimmel and Wang, 1999). Ces enzymes sont potentiellement impliquées dans l'aminocyclation des ARNt dans ce compartiment, permettant ainsi la finalisation de leur maturation. En parallèle, ces aaRS pourraient également assurer des fonctions nucléaires additionnelles, mais la capacité des aaRS cytosoliques à se relocaliser dans le noyau et leurs potentielles fonctions nucléaires restent encore à être élucidées.

VI. Les complexes multi-synthétasiques

Les aaRS, dont la fonction principale est de synthétiser leurs aa-ARNt homologues, ont drastiquement évolué pour donner naissance aux enzymes que l'on connaît à l'heure actuelle. En effet, il est maintenant admis que les aaRS dites primitives étaient réduites au seul domaine catalytique capable d'activer l'aa (Rould et al., 1989). Ce n'est que plus tard que les aaRS primitives ont acquis des domaines spécialisés fixant la mini-hélice retrouvée chez les ARNt primitifs (Schimmel and Ribas de Pouplana, 1995). Il semblerait également que les ARNt et les aaRS aient co-évolué par l'acquisition de la mini-hélice de l'anticodon sur l'ARNt et du domaine

A



B

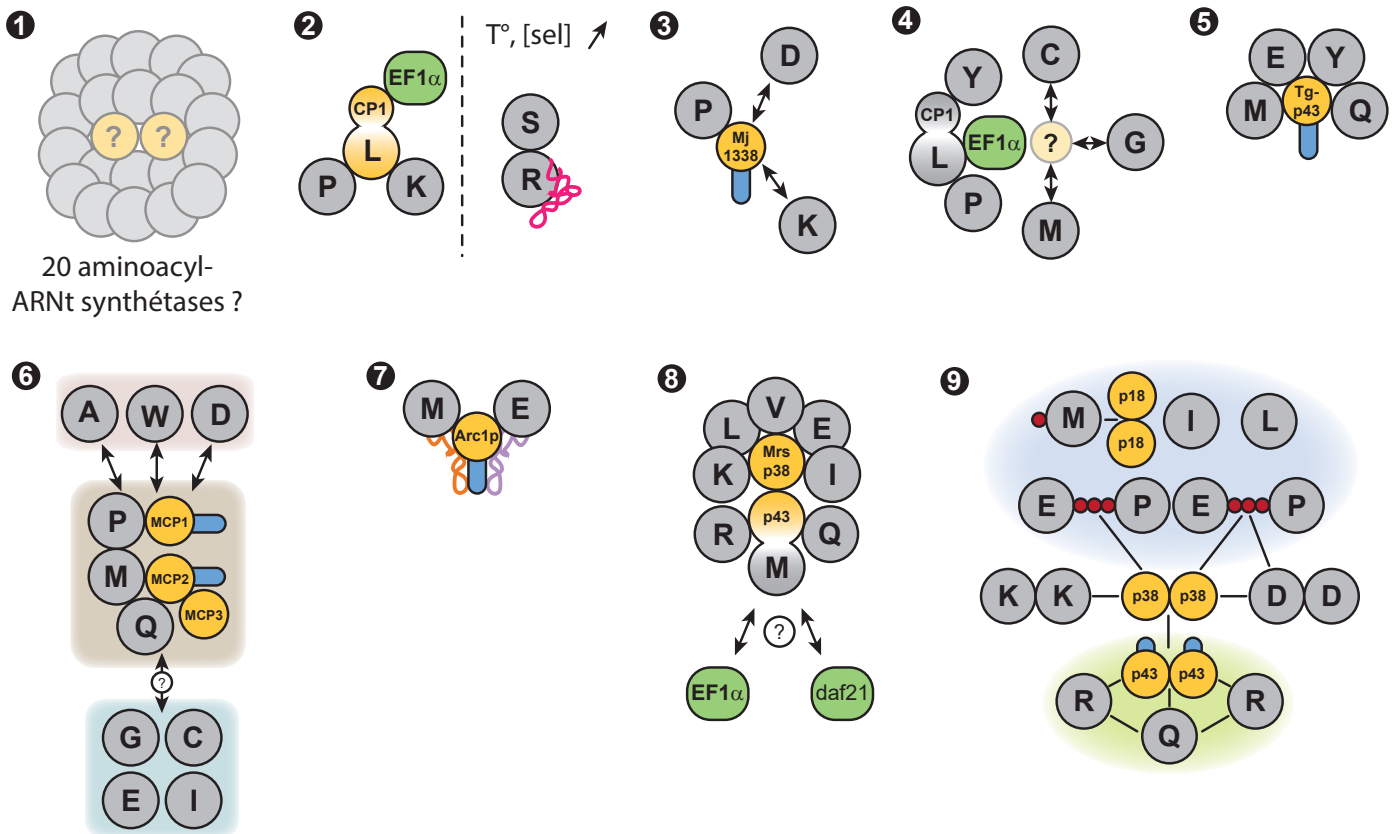


Figure 11 : Les MSC au cours de l'évolution

(A) Arbre phylogénétique comprenant les espèces où un MSC a été identifié à ce jour. (B) Schéma de la composition et l'organisation des MSC. Ici, le code à une lettre est utilisé pour chaque aaRS. Chaque chiffre retrouvé en (A) correspond à un MSC en (B).

de reconnaissance de cet anticodon au niveau de l'aaRS, augmentant ainsi l'efficacité et la spécificité de reconnaissance des deux partenaires. De plus, des domaines permettant d'augmenter la fidélité de la traduction ont également été acquis, notamment les différents domaines d'édition, favorisant la correction des aa-ARNt mésacylés. Au cours de l'évolution des aaRS, ces enzymes se sont complexifiées davantage pour acquérir des domaines spécialisés, n'intervenant pas forcément dans leur fonction primaire d'aminocyclation, mais leur permettant (i) de s'associer avec des partenaires traductionnels (EF...) ou n'intervenant pas dans la traduction, ou (ii) d'assurer des fonctions additionnelles non traductionnelles, tout aussi indispensables que leur fonction d'aminocyclation. Dans certains cas, ces domaines favorisent l'interaction entre les aaRS et sont impliqués dans la formation de complexes dits multi-synthétasiques (MSC). Ces domaines sont apparus en parallèle de la complexification des organismes et leur ont permis d'augmenter leur vitesse de prolifération, mais également leur résistance à certains stress ou leur capacité à se défendre contre les menaces extérieures.

C'est dès le début des années 1970 que l'hypothèse selon laquelle les aaRS s'associent sous forme de complexe a émergé au sein de la communauté scientifique. En effet, les premières expériences de co-sédimentation réalisées sur des cellules de foie de rats suggèrent que les aaRS peuvent former des complexes hautement structurés au sein de la cellule et ne sont donc pas retrouvées sous forme solubles (Bandyopadhyay and Deutscher, 1971). De plus, l'association des aaRS sous forme de complexes de haut poids moléculaires est également décrite chez la levure (Simos et al., 1996), chez la bactérie *E. coli* (Harris, 1987, 1990) et chez les archées (Praetorius-Ibba et al., 2007). Des études ont également montrés que certaines protéines de la machinerie traductionnelle sont capables de s'associer de manière transitoire avec des polysomes ou des filaments d'actine (Barbarese et al., 1995; Kaminska et al., 2009a; Kaminska et al., 2009b; Negrutskii et al., 1994). Toutes ces études suggèrent que les aaRS assurent un rôle central dans la cellule de par leur interaction au sein de MSC et cela dans toutes les branches du vivant. Malgré la conservation des fonctions des MSC au sein de l'arbre phylogénétique, la composition, ainsi que le mode d'assemblage de ces complexes varient fortement en fonction des espèces. Ces variations ont fait l'objet d'une revue dont je suis co-auteur intitulée : "*Exploring the evolutionary diversity and assembly modes of multi-aminoacyl-tRNA synthetase complexes: lessons from unicellular organisms*" (voir **Annexe 1**). De plus, cette variabilité est présentée dans la **Figure 11**.

	<i>S.cerevisiae</i>	<i>C.elegans</i>	<i>D.melanogaster</i>	<i>D. rerio</i>	<i>H.sapiens</i>
AlaRS					
ArgRS		LZ	LZ	LZ	LZ
AspRS	N-Hélice	N-Hélice	N-Hélice	N-Hélice	N-Hélice
AsnRS	N-Hélice, UNE-N	N-Hélice, UNE-N	N-Hélice, UNE-N	N-Hélice, UNE-N	N-Hélice, UNE-N
CysRS	UNE-C ₁ , UNE-C ₂	UNE-C ₁ , UNE-C ₂	UNE-C ₁ , UNE-C ₂	UNE-C ₁ , UNE-C ₂	UNE-C ₁ , UNE-C ₂ , GST
GlnRS	UNE-Q	UNE-Q	UNE-Q	UNE-Q	UNE-Q
GlyRS		WHEP	WHEP	WHEP	WHEP
HisRS		WHEP	WHEP	WHEP	WHEP
IleRS				UNE-I	UNE-I
LeuRS		UNE-L	UNE-L	UNE-L	UNE-L
LysRS	N-Hélice	N-Hélice	N-Hélice	N-Hélice	N-Hélice
MetRS	GST	GST, LZ, EMAPII	GST, WHEP	GST, WHEP	GST, WHEP
PheRS	UNE-F	UNE-F	UNE-F	UNE-F	UNE-F
SerRS				UNE-S	UNE-S
ThrRS	UNE-T	UNE-T	UNE-T	UNE-T	UNE-T
TrpRS				WHEP	WHEP
TyrRS			ELR, EMAPII	ELR, EMAPII	ELR, EMAPII
ValRS				GST	GST
GluRS	GST	GST, WHEP	GST, WHEP	GST, WHEP	GST, WHEP
ProRS	-----	----- WHEP			
p18			GST	GST	GST
p38			LZ, GST	LZ, GST	LZ, GST
p43			LZ, EMAPII	LZ, EMAPII	LZ, EMAPII
Arc1	GST, EMAPII				

Figure 12 : Domaines et motifs retrouvés chez les aaRS

Conservation des domaines et motifs intervenant dans l'interaction entre aaRS et AIMP et leurs fonctions additionnelles. GST : *Glutathion S-transférase* ; LZ : *Leucine Zipper* ; WHEP : Domaine WRS-HRS-EPRS ; EMAPII : *Endothelial Monocyte-Activating Polypeptide II* ; ELR : Domaine Glu-Leu-Arg ; UNE : Domaine Unique ; N-Hélice : Domaine N-ter en hélice. Les pointillés indiquent la fission entre la GluRS et ProRS chez *S. cerevisiae* et *C. elegans*.

1. Les domaines d'interaction retrouvés chez les aaRS

L'assemblage sous forme de MSC est médié, soit par des interactions directes entre aaRS, soit par l'établissement d'interactions entre des aaRS et des protéines d'assemblage (AIMP : *Aminoacyl tRNA synthetase complex Interacting Multifunctional Protein*). Néanmoins, de nombreux MSC identifiés ne contiennent pas de protéines d'assemblage ; la formation de ces complexes est donc possible uniquement par des interactions aaRS/aaRS. Les différentes interactions intermoléculaires permettant la formation de MSC sont assurées par des domaines apparus au cours de l'évolution, présents chez les aaRS et les AIMP. Ceux-ci peuvent être dédiés aux interactions protéine/protéine (*GST-like*), mais peuvent également avoir une fonction toute autre (domaine CP1, WHEP).

a. Les domaines d'interaction protéine/protéine

Les domaines *GST-like* (*Glutathion S-transferase-like*) sont retrouvés au niveau d'aaRS de nombreuses espèces différentes et sont souvent conservés de par leur distribution. En effet, le domaine *GST-like* de la GluRS est présent chez la levure, le nématode, mais également la mouche et les mammifères. Il en va de même pour celui de la MetRS et celui des AIMP permettant l'assemblage des MSC (**Figure 12**). Chez *S. cerevisiae*, c'est l'interaction entre les domaines *GST-like* des trois protéines du MSC qui favorise son assemblage. Chez les mammifères, le complexe est formé de neuf aaRS et quatre d'entre elles possèdent un domaine *GST-like* (**Figure 12**).

Les domaines WHEP sont présents uniquement chez les eucaryotes supérieurs. Leur nom vient de la présence de ces domaines dans la TrpRS (W), l'HisRS (H), la GluRS (E) et la ProRS (P) et correspond à une succession de 46 aa organisés en un motif *Helix-Turn-Helix* (HTH). Cependant, ces domaines sont également retrouvés chez la GlyRS et la MetRS. Les domaines WHEP sont présents de manière préférentielle au niveau de la région N-ter (HisRS, GlyRS et TrpRS), mais peuvent être retrouvés en C-ter (MetRS chez l'Homme et GluRS chez *C. elegans*) et en position centrale (fusion entre la GluRS et la ProRS). Ces domaines sont capables de lier l'ARN, mais également des protéines (voir introduction § VI.4.b). Néanmoins, ces domaines d'interaction ne semblent pas être impliqués préférentiellement dans la formation des MSC, mais plutôt dans la régulation de certaines fonctions non canoniques des aaRS.

Les glissières à Leucine (*leucine zipper* : LZ) sont des domaines d'interaction protéines/protéines caractérisés par la présence d'une Leu en quatrième position au niveau d'une répétition de sept résidus. Ces domaines sont présents dans environ 10 % des protéines eucaryotes. Toutes les aaRS possédant des LZ sont présentes au sein de MSC et ce motif induit leur incorporation dans ces complexes (**Figure 12**).

Il existe également d'autres domaines d'interactions entre protéines chez les aaRS mais ces motifs sont bien moins caractérisés. C'est le cas de l'extension C-ter de l'IleRS qui interagit avec les domaines WHEP de la GluProRS (Rho et al., 1996). Une extension C-ter similaire est retrouvée chez la LeuRS et permet de stabiliser son interaction avec l'ArgRS au sein du complexe (Ling et al., 2005). On retrouve également des domaines uniques, spécifiques d'une seule aaRS, nommés UNE-X (UNique Element) ou X correspond au code à une lettre de l'aa homologue de l'aaRS dans laquelle ce domaine est retrouvé (**Figure 12**). Au sein de la SerRS, il existe une séquence UNE-S qui joue le rôle de signal d'import nucléaire et est indispensable au développement cardio-vasculaire (Fukui et al., 2009; Herzog et al., 2009). Cependant, les fonctions de la majorité des séquences UNE-X ne sont pas encore bien caractérisées.

b. Les domaines d'interaction protéine/ARN

Les extensions N-ter sous forme d'hélices (N-hélice) sont des régions absentes des aaRS procaryotiques et présentes chez la LysRS, l'AspRS et l'AsnRS de tous les eucaryotes. Chacune de ces aaRS contient dans sa région N-ter, une hélice de taille variable (20 à 40 aa) ayant la capacité de fixer leur ARNt homologue, mais également différentes espèces d'ARNt. La N-hélice de la LysRS est impliquée dans la fonction non canonique de cette aaRS et permet, lors de l'infection par le VIH, l'incorporation de l'ARNt^{Lys} au sein de la particule virale.

Les domaines EMAPII (*cytokine-Endothelial Monocyte-Activating Polypeptide II*) sont des domaines retrouvés uniquement chez les aaRS dont la TyrRS chez les mammifères et la MetRS chez *C. elegans*, ainsi que chez les AIMP des MSC : p43 chez les mammifères et Arc1 chez la levure (Kaminska et al., 2000; Kim et al., 2000; Quevillon et al., 1997). Ce domaine EMAPII est un homologue du domaine Trbd111 (*tRNA-binding domain*) bactérien qui est un domaine de liaison à l'ARNt (Renault et al., 2001). Il favorise l'activité catalytique des aaRS au sein du MSC de levure, mais également celle de la MetRS lorsqu'il lui est fusionné

artificiellement (Karanasios et al., 2008). Finalement, ce domaine d'interaction à l'ARN semble avoir évolué pour agir en tant que cytokine (Shalak et al., 2001).

2. Les MSC chez les unicellulaires

a. Les MSC chez les bactéries

Chez les bactéries, la formation d'un MSC a été mise en évidence chez *E. coli* (Harris, 1987). De manière surprenante, la taille de ce complexe varie de 400 kDa à 1000 kDa en fonction de la méthode d'obtention de l'extrait cellulaire utilisé, et la taille de ce complexe ne correspond pas à celle que l'on peut déduire s'il était constitué de l'ensemble des 20 aaRS. Cette divergence de poids moléculaire suggère cependant que les interactions entre les différentes aaRS du complexe sont très faibles. De plus, malgré un nombre très important d'études ultérieures portant sur les aaRS bactériennes (leur purification, l'étude de leurs caractéristiques cinétiques et l'établissement de leurs structures 3D), aucune preuve supplémentaire de la formation de MSC ou de la co-purification de plusieurs aaRS n'a été publiée. L'existence de MSC bactériens semble donc controversée, d'autant plus que des études plus récentes montrent l'interaction possible d'aaRS avec d'autres protéines non synthétasiques chez ces mêmes organismes, mais pas d'interaction entre plusieurs aaRS (An and Musier-Forsyth, 2005; Bailly et al., 2007; Buddha et al., 2004a).

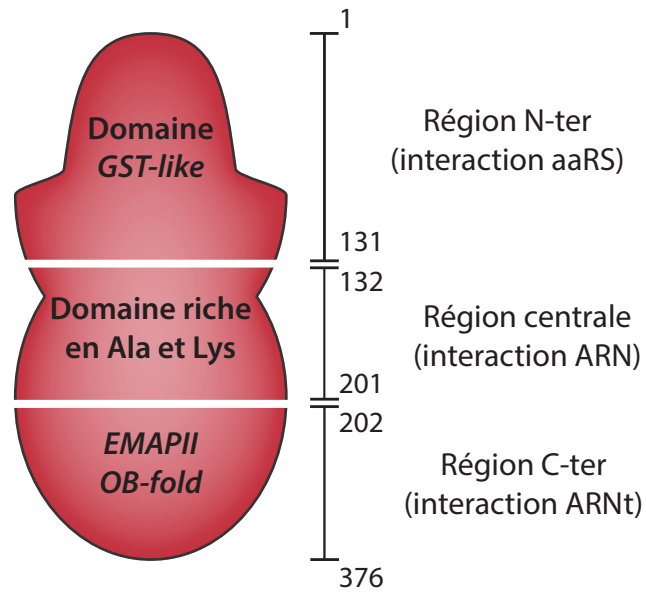
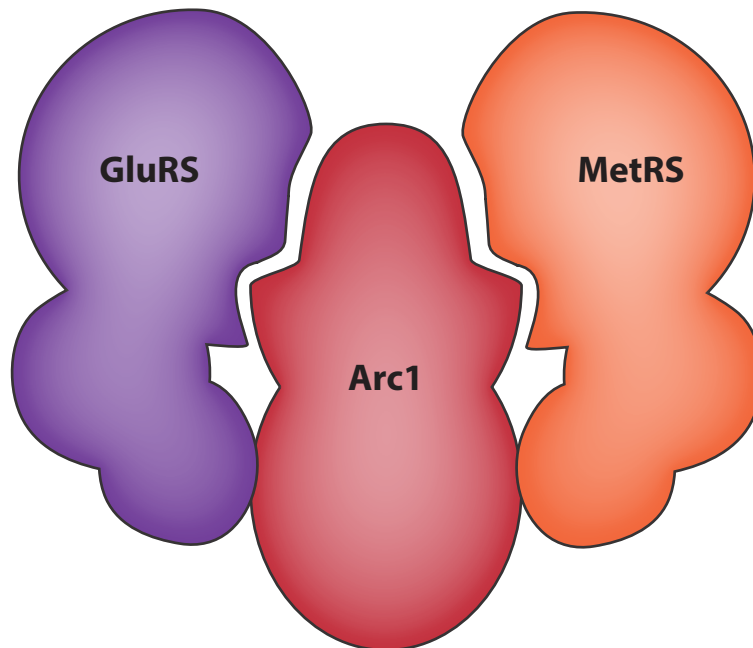
b. Les MSC chez les archées

Durant les premières années de l'étude des MSC, la communauté scientifique pensait que ces complexes étaient restreints aux cellules eucaryotes. C'est uniquement dans les années 1990 que le premier MSC d'archées a été découvert chez *Haloarcula marismortui* (**Figure 11**). Celui-ci est composé de la quasi-totalité des aaRS qui forment un ou deux sous-complexes très résistants à la dissociation. Cependant, aucune information ne permet de spéculer quant au rôle fonctionnel de ce MSC (Goldgur and Safro, 1994). Plus de dix ans après la découverte du premier MSC d'archées, un second complexe a été mis en évidence chez *Methanocaldococcus jannaschii* (*Mja*) (**Figure 11**). En effet, chez cet organisme, la ProRS et l'ARNt^{Pro} sont co-purifiés avec la déshydrogénase putative Mj1338, probablement impliquée dans le métabolisme du méthane. Cependant, le mode de formation, ainsi que l'architecture de ce complexe n'ont pas encore pu être élucidés (Lipman et al., 2003). De plus, Mj1338 est

également capable de s'associer à d'autres aaRS (la LysRS et l'AspRS), suggérant l'existence d'un autre MSC ou la formation d'un MSC incluant les quatre aaRS (**Figure 11**). Le MSC mis en évidence chez cet organisme pourrait jouer un rôle dans le métabolisme ou la synthèse protéique du fait de l'implication de Mj1338 dans la méthanogenèse. Néanmoins, l'existence de ce MSC est discutable car Mj1338 est également capable d'interagir avec les LysRS et ProRS humaines, suggérant que son interaction avec les aaRS de *Mja* pourrait être aspécifique.

Chez *Methanothermobacter thermoautotrophicum* (*Mth*), un MSC relativement stable (K_D de l'ordre du μM) formé de trois aaRS, la LeuRS, la LysRS et la ProRS, a été isolé (Praetorius-Ibba et al., 2007). Les indications structurales de ce complexe montrent que la LeuRS joue un rôle central dans la formation du MSC. En effet, la LeuRS s'associe à la fois à la LysRS, *via* son domaine N-ter, et à la ProRS, de par sa région C-ter (Hausmann and Ibba, 2008). De plus, ce MSC interagit avec la machinerie traductionnelle et plus particulièrement avec le facteur d'élongation EF1A, qui s'associe au domaine d'édition CP1 de la LeuRS (Hausmann et al., 2007). A l'intérieur du complexe, les activités d'aminocyclation des trois enzymes sont augmentées et une légère augmentation de la capacité d'hydrolyse de d'EF1A-GTP est également observée lorsque ce facteur est associé à la LeuRS (**Figure 11**). Ce MSC a donc potentiellement un rôle dans la canalisation (*channeling*) des aa-ARNt (Leu-ARNt^{Leu}, Pro-ARNt^{Pro}, Lys-ARNt^{Lys}) vers le ribosome. Chez ce même organisme, un second MSC associe la SerRS (*via* son domaine C-ter) à l'ArgRS, ce dernier étant associé à l'ARNt^{Arg} (**Figure 11**). Au sein de ce complexe, l'activité de l'ArgRS reste inchangée, alors que celle de la SerRS est augmentée de quatre fois en conditions physiologiques de croissance, mais également à forte température et concentration en sels (Godinic-Mikulcic et al., 2014). Ce MSC est également associé au ribosome, même s'il ne semble pas avoir la même fonction que le MSC décrit précédemment. En effet, il a été proposé que ce dernier sert de plateforme de recyclage des ARNt afin d'éviter leur diffusion et de favoriser une ré-aminocyclation rapide en vue de leur utilisation dans la synthèse protéique.

Finalement, le dernier MSC d'archées mis en évidence à ce jour a été isolé chez *Thermococcus kodakarensis* (Raina et al., 2012). Celui-ci contient un nombre plus important d'aaRS : la GlyRS, la TyrRS, la ProRS, la LeuRS, la CysRS et la MetRS (**Figure 11**). Ce MSC interagit directement avec des protéines de la machinerie traductionnelle et notamment avec le

A**B****Figure 13 : Organisation modulaire d’Arc1 et complexe AME**

(A) Arc1, la protéine d’assemblage du complexe AME peut être divisée en trois sous-parties : (i) la région N-ter, contenant un domaine *GST-like*, impliquée dans l’interaction avec les aaRS ; (ii) une région centrale, riche en Ala et Lys, impliquée dans l’interaction aspécifique avec les ARN ; et (iii) un domaine de type EMAPII, responsable de l’interaction avec différents ARNt. (B) Le complexe AME associe la GluRS, la MetRS et Arc1 avec une stoechiométrie de 1:1:1. Le domaine EMAPII fixe spécifiquement les ARNt^{Glu} et ^{Met} lorsqu’Arc1 est en complexe avec les aaRS.

ribosome (s-u 30 S), ainsi que les facteurs d'initiation IF2 et IF2B et d'élongation EF1A et EF2. De plus, l'activité d'aminocyclation de la LeuRS est augmentée en présence de la ProRS, de la TyrRS et du facteur d'élongation EF1A-GTP, suggérant son implication dans le *channeling* des aa-ARNt. A l'heure actuelle, le mode de formation de ce MSC reste encore inconnu et l'hypothèse selon laquelle l'assemblage est possible *via* l'interaction entre les aaRS est privilégiée.

Chez les archées, plusieurs MSC ont été mis en évidence et ils interagissent, pour la plupart, avec la machinerie traductionnelle, suggérant leur implication dans le *channeling* ou le recyclage des aa-ARNt afin d'optimiser la synthèse protéique et, de ce fait, la croissance de ces organismes. De plus, au sein de ces MSC, il semble y avoir uniquement des aaRS et pas de protéines auxiliaires permettant la formation du complexe, comme c'est par exemple le cas chez les eucaryotes.

c. Le MSC chez *Saccharomyces cerevisiae*

Chez la levure, on retrouve le plus petit MSC eucaryotique répertorié à ce jour. Celui-ci associe une protéine d'assemblage, Arc1 à deux aaRS, la MetRS et la GluRS (Simos et al., 1996) (**Figure 11 & 13**). Ce complexe réduit, encore appelé le complexe AME (Arc1-MRS-ERS, pour Arc1-MetRS-GluRS), est l'un des MSC le mieux caractérisé. La protéine d'assemblage Arc1 est un homologue de la protéine d'assemblage p43, retrouvée au sein du MSC des mammifères. De manière intéressante, Arc1 peut être remplacée par son homologue humain p43 mais celui-ci n'est pas capable de lier la GluRS et la MetRS de levure (Golinelli-Cohen et al., 2004). La délétion du gène *ARC1* n'est pas létale ; elle induit uniquement un retard de croissance important en fermentation et un défaut de croissance en respiration (Debard et al., in prep). L'absence de cette protéine est donc délétère pour la cellule en fonction du type de métabolisme utilisé et ce phénotype est probablement dû à la capacité de cette protéine à intervenir dans divers processus cellulaires indispensables. En effet, Arc1 ne sert pas uniquement de protéine de liaison des deux aaRS, puisqu'elle assure également une fonction nucléaire principale. Une étude de létalité synthétique a montré que la délétion concomitante de *LOS1* et d'*ARC1* n'est pas viable. Comme *Los1* est une exportine impliquée dans l'export des ARNt du noyau, cette létalité synthétique suggère une potentielle participation d'*Arc1* à cet export (Grosshans et al., 2000). De plus, Arc1 est capable d'interagir avec les

phosphatidylinositol 3 phosphate et 3,5 bi-phosphate (PI3P et PI3,5P₂) (Fernandez-Murray and McMaster, 2006). Ces deux espèces étant présentes au niveau des endosomes et de la vacuole, la relocalisation d'Arc1 au niveau de ces compartiments pourrait également avoir un rôle fonctionnel crucial pour la levure.

Au sein du complexe AME, Arc1 est la protéine centrale qui lie les deux aaRS. Cette protéine d'assemblage, composée de 376 aa (42 KDa), peut être divisée en trois domaines principaux (**Figure 13**) : (i) la région N-ter (résidus 1-131) contenant un domaine d'interaction protéines/protéines du type *GST-like*, (ii) un domaine central (résidus 132-201) riche en Lys et en Ala impliqué dans l'interaction non spécifique avec des ARN et (iii) une région C-ter (résidus 202-376) adoptant une conformation d'*OB-fold* (similaire aux domaines EMAPII de mammifères), permettant de lier des ARNt (Frechin et al., 2010). En effet, ce dernier est capable de fixer les ARNt^{Glu} et ARNt^{Met}, mais également les ARNt^{Phe}, ARNt^{Lys}, ARNt^{Ser}, ARNt^{Pro} et ARNt^{Ala} (Deinert et al., 2001). Par contre, lorsqu'Arc1 est en complexe avec la MetRS et la GluRS, la fixation simultanée des ARNt^{Glu} et ARNt^{Met} est privilégiée (Koehler et al., 2013).

La formation du complexe AME est médiée par les domaines N-ter des trois protéines. Le domaine N-ter *GST-like* d'Arc1 interagit avec les domaines N-ter *GST-like* respectifs de chacune des deux aaRS (Simader et al., 2006). Au sein du complexe, Arc1 sert de cofacteur d'aminocyclation et permet de stimuler l'activité de la MetRS de 4 à 400 fois (Golinelli-Cohen and Mirande, 2007; Simos et al., 1998) et celle de la GluRS d'environ deux fois tout en augmentant son affinité pour l'ARNt^{Glu} (Graindorge et al., 2005). Le complexe a donc un rôle "d'amplificateur" de reconnaissance des ARNt^{Glu} et ARNt^{Met}. Ceux-ci s'associent au niveau de la région C-ter d'Arc1 dans une conformation spécifique, favorisant leur reconnaissance et leur aminocyclation par la GluRS et la MetRS respectivement. Le complexe trimérique AME possède une stœchiométrie de 1:1:1 et adopte une structure allongée en forme de Y. Après liaison des ARNt, ce complexe devient plus compact et adopte une structure en forme de cloche (Koehler et al., 2013), suggérant que la formation d'un complexe pentamérique nécessite une flexibilité du complexe AME trimérique et un réarrangement conformationnel.

Les régions, ainsi que les résidus impliqués dans l'interaction entre Arc1 et la GluRS ou la MetRS ont été déterminés par établissement des structures 3D des complexes binaires entre les domaines *GST-like* d'Arc1 et de la GluRS ou de la MetRS (Karanasios et al., 2007;

Simader et al., 2006). La délétion de ces régions d'interaction ou des résidus impliqués dans l'interaction induit une relocalisation des deux aaRS dans d'autres compartiments (Galani et al., 2001). Le complexe AME cytosolique sert donc de point d'ancrage pour les aaRS ayant la capacité de se relocaliser dans des compartiments différents pour assurer des fonctions additionnelles.

d. Le MSC chez *Toxoplasma gondii*

Ce n'est que récemment qu'un MSC a été mis en évidence chez le parasite intracellulaire *Toxoplasma gondii* (*Tgo*) (van Rooyen et al., 2014). Chez cet organisme, une protéine appelée Tg-p43 possède 35% de similarité de séquence avec la protéine p43 humaine et l'homologie est plus importante au niveau de son domaine C-ter, identifié comme étant un domaine de liaison aux ARNt de type EMAPII. De manière intéressante, un domaine *GST-like* est retrouvé dans la région N-ter de Tg-p43. Pour rappel, ce type de domaine est présent chez Arc1, la protéine auxiliaire du complexe AME de levure et chez certaines protéines auxiliaires du MSC de mammifères (**Figure 12**). De plus, la délétion de Tg-p43 n'impacte pas la viabilité du parasite et ne diminue pas sa pathogénicité. Le MSC de *Tgo* est formé de la GluRS et de la MetRS (comme chez la levure), mais également de la GlnRS et de la TyrRS (**Figure 11**). L'assemblage de ce complexe est médié par la protéine Tg-p43 et les images de microscopie électronique suggèrent une flexibilité du MSC (van Rooyen et al., 2014). Néanmoins, la fonction de ce complexe, tant au niveau de l'activité catalytique et de l'ancrage des aaRS au niveau du cytosol, n'est pas encore définie.

e. Le MSC chez *Trypanosoma brucei*

Un MSC a été identifié chez un autre apicomplexe, *Trypanosoma brucei* (*Tbu*) (Cestari et al., 2013). Celui-ci est composé d'au minimum six aaRS (MetRS, ProRS, GlnRS, AlaRS, AspRS et TrpRS) associées à trois AIMP : MCP1-2-3 (*MARS Complex-associated Proteins*) (**Figure 11**). La protéine MCP2 est constituée d'un domaine de liaison aux ARNt (EMAPII), retrouvé chez Arc1 de levure et p43 humain. La protéine MCP1 comporte, en plus du domaine de liaison aux ARNt, un domaine *GST-like*, aussi présent chez son homologue de levure. Finalement, la protéine MCP3 n'a pas d'homologie avec les facteurs d'assemblage des MSC des autres espèces, mais est l'homologue de YbaK, une protéine ayant une activité d'édition des Cys- et Ala-ARNt^{Pro} mésacylés (Wong et al., 2003). Globalement, la MetRS, la ProRS et la GlnRS

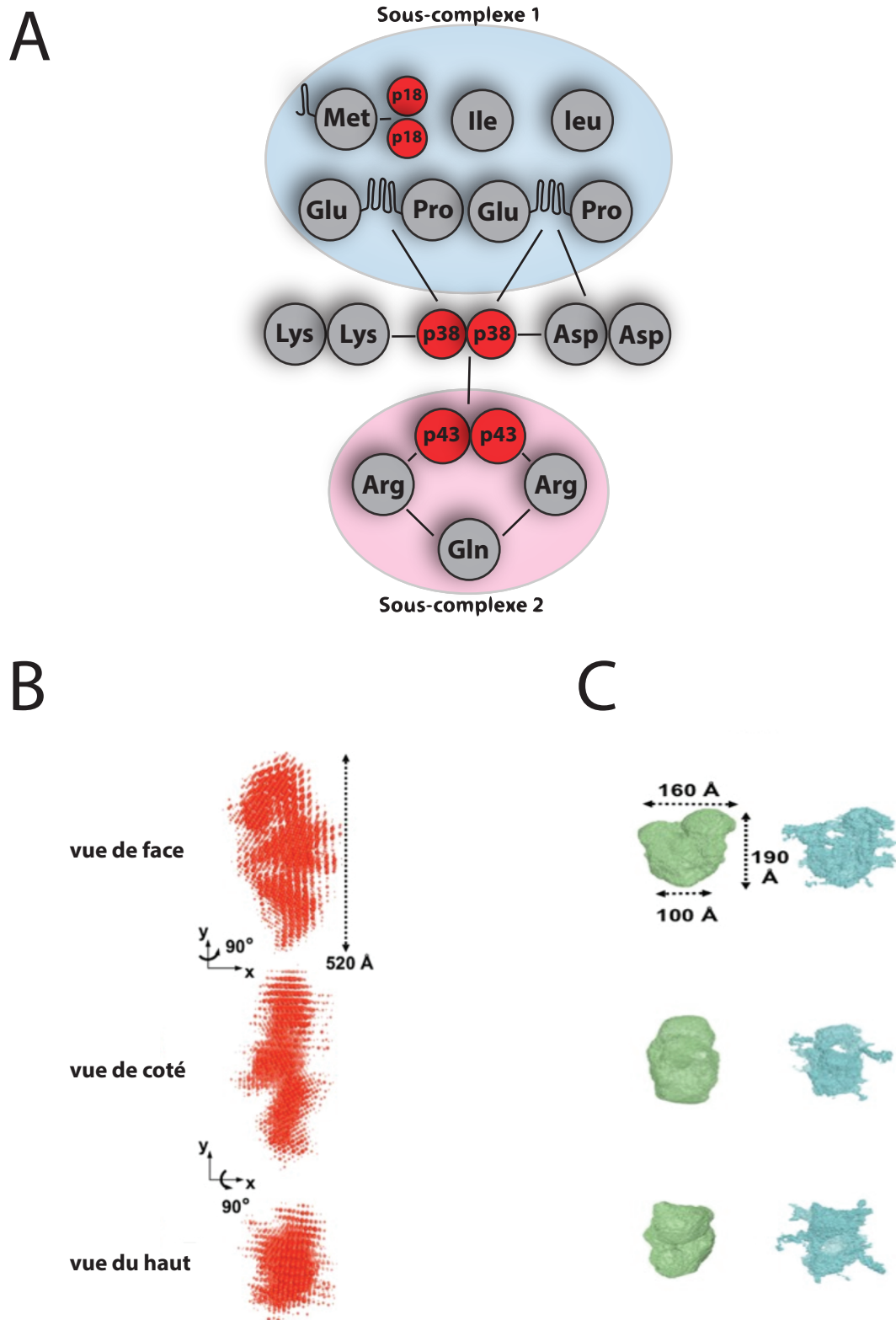


Figure 14 : Composition du MSC de mammifères et modélisation 3D

(A) Le MSC de mammifères est formé de deux sous-complexes contenant chacun une AIMP et un nombre différent d'aaRS. Ceux-ci sont associés par p38 qui interagit également avec deux aaRS. (B) Enveloppe SAXS et (C) modélisation 3D du MSC purifié à partir de foie de lapin (Diaz et al., 2013).

forment le corps du complexe en association avec les protéines MCP1, 2 et 3. L'AlaRS, la TrpRS et l'AspRS, ainsi que d'autres aaRS telles que la GluRS, la CysRS, la GlyRS et l'IleRS pourraient également s'associer au complexe de façon plus transitoire. Au sein du complexe, la protéine MCP2 joue un rôle de cofacteur d'aminocyclation pour les aaRS et notamment la MetRS. La répression de l'expression de MCP2 réduit la croissance du parasite et son infectiosité chez la souris (Cestari et al., 2013). La protéine MCP3, quant à elle, pourrait augmenter la fidélité d'aminocyclation des aaRS présentes dans le complexe *via* son activité d'édition des aa-ARNt mésacylés. En effet, ces espèces pourraient être prises en charge par cette protéine qui favoriserait alors leur hydrolyse et permettrait la réutilisation des ARNt correspondants.

3. Les MSC des eucaryotes multicellulaires

a. Le MSC chez les mammifères

La co-purification de plusieurs aaRS a été mise en évidence dans des foies de rat, de lapin et de chèvre ou encore dans du placenta humain et des cellules de mammifères en culture (Denney, 1977; Godar et al., 1988; Kellermann et al., 1979; Kellermann et al., 1982; Mirande et al., 1982; Norcum, 1989; Ussery et al., 1977; Vennegoor et al., 1972). Quel que soit le mammifère, le MSC est toujours constitué de neuf aaRS : l'ArgRS, la GluRS, la ProRS, l'AspRS, la GlnRS, l'IleRS, la LeuRS, la LysRS et la MetRS. De plus, les trois protéines additionnelles p43, p38 et p18, encore appelées AIMP1-2-3 respectivement (Bandyopadhyay and Deutscher, 1971; Kerjan et al., 1994), sont co-purifiées avec le complexe (Quevillon et al., 1997; Quevillon and Mirande, 1996; Quevillon et al., 1999) (**Figure 14**). Contrairement aux MSC d'unicellulaires, deux activités d'aminocyclation sont retrouvées au sein de la même enzyme, la GlutamylProlyl-ARNt synthétase (GluProRS) ; ces deux domaines catalytiques sont séparés par trois domaines WHEP. Ces derniers sont requis à la fois pour l'incorporation de cette enzyme au sein du MSC et pour la régulation de ses fonctions additionnelles.

Les trois AIMP ont un rôle structural essentiel pour la formation du MSC. La protéine p38 (AIMP2), dont aucun homologue n'a pu être identifié chez la bactérie, la levure ou les archées, a la capacité d'interagir avec de nombreuses aaRS du fait de la présence d'un domaine *GST-like* dans sa région C-ter et de LZ au niveau de sa région N-ter (Ahn et al., 2003; Quevillon et al., 1999). De plus, l'assemblage du complexe nécessite ce facteur auxiliaire, car

la réduction de son expression ou l'ajout de mutations entraînent un défaut d'assemblage du MSC (Kaminska et al., 2009a). On retrouve ensuite la protéine p18 (AIMP3), contenant un domaine *GST-like*, identique à celui de la protéine Arc1 de levure, qui permet l'intégration de la MetRS au sein du MSC, mais également le transfert du Met-ARNt^{Met} au niveau du facteur d'initiation eIF2 (Kang et al., 2012). Finalement, la dernière protéine auxiliaire, p43 (AIMP1), est un homologue d'Arc1 qui possède un domaine EMAPII, similaire au Trbp111 de bactérie, qui se replie sous forme d'*OB-fold* (Renault et al., 2001). Celui-ci est probablement impliqué dans la fixation des ARNt homologues de plusieurs aaRS du complexe, alors que son extrémité N-ter induit l'incorporation de la GlnRS et de l'ArgRS au sein du MSC. Ce facteur possède également un LZ qui interagit directement avec les LZ de p38, favorisant le maintien du complexe (Ahn et al., 2003). De manière intéressante, il existe deux variants traductionnels de p43, dont l'un est importé au sein des mitochondries. Cependant, sa fonction dans ce compartiment n'est pas encore élucidée (Shalak et al., 2009).

Le MSC se décompose en deux sous-complexes liés entre eux par la protéine p38, le tout formant un complexe très stable. Le sous-complexe 1 contient la MetRS associée à p18, la GluProRS, l'IleRS et la LeuRS, alors que le sous-complexe 2 est constitué de l'ArgRS et la GlnRS associées à p43, ainsi que la LysRS et l'AspRS, interagissant directement avec p38 (**Figure 14**). Les modèles à basse résolution obtenus par *small-angle X-ray scattering* (SAXS) ou par microscopie électronique suivie d'une modélisation 3D, suggèrent que le complexe est de grande taille (à titre d'exemple, il est presque deux fois plus gros que le ribosome 80S des eucaryotes) et mesure environ 50 nm de long lorsqu'il est sous forme allongée (Dias et al., 2013). Cette structure dite ouverte pourrait favoriser le relargage des aaRS assurant des fonctions additionnelles.

b. Le MSC chez le nématode

Chez le nématode *Caenorhabditis elegans* (*Cel*), l'unique MSC identifié est composé de la MetRS, la LeuRS, l'IleRS, la GluRS, la ValRS, la GlnRS, l'ArgRS et la LysRS, ainsi que d'une protéine d'assemblage mrsp-38 (*MARS Scaffold Protein 38*) (Havrylenko et al., 2011) (**Figure 11**). Cette dernière possède une très faible homologie avec la protéine p38 (16 %). De plus, aucun homologue des protéines auxiliaires d'autres organismes n'a pu être identifié. Néanmoins, la MetRS de *Cel* est fusionnée au niveau de sa région C-ter à un domaine similaire

à celui de p43, retrouvée chez les mammifères (Havrylenko et al., 2010). Cette aaRS remplace donc la fonction de p43 pour l'incorporation de l'ArgRS et de la GlnRS, et la fonction de p18 pour sa propre incorporation. Contrairement aux MSC de mammifères, celui du *Cel* ne contient, ni l'AspRS, ni la ProRS (probablement du fait de la fission entre la GluRS et la ProRS), mais contient la ValRS, absente du complexe des mammifères. La raison biologique des différences entre le MSC de *Cel* et le MSC de mammifères n'est cependant pas connue. Décrypter le mécanisme d'assemblage de ce complexe pourrait permettre de comprendre les règles gouvernant l'évolution des MSC au sein de l'arbre phylogénétique.

Finalement, la délétion des 67 résidus C-ter de mrsp-38 induit le relargage de nombreuses aaRS, suggérant qu'elle est indispensable pour l'assemblage du complexe. En plus des neuf protéines formant ce MSC, une protéine chaperonne Hsp90 (*Heat shock protein*) et le facteur d'élongation EF1 sont co-purifiés avec le complexe. Bien que la fonction de ce complexe ne soit pas encore caractérisée, il se pourrait qu'il serve de plateforme d'aminocyclation et que l'interaction avec ces deux protéines supplémentaires influe sur l'assemblage du complexe et le *channeling* des aa-ARNt vers la synthèse protéique.

c. Le MSC chez les arthropodes

La présence d'un MSC a été décrite chez les insectes et les crustacés. En effet, un MSC, identique à celui des mammifères, a été mis en évidence chez *Drosophila melanogaster* (*Dme*) et *Artemia salina* (*Asa*) (Kerjan et al., 1994). Le complexe purifié chez la mouche est strictement identique à celui retrouvé chez les mammifères et associe huit aaRS (neuf activités d'aminocyclation) à trois protéines d'assemblage. Ce complexe étant similaire au complexe de mammifères, son organisation structurale n'a pas été étudiée. Cependant, il existe des différences entre les deux MSC et notamment la présence de six domaines WHEP, à la place de trois, permettant la jonction des domaines catalytiques de la GluProRS.

4. Les fonctions non canoniques des aaRS relâchées du MSC

Au sein du MSC, les aaRS catalysent l'aminocyclation de leurs ARNt homologues. Ces complexes sont néanmoins considérés comme des plateformes de rétention pour des aaRS à localisations subcellulaires et à fonctions multiples. L'étude de ces fonctions additionnelles a connu un essor, lorsque l'implication des aaRS a été mise en évidence dans divers processus

non traductionnels essentiels. La plupart des fonctions non traductionnelles des aaRS provenant du MSC ont été décrites chez l'Homme dont certaines sont décrites dans les paragraphes qui suivent (Guo and Schimmel, 2013; Smirnova et al., 2012).

a. Les fonctions additionnelles de la MetRS humaine

La MetRS humaine est la première aaRS dont le relargage du MSC a été mis en évidence. Cette enzyme contient, notamment, un domaine *GST-like* permettant son intégration au sein du complexe *via* l'AIMP p18 (Kwon et al., 2011; Rho et al., 1999). Par ailleurs, la MetRS humaine contient un domaine WHEP en C-ter, absent chez son homologue de levure et de bactéries et deux séquences putatives d'import nucléaire situées également dans la région C-ter (Ko et al., 2000; Rho et al., 1996). Les fonctions additionnelles de la MetRS humaine sont dépendantes des conditions de croissance. En réponse à divers signaux, elle peut ainsi quitter le MSC et se relocaliser dans le noyau et plus précisément dans le nucléole pour participer à la régulation de la synthèse des ARNr (Ko et al., 2000). De plus, la MetRS peut être phosphorylée en réponse à un stress aux UV, et le changement conformationnel subséquent induit le relargage de p18. Cette dernière va alors être importée dans le noyau pour activer certaines voies de réparation de l'ADN endommagé *via* la protéine p53 (Kwon et al., 2011; Park et al., 2005).

b. Les fonctions additionnelles de la GluProRS humaine

Au sein du MSC de mammifères, la GluProRS permet la synthèse des Glu-ARNt^{Glu} et Pro-ARNt^{Pro}. Néanmoins, en plus de sa fonction d'aminacylation, cette enzyme bifonctionnelle participe à de nombreux processus biologiques essentiels pour la cellule. En effet, la GluProRS est impliquée dans la régulation de l'expression des gènes pro-inflammatoires du type *vascular endothelial growth factor A* (VEGF-A), *ceruloplasmin* (Cp), mais également de nombreuses autres cytokines et leurs récepteurs (Mukhopadhyay et al., 2009; Vyas et al., 2009). Chez les cellules myéloïdes, 2 à 4 heures suivant la stimulation par l'interféron γ (IFN γ), les Ser886 et Ser999, présentes au niveau des WHEP de la GluProRS, sont phosphorylées, permettant ainsi son relargage du MSC (Sampath et al., 2004). La phosphorylation de la Ser886 est catalysée par la *cyclin-dependent kinase 5* (Cdk5), entraînant finalement la phosphorylation de la Ser999 par la kinase Cdk5-dépendante AGC (S6K1) (Arif et

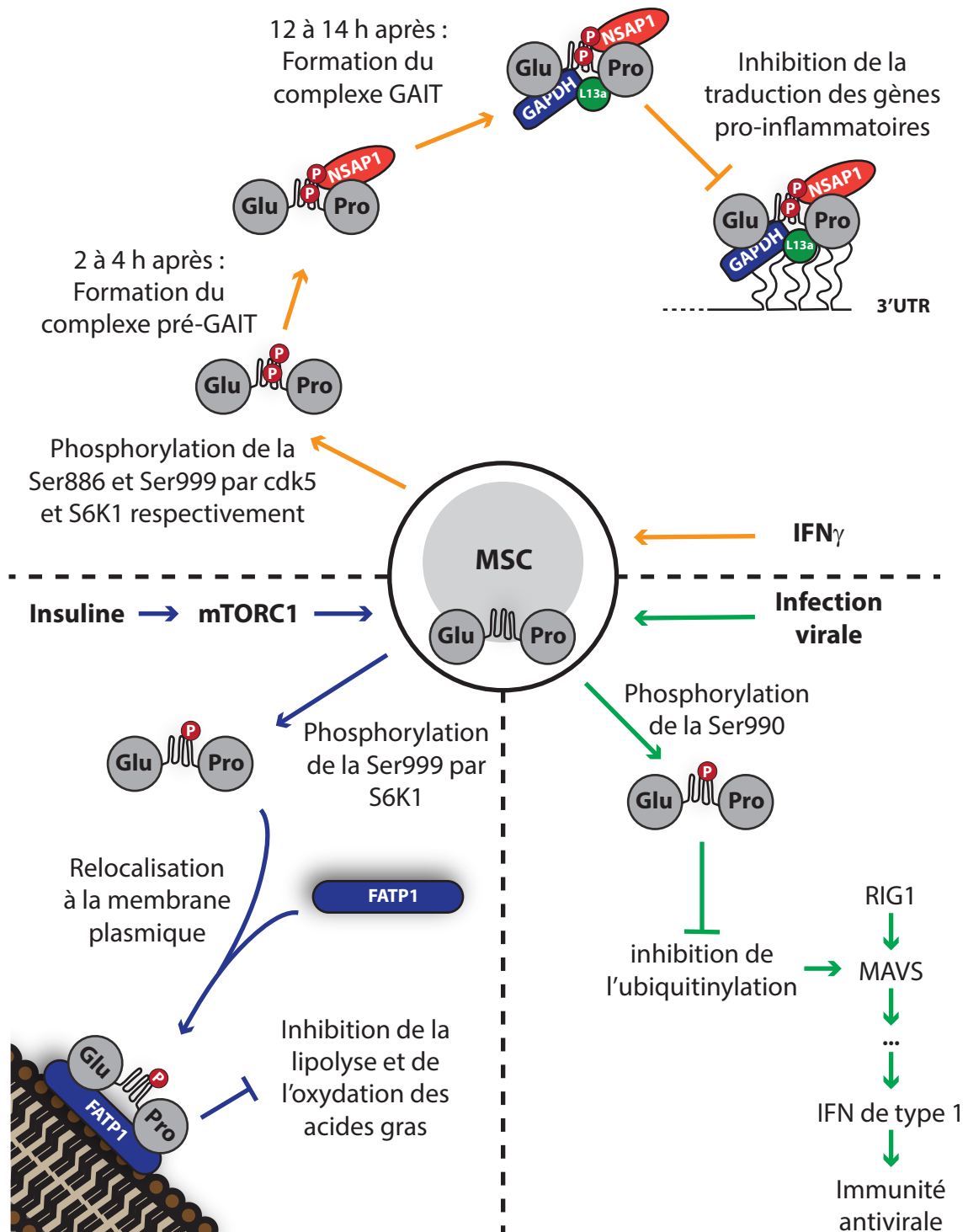


Figure 15 : Fonctions additionnelles de la GluProRS de mammifères

(A) En réponse aux IFN γ , la GluProRS est phosphorylée deux fois (Ser886 et 999), se détache du MSC et forme le complexe GAIT (*IFN- γ -activated inhibitor of translation*) afin d'inhiber la synthèse de protéines pro-inflammatoires. (B) En présence d'insuline, la voie mTORC1 est activée et induit la phosphorylation de la Ser999 de la GluProRS. Celle-ci va alors se relocaliser à la membrane plasmique pour interagir avec un transporteur d'acide gras (FATP1, *Fatty Acid Transporter Protein 1*) et contribuer à l'accumulation lipidique. (C) Lors de l'infection virale, la GluProRS phosphorylée (Ser990) et libérée du MSC est capable d'inhiber l'ubiquitinylation de la protéine MAVS (*Mitochondrial Antiviral Signaling Protein*), et permet de stimuler l'immunité antivirale.

al., 2011; Arif et al., 2009). La forme phosphorylée et libre de la GluProRS s'associe alors à NSAP1 (*NS1-associated protein 1*) pour former le complexe pré-GAIT (*IFN- γ -activated inhibitor of translation*). Il est à noter que la phosphorylation de la Ser886 est indispensable pour la formation de ce complexe. Environ 12 à 14 heures après l'activation par l'IFN γ , la protéine ribosomale L13a est phosphorylée au niveau de la Ser77 *via* une voie faisant intervenir DAPK (*Death-associated protein kinase*) et ZIPK (*Zipper interacting protein kinase*), induisant sa dissociation de la grande s-u 60S du ribosome (Mukhopadhyay et al., 2008). Celle-ci forme alors un dimère avec la GAPDH (glycéraldéhyde-3-phosphate déshydrogénase), qui va interagir avec le complexe pré-GAIT, formant finalement le complexe GAIT ; la phosphorylation de la Ser999 de la GluProRS est cruciale pour cette étape. Ce complexe va reconnaître une structure en tige-boucle, appelée élément GAIT, retrouvée au niveau des 3' UTR des gènes pro-inflammatoires (Sampath et al., 2003). La liaison à cet élément permet (i) l'interaction avec le facteur d'initiation eIF4G, empêchant la fixation d'eIF3, et (ii) de bloquer le recrutement du ribosome, induisant l'inhibition de la traduction de ces gènes (Arif et al., 2009; Kapasi et al., 2007) (**Figure 15**).

Récemment, une étude a permis de comprendre l'implication de la GluProRS dans la réponse immunitaire antivirale (Lee et al., 2016). En effet, lorsque des cellules sont infectées avec des virus, la réduction de l'expression de cette aaRS favorise le développement viral et plus particulièrement celui des virus à ARN. De plus, dans des souris hétérozygotes (*Eprs^{+/-}*) pour la GluProRS, les taux d'IL 6 (InterLeukine), d'IFN β et IFN α diminuent drastiquement et ces souris sont plus sensibles aux infections virales. Comme lors de l'activité de cette aaRS au niveau de l'inflammation, c'est également une phosphorylation qui favorise son relargage du MSC (Ser990), lui permettant ainsi de participer à la réponse antivirale. Lorsque la cellule est infectée par un virus, le senseur intracellulaire RIG-I (*Retinoic acid-Inducible Gene 1*) est activé et interagit avec MAVS (*Mitochondrial Antiviral Signaling Protein*). Cela induit l'activation d'une cascade de signalisation menant à la production d'IFN de type 1. La phosphorylation de la Ser990 de la GluProRS est cruciale pour réprimer l'ubiquitinylation de MAVS et favoriser l'immunité antivirale (**Figure 15**).

Finalement, la phosphorylation de la Ser999 de la GluProRS, par la kinase S6K1, pourrait également jouer un rôle dans l'obésité (Arif et al., 2017). En effet, des souris exprimant un mutant Ser999Ala (non phosphorylable) sont 15 à 20 % plus maigres et leur

durée de vie est augmentée de 15 % comparée à celle de souris exprimant le mutant phosphomimétique Ser999Asp ou la GluProRS sauvage. Cette enzyme joue donc un rôle dans l'obésité *via* sa phosphorylation et son décrochage du MSC (Arif et al., 2017). De plus, l'insuline participe activement à cette fonction. Cette molécule est un agoniste de l'activation de mTORC1/S6K1 et de l'activité anabolique des adipocytes. Lorsque l'insuline se fixe sur son récepteur à la surface des cellules, il y a activation de la phosphorylation de la Ser999 de la GluProRS. Cette phosphorylation, médiée par l'insuline et S6K1, favorise alors l'interaction phosphorylation-dépendante, entre la GluProRS et un transporteur d'acide gras (FATP1 : *Fatty Acid Transporter Protein 1*). Ce complexe de deux protéines se relocalise alors spécifiquement à la membrane plasmique et contribue à l'accumulation lipidique par inhibition de la lipolyse et de l'oxydation des acides gras (**Figure 15**).

c. Les fonctions additionnelles des autres aaRS du MSC

D'autres aaRS sont capables de se détacher du MSC pour assurer des fonctions additionnelles. C'est notamment le cas de la LeuRS humaine qui permet de moduler la localisation de TORC1, un régulateur des voies TOR. Lorsque la concentration en Leu est suffisante, la LeuRS va s'associer à TORC1 et induire sa relocalisation à la membrane lysosomale, favorisant ainsi la régulation de la traduction *via* TOR (Han et al., 2012; Segev and Hay, 2012). De manière intéressante, la fonction additionnelle de la GlnRS dépend également de la concentration de son aa homologue. Lorsque la concentration en Gln est suffisante, cet aa va se fixer dans la poche de reconnaissance de la GlnRS. Celle-ci va alors supprimer l'activité pro-apoptotique d'ASK1 (*Apoptosis Signal-regulating Kinase*) et donc empêcher la mort cellulaire (Ko et al., 2001).

De plus, la LysRS est un régulateur de la réponse immunitaire. Toutes les aaRS activent leur aa homologues en aa~AMP et ces aa activés sont généralement transférés sur l'ARNt lors de l'étape d'aminacylation. Cependant, dans certains cas, l'aa~AMP peut réagir avec une molécule d'ATP pour former l'Ap₄A. Lorsqu'un antigène se fixe sur son récepteur membranaire à la surface d'un mastocyte, la cascade MAPK (*Mitogen Activated Protein kinase*) est activée et induit la phosphorylation de la LysRS qui va alors se relocaliser dans le noyau et synthétiser de l'Ap₄A en grande quantité. Ces molécules activent alors les gènes

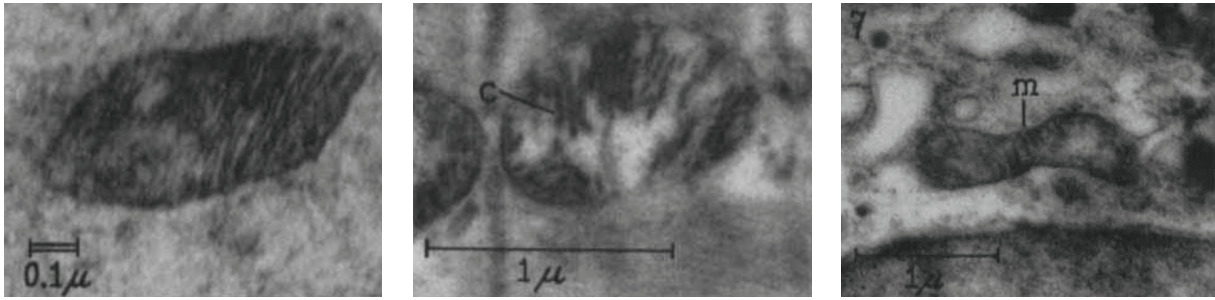
cibles de MITF (*Microphthalmia associated Transcription Factor*) et permettent de réguler le système immunitaire (Lee et al., 2004; Ofir-Birin et al., 2013; Yannay-Cohen et al., 2009).

5. Les fonctions non canoniques des AIMP

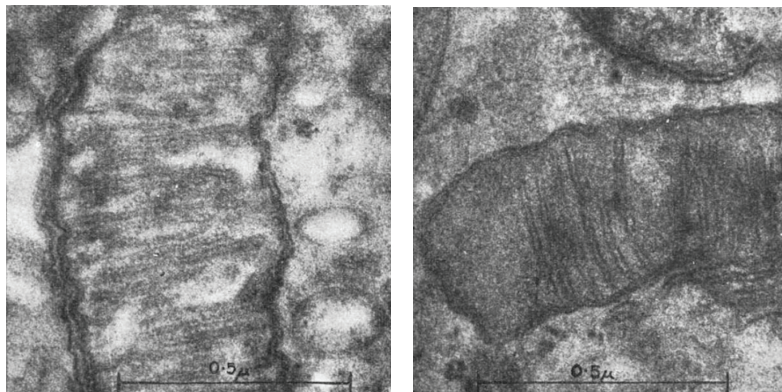
Les protéines auxiliaires des MSC sont cruciales pour la formation du complexe et participent également au relargage des aaRS, qui vont alors assurer leurs fonctions additionnelles. En parallèle, ces protéines peuvent être relâchées du complexe pour assurer elles-mêmes des fonctions additionnelles. En effet, p43 peut être sécrétée en réponse à des conditions hypoxiques ou suite à une stimulation du système immunitaire et joue le rôle de cytokine dans différents types cellulaires. Ainsi, p43 va activer la réponse pro-inflammatoire et la sécrétion de TNF α (*Tumor Necrosis Factor*) ou IL 8 chez les monocytes et les macrophages (Park et al., 2002). De plus, p43 favorise la maturation et l'activation des cellules dendritiques et induit l'expression d'IL 8 et 12, impliquées dans la prévalence aux cancers (Kim et al., 2006). En parallèle, p38 favorise la mort cellulaire de diverses manières. En outre, ce facteur peut se relocaliser dans le noyau, se fixer à p53 et empêcher sa dégradation *via* MDM2 (*Mouse Double Minute 2 homolog*) (Han et al., 2008). Il peut également réguler négativement la protéine TRAF2 (*TNF Receptor–Associated Factor*), un médiateur de la signalisation TNF pour favoriser l'apoptose (Choi et al., 2009). En effet, p38 est également une cible de la *Parkin* qui favorise son ubiquitinylation dans les cellules neuronales. Si l'action de cette protéine est inhibée, p38 s'accumule dans les neurones et induit l'apoptose accrue de ces cellules (Corti et al., 2003; Ko et al., 2005). La protéine p38 a potentiellement un rôle important dans la répression tumorale *via* la régulation de l'apoptose. Pour finir, le plus petit facteur du MSC, p18, assure également des fonctions additionnelles. Il est notamment importé dans le noyau en réponse aux dommages à l'ADN ou lors de stress oncogéniques (Park et al., 2005; Park et al., 2006). L'action de p18 dans le noyau peut être coordonnée à celle de p38 et permet, *via* p53, de lutter contre les dommages. Cette fonction est très importante pour l'hémostasie cellulaire, car des mutants de p18 incapables de réguler p53 sont retrouvés chez des patients atteints de leucémie (Kim et al., 2008).

L'importance des fonctions additionnelles de ces trois facteurs est également démontrée dans des souris mutantes. D'ailleurs, les souris hétérozygotes pour p43 (p43^{+/-}) ont une déficience neuronale (Zhu et al., 2009). Les souris déficientes pour p38 sont sujettes à la

A



B



C

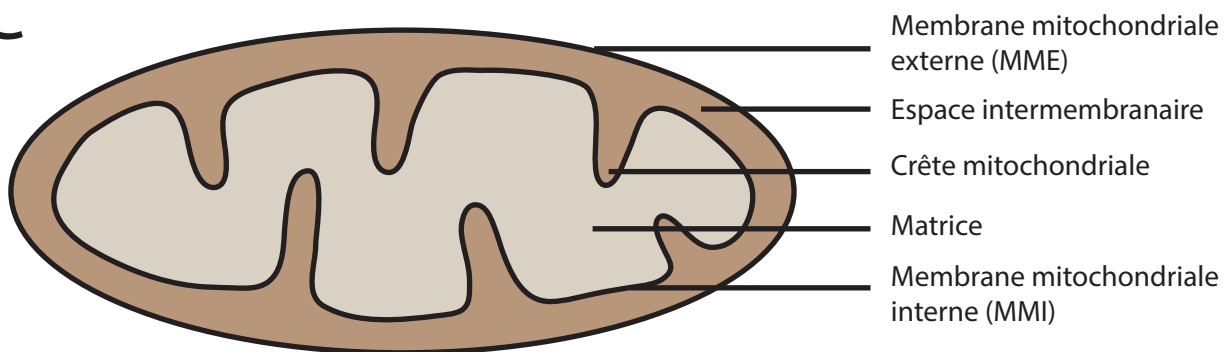


Figure 16 : Les observations des mitochondries et leur structure

Les premières images de microscopie électronique des mitochondries obtenues par le Pr. Palade (Palade, 1953) (A) et le Dr. Sjöstrand (Sjöstrand, 1953) (B). (C) Schéma de la mitochondrie, subdivisée en cinq sous-compartiments ayant tous des fonctions essentielles.

formation de tumeurs et les souris homozygotes pour ce facteur ont une mortalité néonatale due à des dysfonctionnements pulmonaires (Choi et al., 2009). Finalement, la déficience de p18 (p18^{-/-}) induit une létalité embryonnaire, alors que les souris hétérozygotes (p18^{+/-}) naissent normalement, mais sont plus susceptibles de développer des tumeurs (Park et al., 2005).

VII. Les mitochondries

1. La découverte et la structure des mitochondries

Lors de l'étude du muscle en 1857, l'anatomiste suisse Albert von Kölliker décrit pour la première fois les mitochondries. Mais c'est en 1886 que l'histologiste allemand, le Dr. Richard Altmann, développe une technique permettant de préserver et de fixer les tissus en vue de leur observation microscopique. C'est en utilisant un nouveau type de marquage à la fuschine qu'il observe des filaments de granules dans la plupart des cellules. Il nomme ces structures des bioplastes et les décrit comme des organismes élémentaires vivant à l'intérieur des cellules et qui assurent différents processus vitaux. Cette hypothèse est néanmoins rapidement réfutée par la communauté scientifique. Le terme mitochondrie est, pour la première fois, utilisé en 1898 par le Dr. Carl Benda lors de l'étude de la spermatogénèse. En effet, en étudiant les bioplastes du Dr. Altmann, il observe, suite à une coloration au cristal violet, des structures formant des filaments, mais également des granules. Il renomme alors ces structures des mitochondries du grec mitos (filament) et chondros (grain), en référence à leur forme.

Les premières images décrivant la forme des mitochondries ont été obtenues par microscopie électronique en 1945 (Claude and Fullam, 1945) et confirment le fait que les mitochondries sont entourées d'une membrane. Néanmoins, l'étude plus précise de cet organelle est possible à partir des années 1950, suite au développement des techniques de coupes fines de tissus. La première image haute résolution de microscopie électronique est publiée en 1953 par le Pr. Palade et permet de décrire les mitochondries comme étant des compartiments entourés par une membrane qui se replie pour former des crêtes (Palade, 1953) (**Figure 16A**). C'est le Dr. Sjöstrand qui observe la présence d'une double membrane mitochondriale (Sjostrand, 1953) (**Figure 16B**). Il décrit alors les mitochondries comme "deux

espaces ou chambres, délimités par des membranes mitochondriales, une chambre externe entre deux membranes et une chambre interne limitée par la membrane interne". Au fil des années, cette description de Palade a été acceptée par la communauté scientifique (**Figure 16C**). La mitochondrie, telle qu'on l'a décrit aujourd'hui, est donc un compartiment formé de deux membranes : (i) la membrane mitochondriale externe (MME), délimitant la mitochondrie et (ii) La membrane mitochondriale interne (MMI). La MME contient de nombreuses protéines différentes, dont des protéines de structure et des porines (ou VDAC, *Voltage Dependant Anions Channel*) impliquées dans l'import de métabolites, d'ions et d'ATP. De plus, la MME contient le complexe TOM (*Translocase of Outer Membrane*) permettant l'import des protéines vers l'intérieur de la mitochondrie (voir introduction § VII.5). La MMI, dans laquelle sont notamment enchâssés les complexes respiratoires responsables de la production de la majorité de l'ATP cellulaire par phosphorylation oxydative. Au sein de cette membrane, on retrouve également les protéines du complexe TIM (*Translocase of Inner Membrane*) permettant l'import des protéines dans la matrice mitochondriale (voir introduction § VII.5).

Ces deux membranes délimitent les deux sous-compartiments de la mitochondrie : l'espace intermembranaire (IMS : *InterMembrane Space*) et la matrice. Cette dernière abrite divers processus indispensables au bon fonctionnement de la mitochondrie, comme l'expression des gènes à partir du génome mitochondrial et la synthèse des protéines mitochondriales, qui sont principalement des s-u des complexes de la chaîne respiratoire (CR). De plus, la matrice joue un rôle primordial dans la production de cofacteurs de la CR (NADH : *Nicotinamide Adénine DiNucléotide* ou succinate) *via* le cycle de Krebs. L'IMS contient peu de protéines (environ 100), mais celles-ci sont impliquées dans des voies importantes pour l'homéostasie de la cellule, notamment dans l'import des protéines mitochondriales, l'assemblage des s-u de la CR et/ou des centres Fer-Soufre, la signalisation cellulaire ou l'apoptose.

2. Les origines des mitochondries

Depuis la découverte des mitochondries, la communauté scientifique a sans cesse tenté d'expliquer l'origine de ces compartiments spécialisés dans la synthèse d'énergie *via* la CR. Certains scientifiques pensent que le Dr. Altmann est le précurseur de la théorie

endosymbiotique de l'apparition des mitochondries dans la cellule. Cette théorie suggère que les mitochondries sont apparues suite à l'endosymbiose entre deux organismes (un hôte anaérobie et un endosymbionte aérobie) et que les proto-mitochondries ont transféré la majorité de leurs gènes à la cellule hôte, pour former le compartiment que l'on connaît aujourd'hui. Cependant, cette théorie n'est pas issue des travaux du Dr. Altmann car selon lui, tout ce qui est retrouvé dans les cellules eucaryotes est considéré comme un bioblaste : le cytosol, le noyau et les chromosomes par exemple. Selon son hypothèse, les bioblastes correspondent à une organisation chimique qui est plus grande qu'une molécule et plus petite qu'une cellule (Altmann, 1890). Ils peuvent donc être comparés aux complexes macromoléculaires que l'on connaît aujourd'hui qui, soit dit en passant, n'étaient pas visualisables avec les microscopes de l'époque. Le précurseur de cette théorie est le Dr. Mereschkowsky qui émet l'hypothèse qu'il y a une relation entre les bactéries et les mitochondries car ces dernières sont impliquées dans de nombreux processus cellulaires (Mereschkowsky, 1905). C'est ensuite le Pr. Wallin qui développe un peu plus cette théorie et suggère que l'endosymbiose d'une bactérie est impliquée dans la formation des mitochondries, mais il ne présume pas de l'origine de l'hôte (Wallin, 1927). Cependant, certains scientifiques de l'époque ne sont pas d'accord avec cette théorie de Mereschkowsky (Lederberg, 1952).

C'est au cours des années 1960 que la théorie d'endosymbiose à l'origine des mitochondries est remise au goût du jour par le Dr. Lynn Margulis (Sagan, 1967), mais également par un article du Dr. Goksøyr (Goksoyr, 1967). Ce dernier propose que deux endosymbioses successives ont permis l'apparition des mitochondries, puis des chloroplastes. Pour lui, un procaryote aérobie a été "ingéré" par un eucaryote anaérobie du fait de l'émergence d'oxygène dans l'atmosphère. La perte d'autonomie du procaryote est le résultat du transfert horizontal des gènes vers le noyau de la cellule hôte. Quant au Dr. Margulis, elle ravive l'idée selon laquelle l'ancêtre de la mitochondrie est d'origine procaryotique. De plus, dans sa théorie, les organismes photosynthétiques ont joué un rôle central de par leur production et leur libération d'oxygène. En effet, les mitochondries seraient alors apparues pour compenser l'accumulation de l'oxygène dans l'atmosphère.

Au début des années 1970, il y avait beaucoup de réticences à la théorie de la symbiose pour l'origine de la cellule eucaryote. C'est notamment le cas des Dr. Raff et Mahler qui

proposent un modèle non symbiotique de formation des mitochondries. Selon eux, le proto-eucaryote est un organisme hétérotrophe aérobie de grande taille qui, par invagination de la membrane interne, a enfermé sa CR dans un compartiment clos, générant ainsi la mitochondrie (Raff and Mahler, 1972). Une autre théorie non symbiotique suggère que la mitochondrie est issue de la fusion et de la restructuration des thylakoïdes respiratoires dans une cyanobactérie (Cavalier-Smith, 1975). De plus, le séquençage du génome mitochondrial humain ne permet pas non plus de conclure quant à l'origine des mitochondries (Anderson et al., 1981). Les études concernant l'origine des mitochondries prennent alors une toute autre tournure. Des analyses phylogénétiques suggèrent l'origine α -protéobactérienne des cytochromes (Cyt) de la CR (Schwartz and Dayhoff, 1978). Dans les années 1980, l'analyse de l'ARNr 16S (Yang et al., 1985) et l'étude des chaperonnes mitochondriales codées par le génome nucléaire (Jindal et al., 1989), confortent l'hypothèse de l'endosymbiose d'une α -protéobactérie. C'est finalement le séquençage du génome de la bactérie intracellulaire obligatoire *Rickettsia prowazekii* (*Rpr*), qui met un point d'orgue à la validation de la théorie endosymbiotique (Andersson et al., 1998). En effet, les protéines codées par le génome de cette bactérie sont très similaires aux protéines mitochondriales et l'endosymbionte serait par conséquent un organisme apparenté à *Rpr* (Kurland and Andersson, 2000). Ces différents arguments sont en faveur de l'hypothèse de l'origine endosymbiotique des mitochondries.

Récemment, d'autres hypothèses ont néanmoins émergé. Ainsi, le Dr. Gray a proposé une théorie pré-mitochondriale. Celle-ci ne prend pas en compte la formation de la cellule eucaryote et selon lui, les mitochondries sont apparues dans une cellule aérobie, dont l'organisation était plus ou moins celle des eucaryotes actuels. En effet, dans cette cellule, un compartiment présentant une seule membrane et produisant de l'ATP était déjà présent et celui-ci a été reconverti en mitochondrie par l'adressage des protéines de l'endosymbionte α -protéobactérien (Gray, 2014). Le dernier modèle décrivant l'apparition des mitochondries est l'œuvre des Dr. David et Buzz Baum en 2014. Dans cette théorie, une relation de mutualisme étroite était fondée entre un hôte d'origine archéenne et un épibionte α -protéobactérien, qui vivait à la surface de la cellule hôte. Cette dernière a alors commencé à faire des protrusions membranaires entourant l'épibionte afin d'augmenter leur surface de contact. Finalement, ces protrusions se sont étendues formant alors la membrane externe (Baum and Baum, 2014).

3. Le génome et le contenu protéique des mitochondries

L'ADN mitochondrial est essentiel pour tout eucaryote se développant en respiration. Chez la levure *S. cerevisiae*, le génome mitochondrial est circulaire, a une taille de 85 779 paires de bases (Foury et al., 1998) et est caractérisé par une faible densité de gènes riches en paires de bases A/T. En parallèle, les régions intergéniques sont formées de longues régions de centaines de paires de bases A/T entrecoupées par des régions riches en G/C (de Zamaroczy and Bernardi, 1986). Ce génome contient les gènes codant pour les s-u du complexe de la Cyt c oxydase I, II et III (*COX1*, -2 et -3), de trois s-u de l'ATP synthase (*ATP6*, -8 et -9), de l'apocyt b (*CYTB*) et de la protéine ribosomale Var1 (*VAR1*). De plus, le génome code également les 24 ARNt, les ARNr 15S et 21S, et la composante ARN de la RNase P (de Zamaroczy and Bernardi, 1986). Ce génome est donc minimal et ne code que pour huit protéines. Néanmoins, l'une des premières études de protéomique a permis d'assigner une localisation mitochondriale à 749 protéines avec une couverture du protéome attendue de 90 % (Sickmann et al., 2003). Les avancées dans le domaine de la spectrométrie de masse ont permis d'intensifier les études du mitoprotéome et en 2007, Reinders et ses collaborateurs ont mis en évidence 102 protéines additionnelles, augmentant ainsi le mitoprotéome à 851 protéines avec une couverture putative de 84% (Reinders and Sickmann, 2007). En mai 2016, sur la base de données de levure *Saccharomyces genome database* (SGD) (Cherry et al., 2012), le nombre de protéines mitochondriales est passé à 1187 (Gonczarowska-Jorge et al., 2017). Par ailleurs, la levure étant capable de se développer dans différentes conditions, notamment en fermentation ou en respiration, les besoins protéiques au sein de la mitochondrie peuvent certainement varier. En effet, en 2004, Ohlmeier et ses collaborateurs ont déterminé la variation de ce protéome suite à la transition diauxique ; passage d'un métabolisme fermentaire à un métabolisme respiratoire. Cette étude a mis en évidence la surexpression de 17 protéines impliquées, notamment, dans la respiration, le cycle de Krebs et le métabolisme de l'éthanol (Ohlmeier et al., 2004). Une autre étude protéomique comparant des souches se développant en fermentation ou en respiration, a également montré des changements protéomiques selon le métabolisme utilisé. Ainsi, en respiration, certaines protéines de la CR et du cycle de l'acide citrique sont surexprimées, tandis que d'autres sont uniquement détectées en présence d'éthanol (Kolkman et al., 2005). Finalement, une étude récente a comparé les mitoprotéomes et phosphoprotéomes de souches utilisant du glucose (fermentation), du

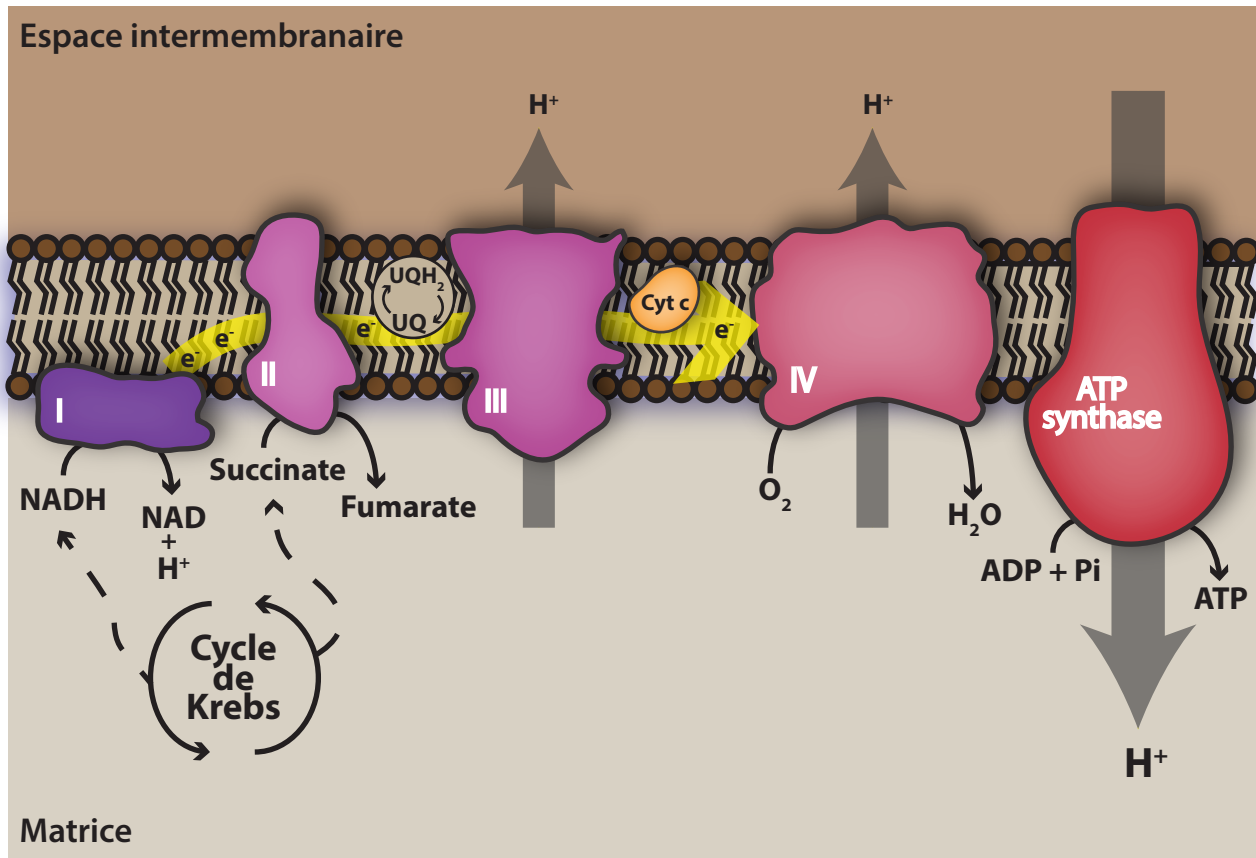


Figure 17 : La chaîne respiratoire de levure

Le NADH et le succinate, produits lors du cycle de Krebs, sont oxydés et les électrons libérés sont transférés à l'ubiquinone (UQ) pour former l'ubiquinol (UQH₂). Cette molécule va alors être oxydée et les électrons vont être transférés sur le cytochrome c (Cyt c). L'oxydation finale du Cyt c, ainsi que l'oxydation de l'UQH₂ permet la formation d'un gradient de proton. Ce gradient électrochimique va alors être utilisé par l'ATP synthase qui va coupler la translocation de protons de l'espace intermembranaire à la matrice avec la phosphorylation d'ADP en ATP.

lactate (respiration) ou du galactose (production glycolytique d'ATP sans répression des mitochondries) pour se développer. Celle-ci confirme la régulation différentielle de 170 protéines en fonction des conditions de croissance (Renvoise et al., 2014). La quantité de protéines impliquées dans le métabolisme énergétique, la synthèse protéique et la biogénèse des mitochondries, est réduite en glucose. En respiration, les protéines abondantes sont celles du métabolisme respiratoire et du cycle de Krebs. Le mitoprotéome est donc variable en fonction des conditions de croissance. En réponse aux besoins de l'organisme, l'expression des protéines mitochondriales codées par le génome nucléaire, ainsi que l'import de ces protéines dans les mitochondries, sont finement régulés.

4. Les fonctions mitochondriales

La fonction principale des mitochondries est bien évidemment la production d'ATP par phosphorylation oxydative. Ce processus est basé sur le transfert d'électrons, générés par l'oxydation du NADH et du succinate, au travers de quatre complexes de la CR, couplé à la formation d'un gradient de proton au travers de la MMI (Fernie et al., 2004; Galluzzi et al., 2012; Reichert and Neupert, 2004). Ce gradient électrochimique ainsi généré est finalement utilisé par la F_1-F_0 -ATP synthase pour synthétiser l'ATP (**Figure 17**). La CR est formée de quatre complexes différents, retrouvés au sein de la MMI, nommés de I à IV, dont la composition varie en fonction de l'organisme décrit. En plus de ces complexes, le transfert d'électrons fait intervenir deux transporteurs, l'ubiquinone et le Cyt c. Les protéines retrouvées au sein des différents complexes proviennent des génomes nucléaire et mitochondrial. Comme évoqué précédemment, chez la levure, sept s-u de ces complexes sont codées par le génome mitochondrial, les 38 autres le sont par le génome nucléaire et sont importées spécifiquement dans la mitochondrie. Chez ce même organisme, et contrairement à la CR des mammifères, il n'y a pas de complexe I, mais c'est une protéine unique, la NADH oxydase, qui capte le NADH produit dans la matrice mitochondriale. Cette protéine permet alors (i) la récupération de deux électrons par oxydation de la molécule de NADH et (ii) leur transfert sur l'ubiquinone (UQ) pour former l'ubiquinol (UQH_2). Le complexe II, encore appelé succinate déshydrogénase, va fixer le succinate et récupérer un électron pour le transférer également à l'UQ. Les électrons fixés à l'UQ issus des complexes I et II, vont alors diffuser au travers de la MMI vers le complexe III (Cyt *bc1*). Il va alors y avoir oxydation de l' UQH_2 et réduction du Cyt c qui va donc prendre en charge les électrons. Ces derniers vont finalement être transférés au

niveau du complexe IV (Cyt c oxydase), par l'intermédiaire du Cyt c qui sert de navette pour le transfert des électrons. L'oxydation de l'ubiquinol et du Cyt c va alors favoriser l'expulsion de protons vers l'IMS, et ainsi induire la formation d'un gradient. Celui-ci est très important car l'ATP synthase va coupler la translocation d'un proton de l'IMS vers la matrice à la phosphorylation d'une molécule d'ADP en ATP (**Figure 17**). L'origine protéique des complexes de la CR étant mixte, il doit y avoir une régulation fine de la quantité de chaque protéine synthétisée à partir des génomes nucléaire et mitochondrial, garantissant le bon déroulement des processus mitochondriaux. Ce mécanisme de coordination de l'expression des gènes implique les voies antérograde (du noyau vers la mitochondrie) et rétrograde (de la mitochondrie vers le noyau). Ces deux voies permettent la régulation de l'expression des gènes en réponse aux conditions environnementales, mais également à des signaux endogènes (voie antérograde) et des signaux mitochondriaux (voie rétrograde). Cette régulation est cruciale pour le respect de la stœchiométrie des protéines des complexes de la CR et un mauvais assemblage de ces complexes peut être néfaste pour la cellule. En effet, lorsque l'assemblage du complexe III n'est pas stœchiométrique ou que le transfert d'électrons *via* le Cyt c n'est pas efficace, il y a production d'espèces réactives de l'oxygène (*Reactive Oxygen Species* ou ROS), impliquées notamment dans l'apparition de certaines mutations au niveau l'ADN (Chen et al., 2003; Drose and Brandt, 2008; Lenaz et al., 2010).

En plus de la production d'ATP par phosphorylation oxydative, la mitochondrie est le siège d'un grand nombre d'autres voies métaboliques. En effet, elle participe au stockage du calcium (Ca^{2+}) cellulaire et au processus de mort cellulaire programmée par apoptose (Orrenius et al., 2003; Pinton et al., 2008). La mitochondrie est également le siège de la synthèse d'acétyl-CoA par β -oxydation des acides gras (Houten and Wanders, 2010), de la synthèse de certains aa (par le cycle de Krebs), ainsi que la biogénèse des centres Fer-Soufre, qui seront ensuite incorporés, soit dans des protéines mitochondriales, soit dans des protéines cytosoliques après export de la mitochondrie (Lill and Muhlenhoff, 2005).

Au niveau de ce compartiment, on retrouve également le catabolisme des aa par le cycle de l'urée (Watford, 1991) et le cycle de l'acide citrique (cycle de Krebs), produisant notamment du NADH et du succinate, utilisés lors de la production d'ATP par phosphorylation oxydative (Krebs and Johnson, 1937). De plus, la mitochondrie génère, *via* sa CR, des ROS qui sont impliquées dans de nombreuses voies de signalisation (Dai et al., 2014; Dalton et al.,

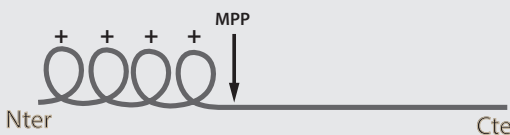
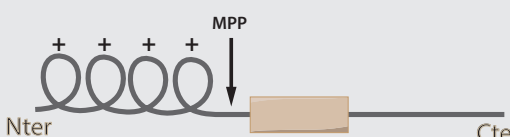
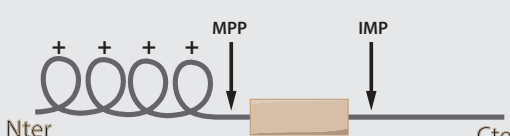

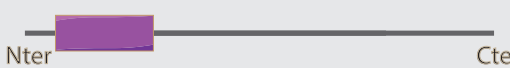
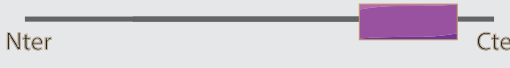
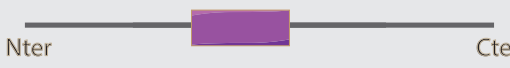

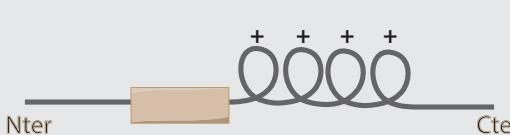
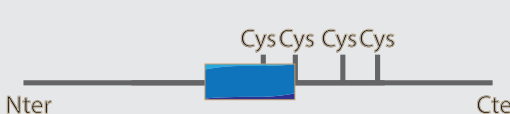
		Type de séquence d'import	Import via	Localisation
Préséquences et variations		Préséquence : hélice amphiphile généralement clivée après import	TOM TIM23 PAM	Matrice
		Préséquence + ancre hydrophobe non clivée	TOM TIM23 (PAM)	MMI
		Préséquence + région hydrophobe clivée	TOM TIM23 (PAM)	IMS
Signaux non clivables des protéines hydrophobes		Signaux β	TOM SAM	MME (tonneau β)
		Ancre N-ter	Mim1	MME (hélice α)
		Ancre C-ter	?	MME (hélice α)
		Signaux internes	(TOM) (SAM)	MME (hélice α)
		Signaux internes multiples	TOM TIM9-10 TIM22	MMI (transporteur de métabolites)
		Presequence-like signaux internes (après région hydrophobe)	TOM TIM23	MMI
Signaux interne pour l'IMS		Signaux riches en cystéines	TOM MIA	IMS

Figure 18 : Les signaux et modes d'import mitochondriaux

Les séquences d'import mitochondriales peuvent être divisées en trois grandes classes : (i) les préséquences en N-ter généralement clivables, (ii) les signaux non clivables hydrophobes et (iii) les signaux internes riches en Cys. Certaines de ces séquences peuvent être clivées par des peptidases mitochondriales comme la *Mitochondrial Processing Peptidase* (MPP) ou l'*Inner Membrane Peptidase* (IMP). La position, ainsi que le type de signal d'import médient le devenir de la protéine au sein de la mitochondrie.

1999; Hamanaka and Chandel, 2010). Néanmoins, ces ROS peuvent provoquer des dommages à l'ADN, la peroxydation des lipides ou l'oxydation des protéines (Imlay, 2008; Ray et al., 2012). La cellule a développé divers mécanismes pour diminuer la réactivité des ROS. Au sein de la mitochondrie, on retrouve ainsi des superoxydes dismutases qui permettent la conversion des ions superoxydes ($O_2^{\cdot-}$) en peroxyde d'hydrogène (H_2O_2). Les catalases et peroxydases prennent ensuite en charge l' H_2O_2 et le convertissent en H_2O (Balaban et al., 2005; Galluzzi et al., 2012).

5. Les mécanismes d'import des protéines mitochondriales

Environ 99 % des quelques 1000 protéines mitochondriales sont codées par le génome nucléaire, traduites dans le cytosol, puis importées dans la mitochondrie par des mécanismes spécifiques. Il a longtemps été admis que l'import des protéines mitochondriales fait intervenir un mécanisme d'import majeur. Néanmoins, les études de biogénèse des mitochondries ont révélé une complexité des modes et des mécanismes d'import des protéines dans ce compartiment. La machinerie d'import présente au niveau des MME et MMI est donc un réseau complexe permettant l'entrée des protéines dans la mitochondrie et contribuant à la biogénèse, la dynamique, ainsi que les fonctions mitochondriales dans les diverses conditions physiologiques. Une protéine mitochondriale possède un signal d'import spécifique appelé le MTS (*Mitochondrial Targeting Signal*), qui dicte la localisation sub-mitochondriale de la protéine et surtout la voie d'adressage qui induit son import dans le bon sous-compartiment. En effet, ces signaux d'import peuvent être de trois types principaux : (i) les préséquences en N-ter, clivées après import qui sont retrouvées principalement au niveau des précurseurs de protéines matricielles, de l'IMS et quelques protéines de la MMI ; (ii) les séquences non clivables de protéines hydrophobes, qui permettent généralement l'import des protéines membranaires et (iii) les signaux internes riches en Cys pour les protéines de l'IMS (**Figure 18**). Ces précurseurs de protéines mitochondriales sont importés selon différentes voies (**Figure 19**).

a. L'import des protéines contenant des préséquences en N-ter

Environ 60 % des protéines mitochondriales sont synthétisées dans le cytosol avec une préséquence d'import clivable retrouvée dans la région N-ter. Après import dans la matrice,

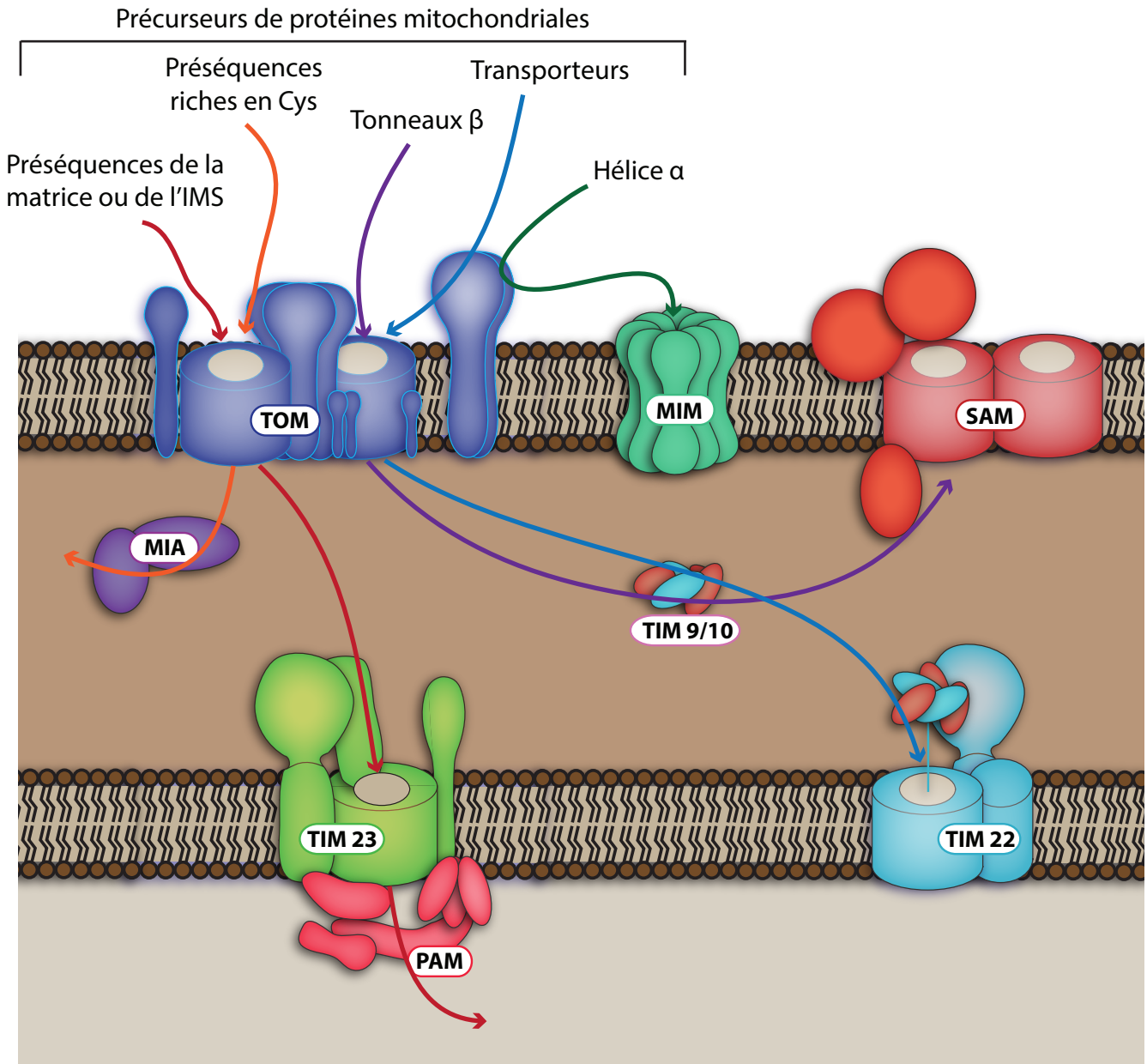


Figure 19 : Les différentes voies d'import mitochondrial

La majorité des précurseurs de protéines mitochondriales est reconnue par le complexe TOM (*Translocase of Outer Membrane*) au niveau de la MME. Le devenir des précurseurs dépend alors du type de signal d'import qu'ils contiennent. Les protéines matricielles et la majorité des protéines de l'IMS seront prises en charge par le complexe TIM (*Translocase of Inner Membrane*) qui permettra leur translocation dans la matrice ou l'IMS. Après le passage au sein de TOM, les précurseurs ayant un signal d'import riche en Cys sont pris en charge dans l'IMS par le complexe MIA (*Mitochondrial Intermembrane space import and Assembly*) qui favorise leur maturation. Certains précurseurs de protéines membranaires seront pris en charge par le complexe de chaperonnes TIM9/10. Les précurseurs de tonneaux β seront alors déposés au niveau du complexe SAM (*Sorting and assembly machinery*) et ceux de transporteurs au niveau de TIM22. Ces deux complexes permettront l'incorporation de ces protéines dans la MME ou MMI. Finalement, les précurseurs des hélices α seront reconnus par le complexe TOM qui les dirigera directement vers le complexe MIM (*Mitochondrial Import*). Il permettra alors leur incorporation dans la MME.

ces signaux sont éliminés par des peptidases mitochondriales spécifiques (Mossmann et al., 2012; Taylor et al., 2001; Vogtle et al., 2009). Cependant, certaines protéines ont des préséquences non canoniques qui peuvent être clivées par les peptidases mitochondriales. C'est notamment le cas de l'hélicase mitochondriale Hmi1 qui possède une structure *préséquence-like* clivable dans sa région C-ter (Lee et al., 1999). Les préséquences ont généralement une taille comprise entre 15 et 60 résidus. Cependant, certains signaux d'import font moins de dix résidus, alors que d'autres peuvent être très longs (une centaine de résidus). Quelle que soit la longueur des préséquences, leur repliement induit la formation d'une hélice amphipatique. Celle-ci est formée d'une face chargée positivement et d'une face hydrophobe (Roise et al., 1986).

Le déterminant d'import des protéines matricielles est la préséquence en N-ter, et plus précisément l'hélice amphipatique. Celle-ci est reconnue par les complexes TOM et TIM23 qui permettent l'import des pré-protéines dans la matrice mitochondriale. Dans un premier temps, la pré-protéine en cours d'import est prise en charge par le complexe TOM. En effet, la région hydrophobe de l'hélice amphipatique est reconnue de manière spécifique par Tom20 (Abe et al., 2000; Becker et al., 2005), alors que la région chargée positivement est, quant à elle, reconnue par Tom22. Finalement, la pré-protéine est transportée vers l'IMS par le canal Tom40. Elle est ensuite recueillie par le complexe TIM23. C'est plus précisément la s-u Tim50 qui fixe cette pré-protéine émergeant du complexe Tom40/Tom22 au niveau de l'IMS (Rahman et al., 2014). Après fixation de la pré-protéine, le canal Tim23 est activé (Meinecke et al., 2006) et la pré-protéine peut continuer son import grâce à l'interaction de Tim50 avec Tim21. Une autre s-u, Tim17, est retrouvée en étroite collaboration avec Tim23. Cependant la fonction de cette protéine n'est pas encore élucidée, elle joue probablement un rôle régulateur pour Tim23. Le potentiel de membrane est également très important dans l'import des pré-protéines matricielles au travers de la MMI. Le potentiel de membrane négatif au niveau de la matrice attire les pré-protéines *via* leur région N-ter chargée positivement (Martin et al., 1991; Turakhiya et al., 2016). De plus, ce potentiel permet l'activation de Tim23 (Truscott et al., 2001). Le potentiel de membrane favorise donc l'import de la pré-protéine au travers de Tim23, mais celui-ci n'est pas suffisant pour l'import de la totalité du précurseur matriciel dans ce sous-compartiment. En effet, le moteur PAM (*Presequence translocase-Associated Motor*) finalise l'import du précurseur dans la matrice mitochondriale.

La chaperonne mitochondriale ATP-dépendante $mHsp70$ forme le core de ce moteur (Horst et al., 1997; Kang et al., 1990) et interagit avec Tim44 qui fait le lien avec le complexe TIM23 (Banerjee et al., 2015). Le précurseur est finalement transféré à $mHsp70$. Les protéines Pam16 et Pam18 régulent alors l'activité d'hydrolyse de l'ATP de $mHsp70$ et l'ADP est recyclé par intervention du facteur d'échange de nucléotides, MgeI. Néanmoins, le mécanisme de fonctionnement du moteur PAM n'est pas encore élucidé et l'association de ce complexe avec d'autres protéines ou complexes, comme le complexe III-IV de la CR, pourrait réguler l'import des protéines matricielles (Wiedemann et al., 2007).

Certaines protéines de l'IMS possèdent également des préséquences en N-ter induisant leur recrutement par les complexes TOM et TIM23 et leur import est initié comme décrit précédemment pour les protéines matricielles. Arrivées au niveau de Tim23, les pré-protéines de l'IMS sont détournées de la voie d'import des protéines matricielles grâce à la protéine Mgr2 (Ieva et al., 2014). En effet, ces protéines, possédant une préséquence N-ter, arborent également un signal d'adressage hydrophobe, reconnu par Mgr2 qui va alors favoriser leur relargage dans l'IMS.

Les préséquences des protéines matricielles, et celles de protéines de l'IMS, sont éliminées par intervention de la peptidase hétérodimérique MPP (*Mitochondrial Processing Peptidase*). Ce clivage n'est pas directement lié à l'import des protéines au travers de la MMI, mais peut intervenir après ou lors du passage au travers du complexe TIM23 en fonction de la protéine importée. Les préséquences sont finalement dégradées par le peptidasome mitochondrial Cym1 (Johnson and Kent, 2006; Mossmann et al., 2014). Le signal d'import hydrophobe de certains précurseurs de l'IMS peut être éliminé par l'IMP (*Inner Membrane Peptidase*). Ces protéines sont alors relarguées dans l'IMS, comme le Cyt *b2*, ou elles restent ancrées dans la MMI, c'est le cas du Cyt *c2* (Glick et al., 1992).

Les néo-protéines matricielles se replient finalement dans leur forme mature et active par intervention de chaperonnes spécifiques ($mHsp70$ soluble). Cette protéine a donc deux rôles : (i) la fraction membranaire de $mHsp70$ permet la translocation de la pré-protéine dans la matrice et (ii) la forme soluble favorise le bon repliement de la protéine et évite l'agrégation (Horst et al., 1997; Kang et al., 1990).

b. L'import des protéines possédant des signaux d'adressage internes

Comme je l'ai décrit dans la partie précédente, les protéines matricielles, et certaines protéines de l'IMS, possèdent des préséquences clivables en N-ter. Cependant, toutes les protéines retrouvées dans ce compartiment n'arborent pas ce type de signaux et sont reconnues par d'autres transporteurs spécifiques. C'est notamment le cas des protéines caractérisées par la présence de motifs riches en Cys (CX₃C, CX₉C) qui vont, dans un premier temps, passer au travers du canal Tom40. Directement à la sortie de ce canal, les pré-protéines riches en Cys sont prises en charge par la protéine Mia40 (*Mitochondrial Intermembrane space import and Assembly*) qui forme des ponts disulfures avec ces pré-protéines afin de les stabiliser (Chacinska et al., 2004). En effet, cette protéine va jouer le rôle de récepteur intermembranaire directement à la sortie de Tom40 et capte les pré-protéines par interaction hydrophobe et formation de ponts disulfures intermoléculaires. L'activité de Mia40 va alors permettre la formation de ponts disulfures intramoléculaires au sein des pré-protéines et ainsi favoriser leur relargage dans l'IMS (Koch and Schmid, 2014). Finalement, Mia40 est réoxydé par Erv1 (*Essential for Respiration and Vegetative growth*) avec l'aide de Hot13 (*Helper Of Tim*) pour recommencer un cycle d'import.

En plus des protéines solubles de la matrice mitochondriale et de l'IMS, il existe également des protéines dont la destination est l'une des membranes mitochondriales (MME ou MMI). C'est notamment le cas des transporteurs de métabolites qui sont principalement retrouvés au niveau de la MMI et qui renferment des signaux d'import internes. En effet, ces transporteurs possèdent plusieurs séquences d'adressage non clivables réparties sur l'ensemble de la protéine (Brix et al., 1999; Endres et al., 1999). Le mécanisme par lequel ces protéines sont importées dans la mitochondrie n'a pas encore été totalement décrypté. Cependant, ces précurseurs sont pris en charge dans le cytosol par des chaperonnes (Hsp70 et 90), empêchant ainsi leur agrégation et favorisant leur interaction avec Tom70 (Bhangoo et al., 2007; Young et al., 2003). Le passage de ces précurseurs au niveau de Tom40 fait intervenir un mécanisme particulier. Les deux extrémités du précurseur restent dans le cytosol pendant que le reste de la protéine traverse le canal. L'extrémité N-ter de Tom40 va alors se transloquer du cytosol à l'IMS pour permettre le recrutement d'un complexe de chaperonnes (TIM9/10). Ce complexe emmène alors les précurseurs hydrophobes au complexe TIM22 qui

va incorporer les transporteurs dans la MMI grâce au potentiel de membrane (Rehling et al., 2003).

Au sein de la MME, on retrouve deux types de protéines transmembranaires : les tonneaux β et les hélices α . Les tonneaux β contiennent des signaux d'adressage caractéristiques en tige-boucle (Jores et al., 2016). Ils sont formés de deux brins β adjacents associés par une boucle et sont généralement retrouvés au niveau de l'extrémité C-ter. Ces précurseurs sont importés *via* le canal formé par Tom40 et sont directement pris en charge par le complexe de chaperonnes TOM9/10, empêchant leur agrégation. L'insertion de tonneaux β au niveau de la MME est catalysée par le complexe SAM (*Sorting and assembly machinery*). L'interaction entre le précurseur et le complexe SAM est médiée par le signal d'adressage en tige-boucle (Kutik et al., 2008), qui induit probablement l'ouverture du canal Sam50, favorisant ainsi l'insertion du tonneau β dans la membrane. Néanmoins, ceci est une hypothèse et la cartographie de l'insertion des tonneaux dans la membrane doit encore être effectuée pour mieux comprendre ce mécanisme. L'insertion des hélices α au sein de la MME est également un mécanisme partiellement connu. Cependant, les signaux d'import de ces protéines sont définis et retrouvés dans la région de la protéine qui est incorporée dans la membrane. En effet, l'insertion peut se faire au niveau des régions N- ou C-ter. Le signal d'adressage est alors retrouvé dans l'hélice insérée dans la membrane. En parallèle, certaines protéines ancrées dans la membrane peuvent incorporer plusieurs segments transmembranaires, et dans ce cas précis, la localisation du signal d'adressage est totalement inconnue. Le complexe MIM (*Mitochondrial IMport*) permet l'insertion des hélices N-ter et des hélices multiples dans la MME. Ces dernières sont d'abord reconnues par Tom70 qui transfère ces protéines au niveau du complexe MIM (Becker et al., 2011; Dimmer et al., 2012). Concernant les protéines ancrées par la région C-ter, aucun complexe d'import n'a pu être identifié. Leur insertion dépendrait donc de la composition lipidique de la membrane (Kemper et al., 2008; Krumpke et al., 2012).

6. Le contact entre la mitochondrie et les autres organelles

Au sein de la cellule, il existe une quantité pléthorique d'organelles ayant des fonctions bien définies et il est indispensable de coordonner ces fonctions cellulaires *via* l'établissement de dialogues inter-organelles. Les échanges inter-compartiments faisant intervenir des

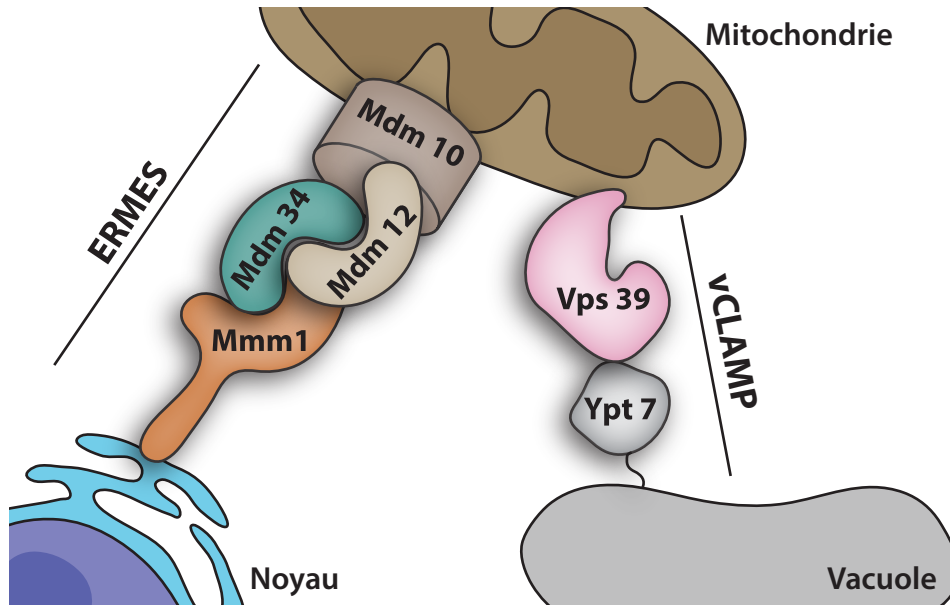


Figure 20 : Sites de contact entre la mitochondrie et les autres organelles

La mitochondrie interagit physiquement avec les membranes nucléaire et vacuolaire par l'établissement de MCS (*Membrane Contact Sites*). Le MCS entre le noyau et la mitochondrie fait intervenir la protéine du RE Mmm1 (*Mitochondrial Morphology Maintenance 1*), deux protéines de la MME, Mdm10 et Mdm34 (*Mitochondrial Distribution and Morphology 10, 34*) et la protéine cytosolique Mdm12. En parallèle, un MCS existe entre la vacuole et la mitochondrie. Celui-ci est médié par l'association entre la protéine Vps39 (*Vacuolar protein sorting*) et la GTPase Ypt7 (*Yeast Protein Two 7*).

vésicules sont étudiés depuis de nombreuses années et concernent la plupart des organelles. Cependant, la mitochondrie ne fait pas partie de ce trafic et sa capacité à communiquer avec d'autres compartiments a été élucidée récemment, grâce à la découverte de contacts directs inter-organelles. Ces interactions nécessitent la présence de MCS (*Membranes Contact Sites*) qui représentent des zones où les deux compartiments sont très proches (10 à 30 nm).

Ce sont des études chez la levure qui ont permis de mettre en évidence les contacts entre le réticulum endoplasmique (RE) et la mitochondrie par l'établissement d'un complexe appelé ERMES (*ER-Mitochondria Encounter Structure*) (Kornmann et al., 2009). Ce dernier, formant un pont entre les deux compartiments, est constitué d'une protéine de la membrane du RE (Mmm 1 : *Mitochondrial Morphology Maintenance 1*), de deux protéines de la MME (Mdm 10 et Mdm 34 : *Mitochondrial Distribution and Morphology 10, 34*) et d'une protéine cytosolique (Mdm 12) (**Figure 20**). En plus de ces protéines, il semblerait également que la GTPase Gem1 participe à la régulation du complexe ERMES (Kornmann et al., 2011; Stroud et al., 2011). Ce complexe a de nombreuses fonctions comme par exemple le transit de lipides du RE à la mitochondrie. A titre d'exemple, la phosphatidylsérine (PS), synthétisée au niveau de la membrane du RE, va être importée au sein de la MMI où la PS décarboxylase va permettre la formation de phosphatidyléthanolamine. Celui-ci est alors réexporté vers le RE pour former la phosphatidylcholine, le phospholipide le plus abondant des membranes de levure. Cependant, le complexe ERMES n'est pas le seul moyen d'échange de lipides entre le RE et la mitochondrie, car la déstabilisation de ce complexe n'est pas létale pour la levure. Ce complexe pourrait également permettre l'homéostasie du Ca^{2+} , et joue également un rôle dans l'autophagie, car il colocalise avec les sites de formation du mitophagosome et favorise l'isolation de la membrane de ce compartiment (Bockler and Westermann, 2014).

La mitochondrie étant un compartiment central, des MCS entre la MME et la membrane vacuolaire sont également retrouvés ; on parle de vCLAMP (*vacuole and mitochondria patch*). Ce complexe s'établit par association de la GTPase Ypt 7 (*Yeast Protein Two 7*) et de la protéine Vps 39 (*Vacuolar protein sorting 39*) (**Figure 20**). De manière intéressante, ce complexe est retrouvé proche du complexe ERMES et permet également le transit de lipides au sein de la mitochondrie. De plus, l'élimination de l'un ou l'autre de ces complexes augmente l'activité de celui restant et l'élimination concomitante de ces deux

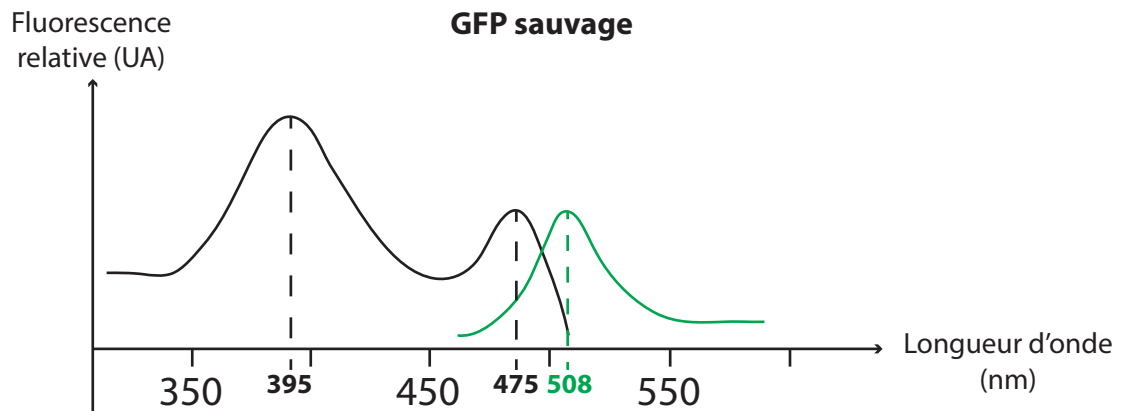
complexes est létale pour la levure (Elbaz-Alon et al., 2014). Le transit de lipides du RE à la mitochondrie, et inversement, est donc probablement médié par l'activité de ces deux MCS.

VIII. La Green Fluorescent Protein et ses applications

1. La découverte de la GFP

Dans les années 1960, les Dr. Johnson et Shimomura se sont intéressés à la lumière fluorescente émise par une méduse du Pacifique, *Aequorea victoria*. A cette époque, la communauté scientifique pensait que la lumière produite par les organismes bioluminescents émanait de la réaction entre la luciférine et la luciférase. Ils tentent alors, par les différents moyens disponibles, d'extraire ces deux molécules de l'ombrelle des méduses. Mais au bout de quelques jours, aucune de ces molécules n'avait pu être purifiée. Dans l'optique d'extraire une molécule luminescente, il est indispensable de trouver les conditions qui permettent l'inhibition du signal lumineux et Dr. Osamu Shimomura essaye donc tous les inhibiteurs connus, mais aucun n'est efficace. C'est finalement en diminuant le pH, que la molécule active est inhibée et que la luminescence est abolie. Grâce au pH acide du tampon d'extraction, il parvient à purifier une substance et lorsqu'il neutralise le pH de la solution, un signal de luminescence apparaît. Néanmoins, à sa grande surprise, lorsqu'il jette la substance à l'évier, il observe une vive luminescence bleue. Il comprend rapidement que les ions Ca^{2+} , présents dans l'eau de mer qu'il avait éliminé à l'évier, stimulent l'émission de cette lumière. Lors de l'été 1961, seuls quelques milligrammes de protéine luminescente purifiée ont été obtenus à partir de 10000 méduses. Cette protéine, émettant une lumière bleue vive en présence de Ca^{2+} , est alors nommée "aequorine" (Shimomura et al., 1962). De manière intéressante, durant la purification de l'aequorine, une autre protéine est co-purifiée et est nommée "*the green protein*" du fait de la production d'une fluorescence de cette couleur. Les Dr. Morin et Hastings la renomment finalement "GFP" (*Green Fluorescent Protein*) quelques années plus tard (Morin and Hastings, 1971). Au sein des méduses, l'ombrelle contient donc deux protéines fluorescentes, l'aequorine et la GFP. L'émission d'un signal par la première molécule permet l'excitation de la GFP, qui émet alors une fluorescence verte (Morise et al., 1974). Malgré le fait que la GFP soit aisément visualisable et cristallise facilement, la quantité de cette protéine est bien plus faible que celle de l'aequorine et des échantillons ont dû être récoltés durant de nombreuses années pour obtenir des quantités suffisantes de GFP permettant

A



B

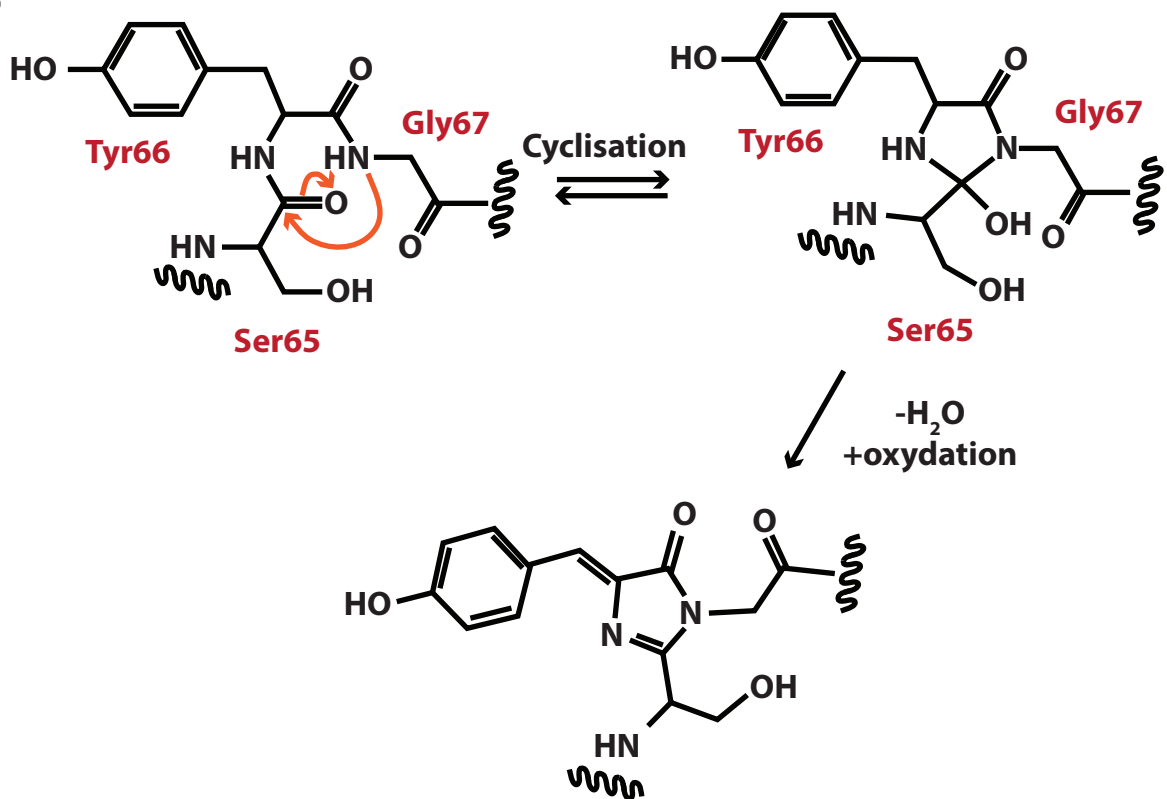


Figure 21 : Caractéristiques spectrales de la GFP sauvage et formation du chromophore

(A) La GFP sauvage possède un spectre d'absorption bimodal ayant un premier pic à 395 nm et un second à 475. Le pic d'émission maximal est, quant à lui, à 508 nm. (B) La formation du chromophore de la GFP est basée sur la cyclisation des résidus 65-66-67 et débute par une attaque nucléophile de la Gly67 sur le groupement carbonyle de la Ser65 suivi par une étape de déshydratation. Il y a alors déshydrogénation de la liaison α - β de la Tyr66 qui permet de mettre son cycle en conjugaison avec les deux autres résidus et ainsi former le chromophore actif.

d'étudier son chromophore. De manière intéressante, ce chromophore est semblable à celui de la luciférine de *Cypridina* (crustacées microscopiques) et Shimomura parvient à le décrire rapidement (Shimomura, 1979). Néanmoins, le chromophore est atypique car il ne ressemble pas à ceux identifiés à l'époque, qui consistent en un complexe entre une protéine et un composant fluorescent. A l'inverse, le chromophore de la GFP est retrouvé à l'intérieur de la structure. La structure du chromophore a été validée dans les années 1990 par Cody (Cody et al., 1993). Cette protéine étant très difficile à purifier, le clonage du gène codant la GFP a été une étape cruciale pour la caractérisation et l'utilisation de cette protéine (Prasher et al., 1985). Cette avancée a permis : (i) l'expression de la GFP dans des cellules vivantes comme *E. coli* ou *C. elegans* (Chalfie et al., 1994; Inouye and Tsuji, 1994) et (ii) l'ingénierie de variants de cette molécule par le Dr. Tsien (Heim et al., 1995; Tsien and Prasher, 1997). Les Dr. Shimomura, Chalfie et Tsien ont obtenu le prix Nobel de Chimie en 2008 pour leurs travaux sur la GFP.

2. La structure et les caractéristiques de la GFP

La GFP est une protéine de 238 aa ayant une taille apparente de 26,8 KDa (Prendergast and Mann, 1978). La GFP sauvage a une longueur d'onde d'excitation maximale de 395 nm et un pic mineur à 475 nm. Le pic d'émission, quant à lui, est à 508 nm (**Figure 21A**). Cependant, le spectre d'excitation a tendance à varier selon la concentration en protéine, suggérant qu'elle a tendance à s'agglomérer. De plus, cette protéine est très résistante à la dénaturation, stable sur une grande échelle de pH (pH 6 à 10) et sa renaturation complète ne dure que quelques minutes (Ward and Bokman, 1982).

Les prémices de l'étude de la structure de la GFP ont été réalisées par Shimomura à la fin des années 1970 lorsqu'il décrit, pour la première fois, le chromophore de cette molécule (Shimomura, 1979). Dans la protéine sauvage, le chromophore est formé par cyclisation de trois résidus, Ser65, Tyr66 et Gly67 (**Figure 21B**). Sa formation débute par une attaque nucléophile de l'amine de la Gly67 sur le groupement carbonyle de la Ser65, suivi directement par une étape de déshydratation. Finalement, l'oxygène permet de déshydrogéner la liaison α - β de la Tyr66 pour mettre son cycle aromatique en conjugaison avec les deux autres résidus et ainsi former le fluorophore actif (Cubitt et al., 1995; Heim et al., 1994; Reid and Flynn, 1997). La formation de ce tripeptide est post-traductionnelle, dépendante de l'oxygène et

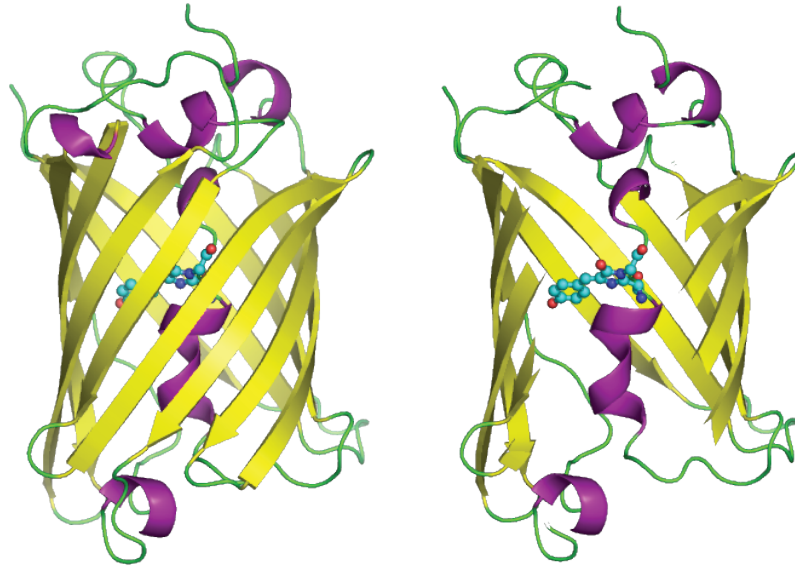


Figure 22 : Structure de la GFP (PDB : 1GFL)

La GFP est composée de onze brins β , représentés en jaune adoptant une structure en forme de tonneau. Les hélices sont colorées en violet. Le chromophore est positionné au milieu de ce tonneau et comprend les résidus 65-66-67.

autocatalytique (Heim et al., 1995; Heim et al., 1994; Niwa et al., 1996). Le résidu Gly67 est un très bon nucléophile. De manière intéressante, il est conservé au sein des chromophores de toutes les protéines fluorescentes du type GFP et favorise sa cyclisation.

Malgré le fait que la GFP ait été cristallisée pour la première fois en 1974 (Morise et al., 1974) et que les premiers clichés de diffraction ont été réalisés en 1988 (Perozzo et al., 1988), c'est au milieu des années 1990 que les premières structures complètes de la GFP ont pu être résolues (Ormo et al., 1996; Yang et al., 1996). Ces structures suggèrent que la GFP adopte une forme de cylindre composé de onze brins β et que le fluorophore est localisé exactement au milieu du tonneau (**Figure 22**). Le résidu 1 et les résidus 230 à 238, sont absents de la première structure résolue ; cette dernière région est probablement trop désordonnée pour être cristallisée. Cependant, leur absence n'impacte pas la production de fluorescence (Dopf and Horiagon, 1996), ce qui suggère que la structure obtenue coïncide avec la structure réelle. De manière intéressante, des molécules d'eau, ainsi que de nombreux groupements polaires sont adjacents au chromophore, notamment la Gln69 et la Thr203. L'environnement au sein du cylindre permet d'expliquer l'effet des différentes mutations existantes et suggère que les chaînes latérales des résidus mutés peuvent influencer sur les propriétés spectrales de la GFP (décrit dans la partie suivante).

Dans la structure de Yang, la GFP est sous forme de dimère et la dimérisation de la molécule implique aussi bien des résidus hydrophobes (Ala206, Leu221 et Phe223) que des résidus hydrophiles (Tyr39, Asn149, Arg168, Asn170, Glu172...). Cependant, la GFP peut également être cristallisée sous forme de monomère (Brejc et al., 1997; Ormo et al., 1996). La dimérisation de la GFP et de ses variants peut être éliminée pour faciliter son étude. D'ailleurs, la mutation de l'Ala206 en Lys (Ala206Lys) est suffisante pour favoriser une structure monomérique.

La GFP sauvage, aussi importante soit elle pour la compréhension du mécanisme de fluorescence, possède des inconvénients et notamment une faible émission de fluorescence, plusieurs pics d'émission et d'absorption et un délai de quatre heures après expression pour obtenir 100 % de fluorescence. Ces inconvénients ont cependant été éliminés par mutagénèses successives et ont permis l'élaboration d'une palette de variants optimisés de la GFP, favorisant également l'émergence de protéines fluorescentes bleue, cyan ou jaune...

	λ Excitation (nm)	λ Emission (nm)	Structure	Fluorescence relative (/eGFP)
GFP sauvage	395/475	509	Monomère*	48%
eGFP	484	507	Monomère*	100%
Emerald	487	509	Monomère*	116%
Azami green	492	505	Monomère	121%
TagGFP	482	505	Monomère*	110%
TurboGFP	482	502	Dimère	102%
mWasabi	493	509	Monomère	167%
AcGFP	480	505	Monomère*	82%
ZsGreen	493	505	Tétramère	117%
T-Sapphire	399	511	Monomère*	79%
mNeonGreen	506	517	Monomère	276%

Figure 23 : Les différents variants optimisés de la GFP

Chacun des variants de la GFP a été optimisé pour améliorer les propriétés spectrales et l'intensité de fluorescence des molécules. Ce tableau décrit les longueurs d'ondes (λ) d'excitation et d'émission, la structure oligomérique, ainsi que la quantité relative de fluorescence par rapport à l'eGFP. * correspond à la tendance à dimériser.

Malgré le fait que la plupart des mutations ponctuelles induisent la formation de protéines non fluorescentes, de nombreuses mutations permettent d'améliorer les propriétés spectrales de la GFP. L'augmentation de l'intensité de fluorescence, ainsi que la variation des longueurs d'ondes d'excitation et d'émission, ont favorisé l'utilisation simultanée de plusieurs molécules fluorescentes. Cette palette de fluorescence a rendu possible la visualisation de plusieurs processus cellulaires en parallèle, mais également la rationalisation des signaux de fluorescence pour la quantification relative d'une protéine donnée à un instant donné dans la cellule. Ces mutagenèses ont également augmenté la rapidité d'oxydation du fluorophore, réduisant le temps entre l'expression de la protéine et l'obtention d'un signal maximal de fluorescence.

La résolution de la structure du célèbre mutant de la Ser65Thr (Ormo et al., 1996) a permis au laboratoire du Dr. Remington de créer rapidement le variant YFP (*Yellow Fluorescent Protein*) par simple mutation de la Thr203 en Tyr. Une version améliorée de la GFP ayant un pic d'excitation unique aux alentours de 484 nm a été créée grâce à la mutation Ser65Thr dans le chromophore. Cette GFP mutée est, pour l'instant, la plus largement utilisée pour la visualisation de protéines dans les cellules vivantes et elle a donné naissance, en combinaison avec la mutation Phe64Leu, au variant appelé *enhanced Green Fluorescent Protein* (eGFP). Celui-ci est utilisé en microscopie à fluorescence avec le filtre fluorescéine (FITC) et présente une fluorescence forte et stable.

3. La palette de variants de la GFP

a. Le super folder GFP

Les études de mutagenèse ont permis de faire évoluer la palette de molécules fluorescentes dérivées de la GFP, afin d'obtenir des molécules ayant une intensité de fluorescence plus importante et des propriétés spectrales différentes (**Figure 23**). L'un d'entre eux est le *foldin reporter GFP* (frGFP) (Waldo et al., 1999), contenant les mutations du "cycle en 3'" (Phe99Ser, Met153Thr et Val163Ala) (Cramer et al., 1996) et les mutations de l'eGFP (Phe64Leu et Ser65Thr) (Patterson et al., 1997). Afin d'améliorer les propriétés spectrales du frGFP, celui-ci a été muté par la méthode d'ADN *Shuffling*. Le variant ainsi obtenu, nommé *super folder GFP* (sfGFP), produit un signal de fluorescence 1,6 fois plus

intense que celui produit par l'eGFP. Ce variant contient, en plus des mutations du frGFP, six autres mutations : Ser30Arg, Tyr39Asn, Asn105Thr, Tyr145Phe, Ile171Val et Ala206Val (Pedelacq et al., 2006). Ces mutations additionnelles ont également été ajoutées aux variants de couleurs de la GFP. Cependant, l'expression de ces variants chez *E. coli* n'induit pas la production d'un signal de fluorescence fort (Pedelacq et al., 2006). Le variant sfGFP est intéressant car après dénaturation, il ré-adopte sa structure native 3,5 fois plus vite que le variant frGFP. De plus, lorsque le sfGFP est étiqueté à des protéines peu ou pas solubles, il permet d'une part, d'obtenir un signal de fluorescence en fusion avec toutes les protéines testées, contrairement au frGFP, et d'autre part, il présente un signal plus intense, démontrant sa capacité à être utilisé pour la visualisation des protéines.

b. La Cyan Fluorescent Protein

La première molécule fluorescente émettant une fluorescence cyan (région spectrale entre 470 et 500 nm) a été découverte en même temps que celle émettant une fluorescence bleue (BFP, *Blue Fluorescent Protein*). Celle-ci possède la mutation Tyr66Trp. Cette molécule a, à la fois, un pic d'absorption bimodal (comme c'est le cas de la GFP sauvage) à 433 et 445 nm et un spectre d'émission bimodal (475 et 503 nm). Le variant *enhanced CFP* (eCFP) est obtenu par combinaison des mutations Phe64Leu, Tyr66Trp et Ser65Thr, qui permettent respectivement d'accélérer la maturation, d'augmenter l'intensité de fluorescence et de changer les propriétés spectrales (Heim et al., 1994). Les pics d'absorption et d'émission de ce variant sont de 439 et 476 nm respectivement. Cependant, ces mutations ne n'induisent pas une augmentation significative de l'intensité de fluorescence, comme c'est le cas de l'eGFP dont l'intensité de fluorescence est 50 % plus élevée que la GFP sauvage. Une autre mutagenèse a permis la création d'un variant de l'eCFP dit "haute performance", nommé la Cerulean, possédant les mutations des résidus 148 et 175 (His et Tyr respectivement) (Rizzo et al., 2004). Cette protéine a un état d'excitation simplifié et une intensité 1,5 à 2 fois plus forte que celle de l'eCFP. Ce variant est particulièrement utilisé lors des études utilisant le FRET (*Förster resonance energy transfer*) en combinaison avec l'YFP, car les diverses mutations permettent de réduire le bruit de fond et les spectres d'excitation et d'émission sont compatibles pour qu'une molécule puisse exciter la seconde. Comme pour la GFP, de nombreuses protéines fluorescentes cyan optimisées ont été élaborées. Une des molécules les plus performantes est la mTFP1 (Ai et al., 2006). Celle-ci possède des pics d'absorption et

d'émission à 462 et 492 nm respectivement et produit un signal de fluorescence très important (trois fois plus que l'eCFP). Cependant, cette molécule ne fait pas partie des variants GFP car elle est dérivée d'une protéine de corail (*Clavularia*).

c. La Yellow Fluorescent Protein

Parmi la multitude de protéines fluorescentes développées à ce jour, les protéines fluorescentes jaunes sont les plus polyvalentes. Leurs longueurs d'ondes d'émission et d'excitation sont comprises entre 514 et 529 nm, et en font les molécules fluorescentes ayant les plus petites régions d'excitations et d'émissions. L'établissement de la structure 3D de la GFP a révélé que la Thr203 est positionnée juste à côté du chromophore et une mutation de ce résidu permet potentiellement de changer les propriétés spectrales de la GFP. L'introduction de la mutation Thr203Tyr, en plus de la mutation Ser65Thr, a permis la création de l'*enhanced* YFP (eYFP) (Wachter et al., 1998). La forte intensité de fluorescence de cette dernière (1,5 fois plus d'intensité par rapport à l'eGFP), ainsi que son spectre d'excitation et d'émission, en font un bon candidat pour une utilisation dans l'étude de localisation en combinaison avec d'autres protéines fluorescentes et ce, malgré sa maturation lente, notamment dans les organelles (Llopis et al., 1998). De plus, l'utilisation de l'eYFP n'est pas recommandée lorsque le pH dépasse 6,5 car son intensité de fluorescence diminue de moitié. Cette protéine est également sensible à la concentration en ions Cl⁻ permettant son utilisation comme senseur (Jayaraman et al., 2000; Kuner and Augustine, 2000). Néanmoins, un variant de l'YFP, possédant la mutation Gln69Met, est plus résistant aux pH acides et à la variation de la concentration des ions Cl⁻. De plus, son expression est plus importante dans des cellules de mammifères, notamment lors de l'import dans les organelles (Griesbeck et al., 2001). Ce variant, nommé Citrine, a un maximum d'absorption de 516 nm et d'émission de 529 nm et a une intensité de fluorescence plus importante que celle de l'eYFP. Finalement, l'introduction de la mutation Phe46Leu permet d'augmenter drastiquement l'efficacité de maturation (Nagai et al., 2002) et donne naissance à Venus (son nom vient de l'intensité lumineuse de la planète).

4. L'utilisation de la GFP et de ses variants

De nos jours, les techniques utilisant des molécules fluorescentes sont multiples et c'est principalement la GFP et ses variants qui sont utilisés. En effet, cette palette de fluorescence offre la possibilité d'étudier la localisation subcellulaire de protéines et d'ARN (associés à une protéine), sur cellules uniques et même au sein d'organismes entiers. De plus, les techniques de fluorescence permettent l'étude de la dynamique cellulaire, mais également de l'interaction protéine/protéine. Les différentes utilisations de molécules fluorescentes sont décrites dans les paragraphes qui suivent.

a. Le marquage de protéines

L'utilisation de la GFP pour visualiser une protéine a été décrite pour la première fois dans les années 90 (Wang and Hazelrigg, 1994). Afin de produire une protéine de fusion avec la GFP (ou une autre protéine fluorescente), il est nécessaire de sous-cloner le gène d'intérêt en phase avec celui codant la GFP ou ses variants. De nos jours, l'expression d'une protéine de fusion pour étudier la localisation ou la relocalisation est une des applications la plus populaire des molécules fluorescentes, et est communément utilisée dans les laboratoires. Malgré cette utilisation récurrente, la fusion des protéines d'intérêts n'est pas aisée et de nombreux paramètres sont cruciaux pour bien choisir la molécule fluorescente à utiliser. En effet, lors de la fusion avec la GFP ou ses variants, l'étiquette fluorescente ne doit interférer, ni avec la localisation initiale de la protéine, ni avec sa fonction. Pour cela, il est conseillé de rajouter une séquence, généralement riche en Gly, associant l'étiquette et la protéine, pour éviter l'encombrement stérique pouvant impacter le repliement de la protéine de fusion. La GFP et ses variants sont principalement retrouvés sous forme de dimères. Néanmoins, la mutation Ala206Lys permet d'obtenir des protéines fluorescentes principalement monomérique qui dimérisent donc faiblement. Ces variants sont privilégiés lors de la fusion avec la protéine d'intérêt. Finalement, il est important de prendre en compte la photostabilité de la protéine fluorescente utilisée pour éviter l'absence de signal due au phénomène de photoblanchiment.

Afin d'étudier la localisation de protéines, l'étiquette peut être associée aux régions N- ou C-ter, mais l'incorporation au sein de la protéine est également envisageable (Baehler et

al., 2002; Moradpour et al., 2004; Rocheleau et al., 2003). De plus, malgré la spécificité du signal obtenu, il est indispensable de valider la localisation avec des anticorps spécifiques dirigés, soit contre la GFP, soit contre la protéine d'intérêt s'ils sont disponibles. Cette validation peut inclure l'utilisation directe d'un anticorps marqué sur des cellules fixées ou la purification du compartiment cellulaire ciblé, suivi de l'immuno-détection de la protéine d'intérêt. Le niveau d'expression d'une protéine donnée varie en fonction des conditions physiologiques de croissance, mais également en fonction des besoins de la cellule. En cas de faible expression, le signal fluorescent peut être difficilement observable. Ce niveau d'expression est d'autant plus important pour la visualisation simultanée de plusieurs protéines. Cependant, une trop forte surexpression de la protéine d'intérêt peut être néfaste pour la cellule et induire la perte de fonction ou encore la perte de la localisation réelle de la protéine. C'est pour cela qu'il est essentiel de bien optimiser tous les paramètres avant leur utilisation.

b. L'étude du photoblanchiment

Bien que le photoblanchiment soit en général un effet indésirable qui rend les études de fluorescence plus difficiles, ce phénomène peut être utilisé pour étudier la mobilité des protéines d'intérêts (Lippincott-Schwartz et al., 2003). La mobilité des protéines dans la cellule est très importante, car la localisation subcellulaire est souvent liée à la fonction de la molécule. Néanmoins, les techniques de photoblanchiment sont limitées pour la visualisation des protéines se déplaçant trop rapidement et peuvent être toxiques principalement à cause de l'utilisation de lampes de trop fortes intensités. La technique de photoblanchiment la plus utilisée est le FRAP (*Fluorescence Recovery After Photobleaching*) (Axelrod et al., 1976; Koppel et al., 1976), qui consiste à blanchir une région de la cellule par une irradiation intense et à observer la cinétique de réapparition du signal fluorescent. Il est également possible d'observer la disparition du signal de fluorescence dans une zone adjacente après blanchiment d'une région dans la cellule ; on parle de FLIP (*Fluorescence Loss In Photobleaching*). L'utilisation de ces différentes techniques nécessite la connaissance approfondie de la ou les molécules fluorescentes utilisées. Un des exemples les plus remarquables de l'utilisation des techniques de FRAP, est l'étude de la forte mobilité des protéines nucléaires qui a permis la compréhension de l'architecture de la dynamique nucléaire (Phair and Misteli, 2000).

c. L'étude de la dynamique cellulaire

L'étude de la localisation d'une protéine, de sa demi-vie, ainsi que du trafic intracellulaire des molécules néosynthétisées, est possible grâce à l'utilisation de molécules fluorescentes ayant des caractéristiques spécifiques. En effet, certaines molécules fluorescentes ont la capacité de changer de couleur au cours du temps et ces protéines peuvent être utilisées pour des études de dynamique sur cellules en culture, mais également au sein d'organismes entiers. La première molécule de ce type décrite est la DsRed-E5, qui produit un signal vert pendant de nombreuses heures, mais à plus long terme, émet un signal rouge. Le chromophore de la DsRed-E5 adopte, dans un premier temps, l'architecture des molécules fluorescentes de type GFP, et induit la production d'une fluorescence verte. Puis le chromophore rouge, ayant une maturation plus lente, se structure, et par un mécanisme de FRET, récupère l'énergie produite par le premier chromophore pour produire une fluorescence rouge (Verkhusha et al., 2004). Par conséquent, les organelles, cellules ou tissus néoformés exprimeront une DsRed-E5 verte, alors que les régions plus anciennes produiront une DsRed-E5 rouge. Cette technique permet de mesurer la dynamique de l'expression des gènes dans différents tissus (Mirabella et al., 2004), d'étudier des promoteurs (Miyatsuka et al., 2009), d'analyser la distribution des organelles (Solimena and Gerdes, 2003) et d'étudier le trafic intracellulaire (Czirok et al., 2006; Kozel et al., 2006). Les protéines fluorescentes multimériques ne sont généralement pas utilisées pour la fusion aux protéines d'intérêts. Pour ce faire, une palette de protéines dérivées de la *mcherry* (molécule fluorescente rouge) a été développée. Ces protéines produisent tout d'abord une fluorescence bleue, puis, à plus ou moins long terme (de quelques heures à plusieurs jours), le chromophore change de conformation et permet la production d'une fluorescence rouge. Ces techniques utilisant la palette de variant de la *mcherry* ont permis une meilleure compréhension du trafic cellulaire (Subach et al., 2009).

d. Le marquage de cellules, tissus ou organes

Sous dépendance d'un promoteur spécifique et dans un contexte génétique approprié, les protéines fluorescentes favorisent la détection de différents types cellulaires, de tissus ou encore d'organes dans un organisme entier. Cette capacité de détection est très importante, autant en immunologie, neurologie, qu'en biologie du développement. En effet, il a été

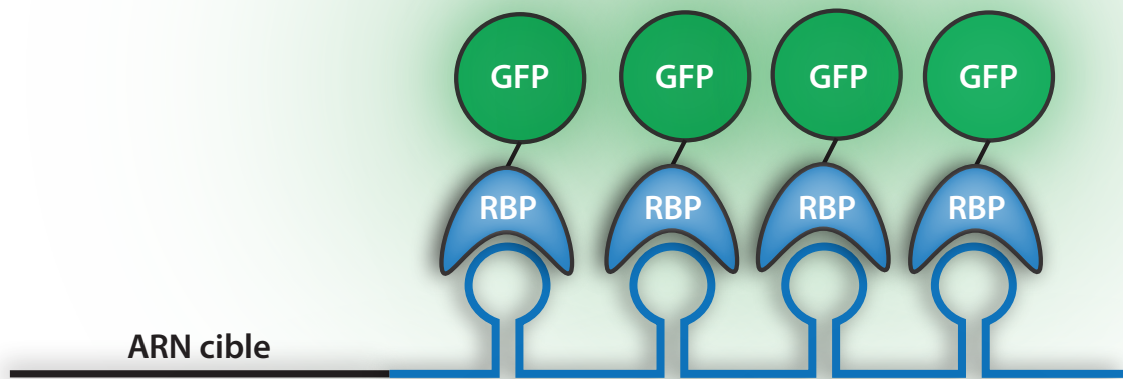
démonstré qu'un mélange de protéines fluorescentes, obtenu par recombinaison aléatoire, permet de produire jusqu'à 100 nuances de signaux de fluorescence détectables au microscope. Dans une étude de 2007, cette capacité a été exploitée grâce au système Cre/lox, pour produire une librairie de molécules fluorescentes exprimées individuellement dans des neurones afin de décrypter le circuit neuronal (Livet et al., 2007).

Il est également possible d'étudier le cycle cellulaire sur organisme entier, comme par exemple chez le nématode ou chez la souris. Pour cela, des protéines fluorescentes vertes et rouges ont été fusionnées à des motifs de dégradations spécifiques des étapes du cycle cellulaire. Les cellules se trouvant dans des cycles différents émettent donc un signal en antiphase ; les cellules en cycle G1 produisent une fluorescence rouge, alors que celles en cycle G2, S ou M sont vertes (Sakaue-Sawano et al., 2008). De manière intéressante, cette technique peut être utilisée sur des animaux transgéniques, permettant l'identification des cellules en division dans le cadre de l'étude du développement et de la progression de tumeurs.

e. Le marquage de l'ARN

L'étude de la synthèse, de la localisation et de la dynamique des ARNm est cruciale pour la compréhension de l'expression des gènes et de nombreux autres processus biologiques. Depuis de nombreuses années, les techniques d'hybridation *in situ* facilitent la visualisation d'ARN cellulaire sur des cellules fixées. Cependant, Ces techniques ne permettent pas d'observer la dynamique de ces molécules. Les techniques alliant les molécules fluorescentes et leur détection pour suivre la localisation d'un ARN, permettent le marquage spécifique d'un ARN donné dans une cellule vivante. Pour faciliter la visualisation des ARN, des molécules fluorescentes sont fusionnées à des protéines capables de reconnaître spécifiquement ces molécules (RBP, *RNA-binding protein*). Pour cibler de manière spécifique un ARN, il est indispensable d'exprimer (i) la protéine fluorescente, dans la plupart des cas la GFP, fusionnée à une RBP, ainsi que (ii) l'ARN cible fusionné au motif reconnu spécifiquement par la RBP. Cette dernière va se fixer à l'ARN pour lequel elle est spécifique et va ancrer la GFP à l'ARN cible. Il est alors facile de visualiser la localisation de la molécule d'ARN par microscopie à fluorescence. Ces techniques peuvent être utilisées, soit par expression

A



B

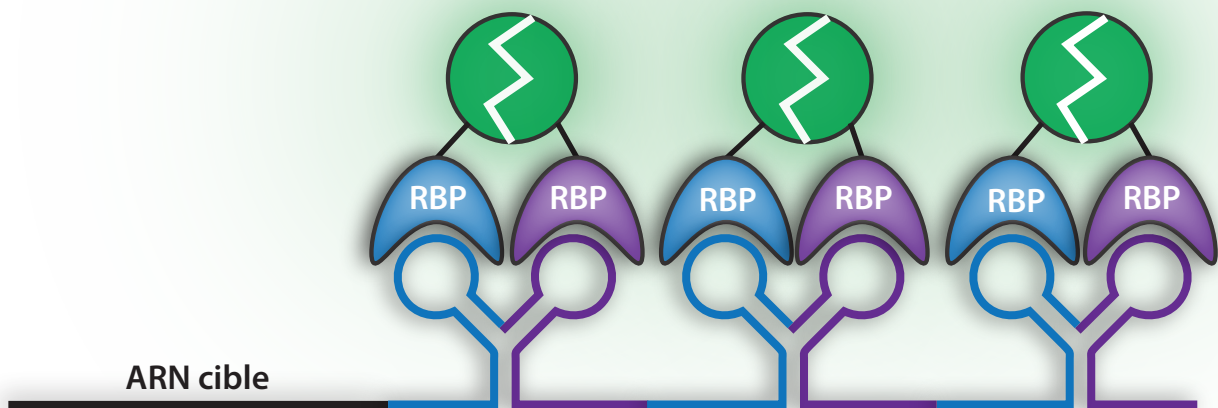


Figure 24 : Marquage des ARN cellulaires

(A) Pour déterminer la localisation d'un ARN donné, il est fusionné à un motif ARN reconnu spécifiquement par une protéine (RBP, *RNA Binding Protein*). Cette dernière est alors exprimée en fusion avec une protéine fluorescente (GFP par exemple). La localisation est finalement visualisée par microscopie à fluorescence. (B) Pour réduire le bruit de fond, une technique de Split-GFP peut être utilisée. Chacun des deux fragments de cette protéine est ainsi fusionné à un domaine de reconnaissance d'un motif ARN. Lors de la fixation des RBP sur l'ARN cible, il y a reconstitution de la molécule fluorescente et émission d'un signal. Les RBP libres n'émettent donc pas de signal.

transitoire dans des cellules en culture, soit par expression stable dans des lignées cellulaires spécifiques ou même des organismes entiers.

Le système MS2 est très utilisé pour détecter des ARN cellulaires. Celui-ci est basé sur la capacité de la protéine d'enrobage du bactériophage MS2 de lier, avec une forte affinité, une structure en tige boucle unique retrouvée au sein du génome à ARN de cet organisme. En pratique, la protéine MS2 est fusionnée à la GFP et la tige-boucle de l'ARN MS2 est associée à l'ARN cellulaire dont la localisation veut être étudiée. De plus, afin de faciliter la détection de l'ARN, une répétition du motif reconnu par la protéine de MS2 est liée à l'ARN cible, permettant ainsi de fixer plusieurs GFP au même endroit dans la cellule et de faciliter sa visualisation par concentration du signal fluorescent (**Figure 24A**). Une technique similaire a été développée à partir du bactériophage λ et est basée sur la capacité d'un peptide de 22 aa, appelé la BoxB, de reconnaître le motif ARN BoxB (Daigle and Ellenberg, 2007). Cette découverte est très intéressante car elle a permis la détection simultanée de deux ARNm par fusion des RBP à deux molécules fluorescentes verte ou rouge (Lange et al., 2008).

Cette technique de marquage des ARN a prouvé son efficacité pour différentes études, tant chez la levure *S. cerevisiae* (Beach et al., 1999; Bertrand et al., 1998), que chez la mouche *D. melanogaster* (Jaramillo et al., 2008; Weil et al., 2006; Zimyanin et al., 2008) ou dans des neurones de rats (Dynes and Steward, 2007; Rook et al., 2000). L'inconvénient principal de la technique reste néanmoins le bruit de fond produit par les molécules fluorescentes, libres dans le cytosol, pouvant interférer avec la capacité de détection des ARN. Pour remédier à cela, il est cependant possible de diminuer l'expression de la protéine fluorescente, ou de forcer son import dans le noyau, lors de l'étude d'ARN nucléaires (Fusco et al., 2003). Finalement, il est également possible d'utiliser des protéines fragmentées, notamment la Split-GFP. Pour mettre en œuvre cette technique, la protéine fluorescente est séparée en deux et chacun des fragments est associé à la RBP reconnaissant l'ARN d'intérêt. Après interaction entre la RBP et l'ARN cible, les deux fragments se trouvent dans un environnement proche et s'autoassemblent pour reconstituer la GFP active et produire un signal fluorescent (Ozawa et al., 2007; Rackham and Brown, 2004; Valencia-Burton et al., 2007) (**Figure 24B**). Les fragments de la protéine fragmentée, libres dans le cytosol ne produisent donc pas de fluorescence et le bruit de fond est considérablement diminué.

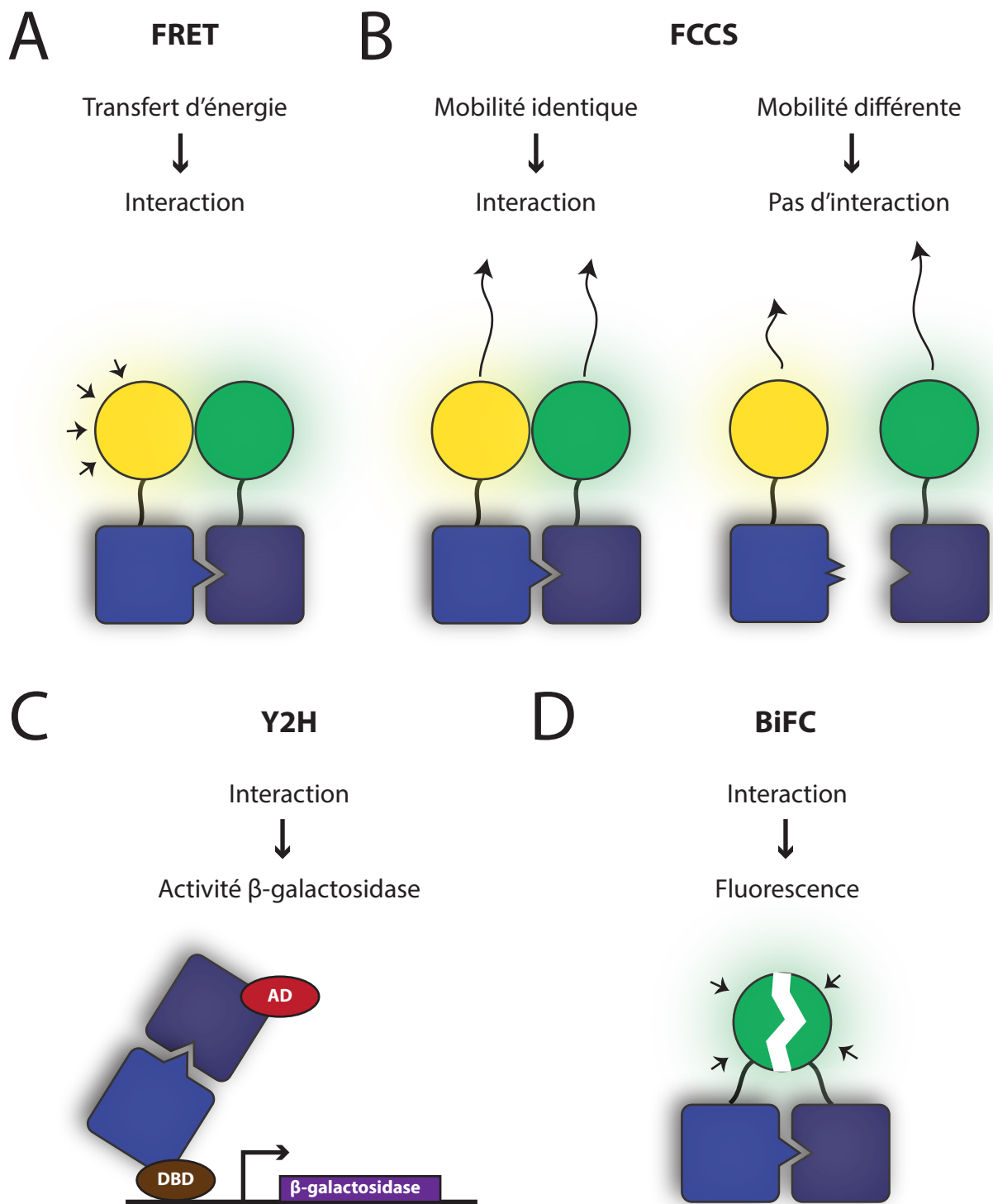


Figure 25 : Méthodes d'étude d'interaction protéine/protéine

L'interaction entre deux protéines peut être étudiée par divers mécanismes : (A) le FRET (*Förster resonance energy transfer*) basé sur le transfert d'énergie d'une molécule fluorescente à la seconde lors de l'interaction entre deux partenaires ; (B) Le FCCS (*Fluorescence Cross-Correlation Spectroscopy*) qui permet de mesurer le taux de diffusion des molécules ; (C) Le Y2H (*Yeast two Hybrid*) qui mesure la capacité d'activation de la transcription lors de l'interaction entre les deux partenaires ; (D) Le BiFC (*Bimolecular Fluorescence Complementation*), basé sur la capacité de reconstitution d'une Split-GFP lors de l'interaction entre les deux partenaires. AD : *Activating Domain* ; DBD : *DNA Binding Domain*.

f. L'étude de l'interaction protéine/protéine

L'identification de partenaires protéiques représente une étape importante dans la compréhension de divers processus indispensables au bon fonctionnement de la cellule. Les protéines fluorescentes peuvent être utilisées de différentes manières pour obtenir des informations sur l'interactome de la protéine d'intérêt. C'est notamment le cas du FRET, de la FCCS (*Fluorescence Cross-Correlation Spectroscopy*) ou encore des PCA (*Protein complementation Assays*) (**Figure 25**).

Le FRET est la technique d'étude de l'interaction entre deux protéines la plus connue et utilisée (Piston and Kremers, 2007). Le principe est le suivant : deux protéines d'intérêts dont on veut tester l'interaction sont chacune fusionnée à une protéine fluorescente (donneuse ou acceptrice). Lors de l'interaction entre les deux protéines, les molécules fluorescentes se trouvent dans un environnement proche ; la distance ne doit pas excéder 10 nm pour permettre l'apparition d'un signal. La première molécule va alors être excitée à une certaine longueur d'onde. Celle-ci va transférer son énergie à la seconde molécule fluorescente qui va passer à un état excité et va produire un signal fluorescent visualisable par microscopie, par spectroscopie ou par cytométrie en flux (Siegel et al., 2000) (**Figure 25A**).

La FCCS est une méthode alternative de détection d'interaction entre deux protéines et est basée sur la *Fluorescence Correlation Spectroscopy* (FCS). La FCS détecte les fluctuations d'intensité de fluorescence à l'échelle de la molécule unique au sein d'un petit volume permettant ainsi la mesure des concentrations moyennes ou le taux de diffusion d'une molécule dans la cellule. Dans la technique FCCS (deux couleurs), la variation de signal est mesurée simultanément pour deux fluorophores (Bacia et al., 2006; Schwille et al., 1997). La corrélation croisée correspond alors au degré où les deux signaux fluctuent ensemble et permet d'estimer si la diffusion des deux protéines est indépendante ou non (**Figure 25B**). Si elle ne l'est pas, les deux partenaires interagissent (Baudendistel et al., 2005; Kim et al., 2005; Kim et al., 2004b). Pour réaliser une excitation efficace, il est nécessaire d'avoir deux lasers alignés sur le même plan focal. Cependant cette opération est techniquement difficile. Ainsi, l'utilisation de deux protéines fluorescentes étant excitées à la même longueur d'onde, mais qui fluorescent à une longueur d'onde différente, est privilégiée. Contrairement au FRET, la

détection de l'interaction entre deux protéines avec la FCCS ne dépend pas de la proximité des deux protéines et donne donc plus de liberté pour la construction des protéines de fusion.

L'interaction entre deux protéines peut également être visualisée par des méthodes utilisant la capacité de deux parties d'une molécule fragmentée à s'associer pour restaurer l'activité de la molécule sauvage. Plusieurs techniques appartiennent à cette famille et c'est notamment le cas du Y2H (*Yeast Two hybrid*), mais également un dérivé de celle-ci, le BiFC (*Bimolecular Fluorescence Complementation*). Le Y2H a été décrit pour la première fois par Fields et Song (Fields and Song, 1989) et a été l'objet d'une attention particulière pour l'étude d'interactions fortes entre deux protéines. Cependant, cette technique peut également être utilisée pour l'étude d'interactions faibles et transitoires. Cet outil repose sur la capacité des facteurs de transcription à induire l'expression des gènes. En effet, ces protéines contiennent deux domaines principaux : un domaine de fixation à l'ADN (DBD, *DNA binding Domain*) et un domaine d'activation (AD, *Activating domain*) permettant le recrutement de la machinerie transcriptionnelle. La technique de Y2H est basée sur la séparation des domaines AD et DBD et leur association (fusion) avec deux protéines dont on veut tester l'interaction. Si ces deux protéines de fusion sont en interaction, le domaine DBD induit la fixation à l'ADN et le domaine AD recrute la machinerie nécessaire à l'expression du gène d'intérêt (**Figure 25C**). Ce gène d'intérêt (β -galactosidase, la plupart du temps) possède, dans ses régions promotrices, les séquences reconnues par DBD et la fixation des deux protéines favorise l'expression de cette protéine et la visualisation d'une coloration bleue en présence de X-Gal et d'IPTG (Isopropyl β -D-1-thiogalactopyranoside) (**Figure 25C**). De manière intéressante, cette technique a évolué pour étudier l'interaction de protéines membranaires, mais également de protéines se fixant à l'ADN ou l'ARN (Causier and Davies, 2002; Petschnigg et al., 2014; Reece-Hoyes et al., 2011). L'utilisation d'un dérivé de la technique Y2H, basé sur la séparation d'une molécule fluorescente à la place du facteur de transcription, a donné naissance au BiFC. Le principe est donc le même et la protéine fluorescente est séparée en deux fragments, incapables de produire de signal fluorescent. Chacun des fragments est alors fusionné à une protéine d'intérêt et c'est uniquement l'interaction entre ces deux protéines qui permet la reconstitution de la molécule fluorescente active et l'émission d'un signal (**Figure 25D**). L'utilisation de cet outil a été publiée pour la première fois par Regan et ses collaborateurs (Ghosh et al., 2000). Dans cette étude, ils ont séparé la GFP entre les résidus 157 et 158

(fragments N-GFP et C-GFP). Ces fragments sont incapables de s'assembler pour former la GFP active lorsqu'ils sont exprimés chez *E. coli*. Cependant, lorsqu'ils sont fusionnés à des glissières à Leucine (LZ), il y a reconstitution de la GFP active et production d'une fluorescence verte (Magliery et al., 2005). Quelques années plus tard, cette technique a été utilisée chez la levure. En effet, comme dans l'étude de Regan, l'eGFP a été séparée entre les résidus 157 et 158. La reconstitution de la molécule active permet alors la détection de l'interaction comme le système Y2H de levure (Park et al., 2007). La même année, celle-ci est utilisée avec une YFP séparée entre les résidus 154 et 155 pour détecter l'interaction de LZ chez la levure. C'est dans cette étude que le terme BiFC est utilisé pour la première fois (Sung and Huh, 2007). Chez cet unicellulaire, l'utilisation d'une BiFC, basée sur la YFP, a permis l'étude de la quasi-totalité de l'interactome des protéines avec des protéines sumoylées, par création de deux banques permettant d'étiqueter, soit le fragment N-ter (VN), soit le fragment C-ter (VC) de la Venus (variant de l'YFP) (Sung et al., 2013). Finalement, comme l'ajout d'une étiquette de grande taille peut modifier la capacité d'interaction entre deux protéines du fait de l'encombrement stérique, une GFP tripartite a été développée. La GFP a donc été séparée en trois fragments distincts : β 1-9, β 10 et β 11. Ces deux derniers fragments sont alors fusionnés à deux protéines qui peuvent potentiellement interagir ensemble et le grand fragment (β 1-9) est libre. C'est uniquement si les deux protéines interagissent ensemble que la reconstitution de la GFP active induit la production d'un signal (Cabantous et al., 2013).

g. L'étude de dynamique cellulaire par la Split-GFP

L'étude de la localisation subcellulaire d'une ou plusieurs protéines d'intérêts est facilitée par les progrès dans la variété des molécules fluorescentes disponibles et la sensibilité des techniques de détection utilisée pour la visualiser. Cependant, l'utilisation de ces molécules a des inconvénients et la protéine de fusion peut être insoluble (Nixon and Benkovic, 2000; Ullmann et al., 1967; Wehrman et al., 2002). De plus, certaines de ces techniques ne fonctionnent pas sur des cellules vivantes (Kelemen et al., 1999; Kim and Raines, 1993; Richards and Vithayathil, 1959). L'utilisation d'une technique de Split-GFP permet de remédier à ces problèmes. Néanmoins, la séparation de la GFP au centre de la molécule n'est pas optimale et d'autres variants de la Split-GFP ont été élaborés. C'est notamment le cas pour l'étude de la solubilité de protéines, où la GFP a été séparée en fragments de tailles très différentes (Cabantous et al., 2005a). En effet, cette molécule,

composée de onze brins β , a été séparée entre le dixième et le onzième brin, formant ainsi un grand fragment β 1-10 et un petit fragment β 11 (Cabantous et al., 2005b). Le petit fragment est, la plupart du temps, fusionné à la protéine d'intérêt et les deux protéines sont séparées par un *linker* (Gly-Gly-Gly-Ser)₂ afin d'éviter l'encombrement stérique. Pour augmenter la capacité de complémentation et leur solubilité, les deux fragments, issus de la sfGFP décrit dans la partie VIII.3.a, ont subi des mutagénèses successives générant les variants β 1-10 opt et β 11 M3. Le fragment β 1-10 opt possède donc les mutations du sfGFP (Ser30Arg, Tyr145Phe, Ile171Val et Ala206Val) et sept nouvelles mutations (Asn39Ile, Thr105Lys, Glu111Val, Ile128Thr, Lys166Thr, Ile167Val et Ser205Thr), augmentant la solubilité de ce fragment de 50 %. En parallèle, le fragment β 11 M3 possède trois nouvelles mutations (Leu221His, Phe223Tyr, Thr225Asn), permettant d'augmenter la solubilité de la protéine d'intérêt fusionnée au fragment β 11 (Cabantous et al., 2005a).

Cet outil de Split-GFP a été utilisé pour diverses applications, tant au niveau de la quantification et l'étude de solubilité de protéines (Cabantous et al., 2005a), que pour l'étude de contact entre cellules (Feinberg et al., 2008) et de l'assemblage de complexes protéiques *in vitro* (Kim et al., 2015). La Split-GFP est également un outil puissant pour l'étude de la localisation subcellulaire des protéines et l'imagerie moléculaire (Hyun et al., 2015; Kaddoum et al., 2010; Pinaud and Dahan, 2011; Van Engelenburg and Palmer, 2010). En effet, pour étudier la localisation subcellulaire chez les cellules de mammifères, la protéine d'intérêt peut être fusionnée au fragment β 11. Après croissance des cellules, celles-ci sont fixées et perméabilisées avant ajout de la protéine recombinante β 1-10 (purifiée de chez *E. coli*). Finalement, après 3 h d'incubation à température ambiante, la localisation de la protéine étiquetée est visualisée par microscopie à fluorescence et les zones fluorescentes correspondent aux compartiments où se trouvent la protéine d'intérêt (Kaddoum et al., 2010). En parallèle, cet outil peut également être utilisé pour étudier la topologie membranaire et la compartimentalisation sur des cellules vivantes. Généralement, c'est le fragment β 1-10 qui est ancré à la membrane et le fragment β 11 est fusionné aux différentes protéines d'intérêts. La Split-GFP a permis l'étude de la relocalisation de protéines lors de l'infection d'érythrocytes par *P. falciparum* (Kulzer et al., 2013), mais est également utilisée dans des cellules de mammifères (Kamiyama et al., 2016). De plus, dans cette dernière étude, le système est

amélioré par fusion du fragment β 11 en tandem (jusqu'à sept copies) et permet d'augmenter fortement la fluorescence produite.

L'utilisation de la Split-GFP est donc en pleine expansion pour l'étude de la dynamique cellulaire. Cependant, la détermination de la localisation subcellulaire des protéines est de plus en plus complexe, car les protéines ne sont plus de simples molécules assurant une seule fonction dans un compartiment unique. En effet, un nombre croissant d'étude suggère la présence simultanée de protéines, notamment les aaRS, dans des compartiments subcellulaires distincts. Au sein de ces compartiments, elles participent, la plupart du temps, à des processus biologiques essentiels pour l'homéostasie cellulaire. A l'heure actuelle, la détection de ces localisations additionnelles est compliquée du fait de la limitation des méthodes de détection disponibles. L'élaboration d'un outil basé sur la Split-GFP pourrait permettre de remédier à ces difficultés et donc faciliter l'étude des localisations et fonctions non canoniques des protéines.

Projet de thèse

Mon travail de thèse s'est décomposé en plusieurs axes de recherches : (i) l'étude de la dynamique du complexe AME de levure ; (ii) l'ingénierie d'un outil, basé sur la Split-GFP, permettant l'étude de la fraction organellaire d'une protéine multi-localisée et son utilisation pour délimiter les régions impliquées dans la relocalisation de diverses protéines d'intérêts ; (iii) l'étude de la GluProRS^{N1}, un variant de l'aaRS humaine, potentiellement impliquée dans la formation du Gln_{-m}ARnt^{Gln}.

i. Etude de la dynamique du complexe AME

Comme je l'ai décrit dans l'introduction, les aaRS forment une famille d'enzymes ubiquitaires dont la fonction principale est la formation des aa-ARnt, indispensables lors de la synthèse protéique. Les aaRS sont capables de s'associer entre elles, mais également avec des protéines non synthétasiques permettant ainsi la formation de MSC. Ces complexes forment alors des plateformes de rétention pour des aaRS capables de se détacher dans des conditions spécifiques afin d'assurer des fonctions additionnelles. La dynamique de libération des aaRS est très bien décrite chez les mammifères et a permis la compréhension de nombreux processus biologiques. Chez la levure *S. cerevisiae*, le petit MSC nommé le complexe AME, formé de la GluRS, de la MetRS et de la protéine d'ancrage Arc1, a également fait l'objet d'études approfondies. En revanche, aucune donnée concernant la dynamique de dissociation du complexe n'a été décrite. Néanmoins, Mathieu Fréchin, ancien doctorant au laboratoire, a montré la diminution de la quantité d'Arc1 lorsque la levure passe d'un métabolisme fermentaire à respiratoire, mais également la capacité d'import mitochondriale de la GluRS (Frechin et al., 2009b).

De nombreux questionnements restent cependant en suspend comme le mode de régulation de la quantité de l'ancre ou encore le devenir de la MetRS, libre dans le cytosol. Le premier projet auquel j'ai participé est basé sur l'étude de la dynamique du complexe AME. En outre, nous avons pu décortiquer le mécanisme de régulation de l'expression de la protéine Arc1 et le devenir de la MetRS cytosolique qui se relocalise dans un autre compartiment subcellulaire. Finalement, toutes les données récoltées nous ont permis de comprendre la fonction biologique de la relocalisation synchrone des deux aaRS et tout ceci est décrit dans la partie I des résultats. De plus, cette étude a fait l'objet d'un article publié en 2014 et intitulé :

" Expression of nuclear and mitochondrial genes encoding ATP synthase is synchronized by disassembly of a multisynthetase complex", dont je suis co-auteur.

ii. Elaboration d'un outil de localisation basé sur la Split-GFP

L'étude de la dynamique du complexe AME nous a permis de comprendre le mode de régulation de l'expression des génomes nucléaire et mitochondrial, mais également la difficulté de l'étude de protéines multi-localisées. En effet, lors des études de localisation subcellulaire, la fusion GFP est majoritairement utilisée pour la visualisation de la protéine d'intérêt. Cependant, ces techniques ne sont pas adaptées pour l'étude de protéines multi-localisées et d'autant plus, lorsque la localisation principale de la protéine est le cytosol. La fluorescence issue de ce compartiment empêche effectivement la visualisation de la fraction organellaire et de ce fait masque les potentielles autres localisations. Au laboratoire, nous avons décidé de créer un outil permettant de visualiser uniquement la fraction organellaire d'une protéine multi-localisée.

En collaboration avec le Dr Yuhei Arais, post-doctorant au laboratoire, nous avons décidé d'utiliser une technique de Split-GFP. Celle-ci est basée sur la séparation de la GFP en deux fragments distincts : le fragment β 1-10, composé des dix premiers brins β de la GFP et le fragment β 11, constitué uniquement du dernier brin. Nous avons donc optimisé tous les composants de l'outil pour leur expression chez la levure, construit les plasmides d'expression et sélectionné les ancres permettant la rétention d'un des fragments (β 1-10) dans un compartiment donné. Nous avons également créé une souche, où le fragment β 1-10 est directement exprimé par le génome mitochondrial. J'ai alors confirmé : (i) la capacité de reconstitution des fragments, (ii) le maintien de l'activité de la protéine de fusion lors de l'étiquetage et (iii) la spécificité de l'outil. D'ailleurs, j'ai utilisé la souche exprimant le fragment β 1-10 à partir du génome mitochondrial pour étudier des protéines provenant d'autres organismes (notamment humaines) et délimiter les signaux d'import mitochondriaux de la GluRS de levure et de GatF, une s-u de l'AdT mitochondriale. Finalement, j'ai également pu montrer une relocalisation inattendue pour la protéine Arc1. La mise au point de l'outil, ainsi que son utilisation pour la visualisation de la relocalisation des aaRS de levure, de la protéine d'ancrage du complexe AME et des protéines hétérologues, sont décrites dans les parties II et III des résultats.

iii. Etude de la fonction mitochondriale de la GluProRSN1 chez la levure

Lors de ma thèse, j'ai également participé à un dernier projet basé sur l'étude hétérologue d'un variant d'aaRS humain chez la levure. Comme je l'ai décrit dans l'introduction, chez certains organismes, deux activités d'aaRS sont présentes au sein de la même enzyme et forment l'aaRS bifonctionnelle nommée la GluProRS. Cette enzyme de mammifère est au centre de toutes les attentions, car elle assure différentes fonctions essentielles pour la régulation de la synthèse protéique lors de l'inflammation, la résistance à l'infection virale et la lipogenèse (**Figure 15**). Lors de la réponse aux IFN γ , un variant de cette enzyme, dépourvu de la partie ProRS, est synthétisé. Celui-ci, nommé GluProRS^{N1}, comprend donc la partie GluRS, ainsi que la quasi-totalité de deux domaines WHEP ; l'association entre les domaines catalytiques de la GluRS et la ProRS étant médiée par la répétition de trois de ces domaines. De manière intéressante, la voie de formation du Gln-mARNt^{Gln} est conservée et indirecte chez la levure et l'Homme. Il y en a en effet coordination entre l'activité d'une GluRS ND qui forme le Glu-ARNt^{Gln} et d'une AdT mitochondriale qui génère le Gln-mARNt^{Gln} par transamidation du Glu en Gln (Frechin et al., 2009b; Nagao et al., 2009).

La GluProRS^{N1} étant très proche de la GluRS de levure, nous sommes partis de l'hypothèse que ce variant pouvait potentiellement (i) être importé dans la mitochondrie et (ii) assurer la formation du Glu-mARNt^{Gln} et ce, malgré les données de la littérature montrant que la GluRS mitochondriale humaine assure cette fonction (Nagao et al., 2009). Dans un premier temps, la localisation mitochondriale de la GluProRS^{N1} étiquetée GFP a été validée dans des cellules HeLa par microscopie confocale. Ces résultats sont également confirmés par des immuno-détections de la GFP sur des extraits totaux mitochondriaux. Nous avons alors étudié la capacité de cette enzyme à remplacer l'enzyme de levure par différentes techniques de chasse et par dissections de tétrades. La localisation subcellulaire de la protéine étiquetée GFP a, de plus, été étudiée par microscopie à fluorescence et j'ai réalisé tous les témoins nécessaires à la validation de nos résultats. Finalement, j'ai également étudié la localisation mitochondriale de la protéine avec la souche exprimant le fragment β 1-10 à partir du génome mitochondrial. Les résultats, somme toute très surprenants, sont présentés dans la partie IV des résultats.

Matériels et méthodes

I. Matériel biologique et milieux de culture

1. Souches bactériennes

	Géotypes	Utilisations
DH5α	F ⁻ endA1, glnV44, thi-1, recA1, relA1, gyrA96, deoR, nupG, purB20, φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK-mK+), λ ⁻	Transformations BP/LR
XL1-Blue	endA1, gyrA96(nalR), thi-1, recA1, relA1, lac, glnV44, F' [::Tn10 proAB+ lacIq Δ(lacZ)M15] hsdR17(rK- mK+)	Transformations clonage classique
DB3.1	gyrA462 endA1 Δ(sr1-recA) mcrB mrr hsdS20 glnV44 (=supE44) ara14 galK2 lacY1 proA2 rpsL20 xyl5 leuB6 mtl1	Transformation vecteurs vides Gateway®

2. Milieux de culture bactériens

Les milieux de culture liquides utilisés pour la croissance des souches bactériennes sont présentés dans le tableau ci-dessous. Les milieux solides sont supplémentés d'Agar-Agar (Euromedex) à raison de 15 g/L. Ces milieux solides ou liquides sont alors autoclavés à 120 °C sous une pression de 1,2 bars pendant 20 min. Si nécessaire, les antibiotiques sont additionnés au milieu solide, lorsque la température est d'environ 50 °C, avant de couler les boîtes. Les antibiotiques sont ajoutés au milieu liquide juste avant leur utilisation. Les milieux sont finalement conservés à 4 °C jusqu'à utilisation.

	Composition	Utilisations
LB	10 g/L de peptone, 5 g/L d'extrait de levures, 10 g/L de NaCl	Régénérations
LB + Kan	10 g/L de peptone, 5 g/L d'extrait de levures, 10 g/L de NaCl, 50 mg/L de Kanamycine	Transformations réactions BP
LB + Amp	10 g/L de peptone, 5 g/L d'extrait de levures, 10 g/L de NaCl, 100 mg/L d'Ampicilline	Transformations réactions LR et clonages classiques

3. Souches de levure

Tout au long de mon doctorat, j'ai été amené à utiliser diverses souches de levure pour la transformation des plasmides d'expression que j'ai construit. Celle-ci sont décrites dans le tableau ci-contre. Les souches *gus1Δ Shuffle* est dérivée d'une souche *RS453*, la *dps1Δ Shuffle* est dérivée de la souche *YAMB4*, tandis que la souche *gus1Δ* diploïde est issue d'une *BY4742*. De plus, au laboratoire, nous possédons un mutant de cette dernière possédant un marqueur d'auxotrophie additionnel (*trp1*). Pour les intégrations génomiques des fragments β1-10 dans la souche *BY4742*, les linéarisations ont été réalisées au niveau des régions promotrices ou terminatrices des marqueurs d'auxotrophies (ces gènes sont délétés de l'ATG au stop).

	Génotypes	Utilisations
RS453	<i>ade2-1 his3-11,15 ura3-52 leu2-3,112 trp1-1</i>	Contrôle tests en gouttes
<i>gus1Δ shuffle</i>	<i>ade2-1 his3-11,15 ura3-52 leu2-3,112 trp1-1, gus1::HIS3, pRS316 GUS1</i>	Chasses de plasmide
<i>dps1Δ shuffle</i>	<i>ura3-52 lys2-801am trp1- 63 his3- 200 leu2- 1 ade2- 450 ade3- 1483, pAL3-DPS1 (URA3, ADE3)</i>	Chasses de plasmide
BY4742	<i>his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 trp1Δ0</i>	Contrôle tests en gouttes et intégration du fragment β1-10
<i>gus1Δ</i> diploïde	<i>his3-11,15/ his3-11,15 ura3-52/ura3-52, leu2-3,112/leu2-3,112, GUS1/gus1</i>	Dissections
NB40-3C	<i>lys2 leu2-3,112 ura3-52 his3ΔHinDIII arg8::hisG (ρ+ cox2-62)</i>	Transformation biolistique
MR6	<i>ade2-1 his3-11,15 trp1-1 leu2-3,112 ura3-1 CAN1 arg8::HIS3 (ρ- atp6::ARG8m)</i>	Transformation biolistique

4. Milieux de culture levure

La croissance des souches de levure est réalisée dans les différents milieux présentés dans le tableau ci-dessous. Les milieux liquides et solides (additionnés d'Agar-Agar à raison de 15 g/L) sont autoclavés à 120 °C sous une pression de 1,2 bars pendant 20 min. Ils sont

finaleme nt conservés en chambre froide. Avant utilisation, ceux-ci sont séchés sous hotte à flux laminaire.

	Composition	Utilisations	
SC	6,7 g/L <i>Yeast Nitrogen base</i> (YNB) supplémenté en sulfate d'ammonium, 20 g/L glucose, 0,79 g/L <i>Complete Supplement Mixture</i> (CSM) aa	Tests en gouttes	Milieux synthétiques
SCGly	6,7 g/L YNB supplémenté en sulfate d'ammonium, 20 g/L glycérol, CSM aa	Tests en gouttes	
SC-aa/ba	6,7 g/L YNB supplémenté en sulfate d'ammonium, 20 g/L glycérol, X g/L CSM dépourvu de l'aa ou ba voulu (X varie en fonction du CSM)	Transformations et tests en gouttes	
5FOA	6,7 g/L YNB supplémenté en sulfate d'ammonium, 20 g/L glucose, CSM aa, 2,5mM de 5FOA	Chasses de plasmide et tests en gouttes	
5FOAGly	6,7 g/L YNB supplémenté en sulfate d'ammonium, 20 g/L glycérol, CSM aa, 2,5mM de 5FOA	Tests en gouttes	
Sporulation	1 g/L d'extrait de levure, 0.5 g/L glucose, 10 g/L d'acétate de potassium	Sporulation des diploïdes	
YPD	20 g/L peptone, 10 g/L d'extrait de levure, 20 g/L glucose	Chasses du plasmide et tests en gouttes	Milieux riches
YPGly	20 g/L peptone, 10 g/L d'extrait de levure, 20 g/L glycérol	Tests en gouttes	

II. Techniques biochimiques utilisées

1. Préparations de bactéries chimio-compétentes

Les souches DH5 α , XL1-blue ou DB3.1 sont inoculées dans du milieu LB et incubées à 18 °C jusqu'à l'obtention d'une densité optique à 600 nm (DO_{600nm}) d'environ 0,6. Après 10 min d'incubation sur glace, les bactéries sont culottées par 10 min de centrifugation à 2500 g à 4 °C. Le culot est alors re-suspendu dans 80 mL de tampon TBjap pH 6,7 froid (10 mM Pipes, 15 mM CaCl₂, 250 mM KCl, 55 mM MnCl₂), additionné de 2 % de DMSO, puis incubé 10 min

dans la glace. Les bactéries sont une nouvelle fois culottées par 10 min de centrifugation à 2500 *g* à 4 °C et re-suspendues dans 20 mL de TBjap contenant 7 % de DMSO. Après une dernière incubation de 10 min dans la glace, les cellules sont finalement aliquotées dans des tubes Eppendorf et plongées dans de l'azote liquide avant leur conservation à -80 °C.

2. Transformations de bactéries

Après décongélation lente des cellules compétentes dans la glace, celles-ci sont incubées 30 min à 4 °C en présence de 5 µL de produit de ligation ou 2 µL de réaction Gateway. Les cellules sont alors incubées 45 sec à 42 °C afin de favoriser l'entrée du plasmide dans la cellule. Finalement, après 2 min sur glace, les cellules sont, soit directement étalées sur milieux LB + Amp pour les ligations classiques et les réactions LR, soit incubées 10 à 15 min à 37 °C (phase de régénération), suivi d'un étalement sur LB + Kan pour les réactions BP. Les boîtes sont alors incubées durant la nuit à 37 °C.

3. Préparation de plasmides

Les colonies apparues sur boîtes sont inoculées dans 2 à 3 mL de milieu LB supplémenté d'Amp ou de Kan et incubées sous agitation durant la nuit à 37 °C. Les bactéries sont alors culotées par centrifugation à 2500 *g* durant 5 min et l'ADN plasmidique est extrait avec le kit *Bio basics Canada inc* suivant les instructions fournisseurs. Après élution de l'ADN sur colonne, la concentration est finalement mesurée au Nanodrop (Labtech France). Pour les vecteurs utilisés pour la clonage Gateway®, la concentration finale est ajustée à 150 ng/µL.

4. Restriction enzymatique

La présence des gènes d'intérêts est vérifiée par restriction enzymatique. Pour cela, 500 ng des préparations de plasmides sont mélangés avec 2 µL de tampon de digestion *Fast digest* (ThermoFisher), 1 µL de chaque enzyme utilisée et de l'eau (QSP 20 µL). Le milieu réactionnel est incubé entre 5 et 30 min à 37 °C et les échantillons sont ensuite déposés sur gel d'agarose 1 % (p/v). Après migration à 120 V durant 20 min, la présence des fragments de bonnes tailles est vérifiée par marquage au BET et visualisation sous UV à la Gbox (Syngene).

5. Préparation d'ADN génomique de levure

L'ADN génomique d'une souche de levure est extrait à partir de 20 mL d'une culture à une DO_{600nm} d'environ 1. Les cellules sont tout d'abord culottées par centrifugation à 3500 *g* pendant 10 min. Le culot est alors lavé avec de l'eau stérile, centrifugé dans les mêmes conditions que précédemment et re-suspendu dans 400 μ L de tampon de lyse (2 % Triton X-100, 1 % SDS, 10 mM Tris-HCl pH 8 et 1 mM d'EDTA). Le mélange est additionné de 400 μ L de phénol/chloroforme et de 0,6 g de billes de verre de diamètre 0,45 mm, et le tout est vortexé pendant 3 min. Après centrifugation à 10000 *g* pendant 2 min, le surnageant est récupéré, additionné de 1,2 mL d'éthanol 100 % et de 40 μ L d'acétate de sodium (AcNa) 3M et incubé à -20 °C pendant 10 min. Le tube est alors centrifugé à 10000 *g* pendant 10 min et le surnageant est éliminé. Le culot est re-suspendu dans 400 μ L de tampon TE et 4 μ L de RNase (10 mg/mL) sont ajoutés. Le tout est incubé à 37 °C pendant 10 min, puis la réaction est stoppée par une extraction au phénol/chloroforme (1:1) et suivie d'une extraction au chloroforme. L'ADN est à nouveau précipité à -20 °C pendant 10 min en présence de 40 μ L d'AcNa 3 M et de 1 mL d'éthanol 100 %. Après une dernière étape de centrifugation à 10000 *g* pendant 10 min, le culot est séché et re-suspendu dans 50 à 100 μ L de tampon TE.

III. Amplification et clonage des gènes d'intérêts

1. Design des amorces

Lors de l'amplification des gènes d'intérêts, les amorces sens contiennent invariablement les séquences 5' (début du gène), tandis que les amorces antisens contiennent les séquences 3' (fin du gène sans le codon stop). La constitution des amorces varie ensuite en fonction de la technique de clonage utilisée. En effet, le clonage classique nécessite la présence de sites de restriction spécifiques à l'extrémité de chaque amorce, permettant un sous-clonage orienté dans le vecteur de destination. Le clonage par la technique Gateway® nécessite, quant à lui, des séquences de recombinaisons spécifiques, AttB1 (GGGG **ACAAGTTTGTACAAAAAAGCAGGCTTC**) et AttB2 (GGGG**ACCACTTTGTACAAGAAAGCTGGGTC**), présentes en amont et en aval de la séquence du gène au niveau de l'amorce sens et antisens respectivement. Finalement, les amorces utilisées pour le clonage par la technique de Gibson sont construites avec les régions plasmidiques bordant le site d'insertion du gène d'intérêt.

Les amorces sont commandées directement chez "Eurofins genomics" à une concentration de 100 µM et purifiées par dessalage.

2. Amplification PCR et purification

Les gènes d'intérêts sont amplifiés directement à partir d'ADN plasmidique, lorsque le gène est disponible au laboratoire, ou à partir d'ADN génomique de levure. L'amplification des gènes sur plasmide est réalisée dans un volume final de 50 µL contenant 25 µL du mix *PrimeSTAR Max*[®] 2X (clontech), 0,2 à 0,3 µM de chaque amorce (sens et antisens) et entre 10 pg et 200 ng d'ADN plasmidique. L'amplification débute par 30 sec de dénaturation à 98 °C. La PCR est alors séparée en trois étapes qui se répètent 30-35 fois : (i) une étape de dénaturation (5 sec à 98 °C), (ii) une étape d'hybridation des amorces (5 à 15 sec à 55 °C) et (iii) une étape d'élongation (5 sec/kb à 72 °C). L'amplification se termine alors par une étape finale d'élongation (1 min à 72 °C).

L'amplification des gènes à partir d'ADN génomique de levure est, quant à elle, effectuée dans un volume final de 50 µL contenant 10 µL de tampon Phusion "*High Fidelity*" 5X, 200 µM de dNTP, 0,5 µM de chaque amorce, 1 U d'enzyme Phusion ADN polymérase et entre 50 et 250 ng d'ADN génomique. Suite à une dénaturation initiale de 30 sec à 98 °C, 30 à 35 cycles d'amplification sont réalisés (5 à 10 sec à 98 °C ; 10 à 30 sec à 55 °C ; 15 à 30 sec/kb à 72 °C). L'amplification se termine par une étape d'élongation de 5 à 10 min à 72 °C.

Finalement, les produits PCR sont déposés sur gel d'agarose 1 % et la présence des bandes d'intérêts est visualisée comme décrit précédemment. Les produits de PCR sont alors purifiés, soit après migration sur gel et découpe de la bande d'intérêt, grâce au kit *Gel extraction* (ThermoFisher), soit directement après à partir du milieu réactionnel par le kit *PCR purification* (ThermoFisher).

3. Amplification PCR par le protocole *Phire Plant direct PCR*

La vérification de la présence des gènes d'intérêts a été réalisée par le kit *Phire Plant direct PCR* (ThermoFisher). Pour cela, une faible quantité de cellules est homogénéisée avec 10 µL de tampon Phire 2X, 0,5 µM de chaque amorce, 0,4 µL d'ADN polymérase Phire *hot start* et de l'eau (QSP 20 µL). Après une dénaturation initiale de 5 min à 98 °C, l'amplification se

décompose en 40 cycles comprenant : (i) une étape de dénaturation à 98 °C pendant 5 sec suivie (ii) d'une étape d'hybridation des amorces de 5 sec à 55 °C et se terminant par (iii) une étape d'élongation à 72 °C durant 20 sec/kb. Une étape finale d'élongation d'1 min à 72 °C est alors réalisée. La présence des bandes d'intérêt est visualisée par migration sur gel d'agarose 1 % et marquage au BET. La photo est prise à la Gbox.

4. Clonage des produits PCR

Durant mon doctorat, j'ai été amené à utiliser principalement trois techniques de clonage (Classique, Gateway® et Gibson®) pour obtenir les différents plasmides d'expression bactérien ou de levure. Préalablement aux clonages classiques et Gateway®, les produits PCR sont sous-clonés dans le vecteur pJET1.2/blunt (ThermoFisher). Pour cela, 3 à 4 µL de produits PCR purifiés sont mélangés avec 50 ng de vecteur pJET1.2 (linéarisé), 10 µL de tampon de ligation 2X, 1 U de T4 ADN ligase dans un volume final de 20 µL. Suite à une incubation de 5 à 30 min à 20-25 °C, 5 µL du milieu réactionnel de ligation sont transformés dans des bactéries XL1-Blue chimio-compétentes selon le protocole décrit précédemment. La présence du gène d'intérêt et l'absence de mutations sont finalement vérifiées par restriction enzymatique et par séquençage (GATC Biotech), respectivement.

5. Sous-clonage classique

Lors du clonage classique, les vecteurs pJET1.2 contenant les gènes d'intérêts sont préalablement digérés avec les enzymes de restriction reconnaissant les sites bordant le gène d'intérêt. En parallèle, le vecteur de destination est, lui aussi, digéré avec les mêmes enzymes. Si les sites de restriction sont (i) les mêmes de part et d'autre du gène ou (ii) compatibles, le vecteur est déphosphorylé lors de la restriction par ajout de 1 µL de FastAP (Thermofisher). Après migration sur gel et découpe des bandes d'intérêts, les deux fragments à assembler sont purifiés par le kit *Gel extraction* (ThermoFisher). Leurs concentrations sont alors déterminées au Nanodrop et un ratio insert:vecteur de 1:5 est utilisé pour la ligation. La ligation débute par le mélange d'eau (QSP 20 µL), des quantités adéquates d'insert et vecteur et de 1 µL de PEG₄₀₀₀. Après 5 min d'incubation à 65 °C, le mélange est incubé 5 min à température ambiante. Finalement, le milieu réactionnel est complété avec 2 µL de tampon T4 ADN ligase 10X et d'1 U de T4 ADN ligase. Après une heure incubation à 20 à 25 °C, 5 µL de produit de

ligation sont transformés dans une souche XL1-Blue et le clonage des gènes d'intérêts dans les vecteurs d'expression est confirmé comme décrit précédemment.

6. Sous-clonage par Gateway (kit invitrogen)

Le clonage par Gateway® est basé sur deux étapes de recombinaison sites-spécifiques successives permettant (i) de transférer le gène d'intérêt dans le vecteur donneur pDONR221 (via la réaction BP) puis (ii) de le sous-cloner dans différents vecteurs d'expression pAGXXX-GPD-CcbB-YYY (via la réaction LR). Dans un premier temps, 1 µg du vecteur contenant le gène d'intérêt (pJET1.2-gène X) est linéarisé. Après inactivation de l'enzyme suivant les instructions fournisseurs, 5 µL de la digestion sont mélangés avec 150 ng de vecteur donneur pDONR221, 2 µL tampon BP 5X, 1,5 µL clonase BP et une quantité donnée d'eau (QSP 10 µL), et incubés durant la nuit à 25 °C. La transformation de 2 µL de la réaction dans la souche DH5α est finalement réalisée et les bactéries sont étalées sur milieu solide LB + Kan, après 10 min de régénération à 37 °C dans du milieu LB. La présence du gène d'intérêt est alors vérifiée par restriction enzymatique et séquençage.

Après recombinaison des gènes d'intérêts dans le vecteur donneur, ceux-ci sont transférés dans les vecteurs d'expression par la réaction LR. Pour cela, 150 ng de pDONR221-Gène X sont mélangés avec 150 ng de vecteur de destination (pAGXXX-GPD-CcbB-YYY), 2 µL tampon LR 5X et 1,5 µL de clonase LR dans un volume final de 10 µL. Après 1 à 2 h d'incubation à 25 °C, 2 µL de réaction sont transformés dans des bactéries DH5α chimio-compétentes et les bactéries sont finalement étalées sur LB + Amp. Comme lors de la réaction BP, les clones sont ensuite testés par restriction enzymatique et séquençage.

7. Sous-clonage par la technique de Gibson

La dernière technique de clonage que j'ai utilisée est l'assemblage par Gibson. Après amplification du gène d'intérêt par PCR, le produit final est constitué de la région codant pour la protéine d'intérêt bordée par les séquences présentes sur le vecteur de destination. Le milieu réactionnel est alors additionné de tampon de digestion *Fast Digest* (concentration finale 1X), ainsi que d'1 U de l'enzyme DpnI et incubé durant 1 à 2 heures à 37 °C. En parallèle, le vecteur de destination est, soit amplifié avec des amorces contenant les régions bordant le produit PCR à incorporer, soit linéarisé par une enzyme de restriction, dont la séquence

reconnue est présente au niveau de ces mêmes régions. Le produit PCR et le vecteur linéaire sont alors purifiés par le kit *PCR Clean-up* (Macherey-Nagel) et élués dans le même volume final (généralement 30 μL). Le tampon IsoT est préparé par mélange de 3 mL de Tris-HCl 1 M pH 7,5, 150 μL de MgCl_2 2M, 600 μL de dNTP (100 mM chacun), 300 μL de DTT 1M, 20 mg de *nicotinamide adénine dinucléotide* (NAD) 100 mM, 1,5 g de PEG₈₀₀₀ et 2,25 mL d' H_2O . Des aliquotes de 320 μL sont alors réalisées et conservées à -80°C . Une aliquote de 320 μL tampon IsoT est alors mélangée avec 1,2 μL de T5 exonucléase (10 U/ μL), 20 μL de *Phusion DNA polymerase* (2 U/ μL), 160 μL de Taq ADN ligase (2U/ μL) et 300 μL d' H_2O pour former le mix IsoT. Le tout est alors aliquoté par tube de 11 μL et conservé à -80°C . Pour réaliser le clonage, 1 μL de chaque fragment purifié est alors incubé entre 30 et 45 min à 50°C avec 11 μL de mix IsoT dans un volume final de 20 μL . Finalement, 5 μL de réaction sont transformés dans des bactéries chimio-compétentes DH5 α et les bactéries sont étalées sur milieu de sélection spécifiques (LB + Amp ou + Kan) selon le marqueur de sélection présent dans le vecteur de destination. L'insertion du fragment d'intérêt dans le vecteur de destination est finalement vérifiée par restriction enzymatique et séquençage.

IV. Utilisation de l'outil levure

1. Transformation de levures

La souche à transformer est mise en culture dans du milieu riche (YPD) ou milieu de sélection si elle contient déjà un plasmide (SC-aa/ba). Le lendemain, la pré-culture est diluée et remise à 30°C jusqu'à l'obtention d'une $\text{DO}_{600\text{nm}}$ d'environ 1 (phase exponentielle). Les levures sont alors culottées à 3500 g pendant 10 min et le surnageant est éliminé. Après lavage avec de l'eau stérile et centrifugation à 3500 g pendant 10 min, le culot est re-suspendu dans un volume adéquat à raison de 50 μL par transformation. Les 50 μL de suspension de levures sont alors mélangés vigoureusement avec 10 μL d'ADN de sperme de saumon (100 $\mu\text{g}/\text{mL}$ soniqué) préalablement chauffé. Finalement, 2 μL d'ADN plasmidique (quelle que soit la concentration), 240 μL de PEG₄₀₀₀ à 50% (p/v), 16 μL d'acétate de lithium 1M et 50 μL d'eau stérile sont ajoutés au mélange précédent. Le tout est légèrement vortexé et incubé à 42°C pendant 45 min. Les levures sont ensuite lavées deux fois avec de l'eau stérile et sont finalement étalées sur le milieu de sélection, dépendant du marqueur d'auxotrophie présent

sur le vecteur utilisé. Après 3 jours d'incubation à 30 °C, des colonies isolées sont réétalées sur une autre boîte du même milieu.

2. Chasse du plasmide au 5FOA

Lors de l'étude de la complémentation, les souches gèneXΔ *Shuffle* contenant le plasmide d'expression du gène d'intérêt sont mises en culture durant la nuit sur milieu riche (YPD), permettant de favoriser la perte de plasmides. Celle-ci sont alors étalées sur milieu de sélection 5FOA permettant la croissance des souches capables de perdre le plasmide *URA3*. Après 3 à 4 jours de croissance, les colonies sont repiquées sur milieu 5FOA et la perte du plasmide est vérifiée par étalement sur milieu SC-Ura. La capacité de croissance est finalement vérifiée par tests en gouttes.

3. Sporulation et dissections de tétrades

Les souches diploïdes, contenant les plasmides d'expression des gènes d'intérêts, sont étalées sur milieu de sporulation et incubés 5 jours à 25 °C. La présence des tétrades est alors vérifiée par microscopie, et lorsque la proportion est suffisante (10%), les tétrades sont disséquées. Pour cela, 45 µL d'eau stérile sont homogénéisés avec 5 µL de Zymolyase (concentration de 0,25 à 2,5 mg/mL) et une petite quantité de cellules ayant sporulées. Le mélange est incubé 5 min à température ambiante et 10 µL du lysat sont finalement déposés sur une boîte de milieu riche. Les spores sont séparées par un micromanipulateur au microscope à dissection et cultivées pendant 2 jours à 30 °C. La ségrégation est finalement vérifiée par étalement sur milieu respiratoire (YPGly) et répliques velours sur différents milieux de sélection (SC-aa/ba) permettant l'établissement du génotype des souches. La capacité de croissance des souches peut également être testée par tests en gouttes.

4. Tests en gouttes

Afin de tester la capacité des souches à se développer sur différents milieux, des pré-cultures de l'ensemble de celles-ci sont inoculées dans des milieux adéquates : riche pour les souches sauvages et de sélection permettant le maintien des plasmides d'expression pour les autres. Le lendemain matin, les souches sont diluées et mises en culture jusqu'à l'obtention d'une DO_{600nm} d'environ 1. Lorsque que les souches se sont suffisamment développées, elles

sont diluées à une DO_{600nm} de 0,5, puis trois fois en cascade à raison de dix. Finalement, 7 μ L de chaque dilution sont déposés en ligne sur les milieux sélectionnés. Les boîtes sont prises en photo après deux et quatre jours de croissance à 30 °C.

V. Techniques de biologies cellulaires utilisées

1. Purifications de mitochondries

Afin de purifier des mitochondries, le protocole utilisé est dérivé de (Meisinger et al., 2006), décrit dans la revue que j'ai co-écrite et disponible dans l'Annexe 1.

2. Fractionnements subcellulaires

Durant mon doctorat, j'ai été amené à séparer les fractions membranaires de levure. Pour cela, j'ai utilisé le protocole décrit dans la revue présente dans l'Annexe 1.

3. Visualisations microscopiques

a. Cultures de levures


Le jour avant l'observation microscopique, les levures sont mises en culture dans 2 mL de milieu de sélection permettant le maintien du plasmide. Le lendemain matin, les cellules sont diluées dans 2 mL de ce même milieu puis remises en culture jusqu'à l'obtention d'une DO_{600nm} d'environ 1. Elles sont alors culottées par centrifugation à 8000 g pendant 2 min et lavées deux fois avec de l'eau stérile ou du PBS. Elles sont finalement remises en suspension dans un volume plus faible (50 à 100 μ L) et 2 μ L sont déposés entre lame et lamelle pour l'observation microscopique.

b. Etudes microscopiques et paramètres utilisés

L'observation des échantillons est réalisée au microscope à fluorescence Axio Observer D1 (Zeiss) à l'aide d'un objectif à immersion à huile (100X/1,45). Les images sont réalisées par l'intermédiaire de la caméra CoolSnap HQ2 photometrix (Roper Scientific), associée au microscope et reliée à l'ordinateur. Le traitement d'image et les montages sont réalisés à l'aide du logiciel ImageJ (Rasband W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA). Les réglages sont les suivants :

Channel Definition

Dye: CFP

Color: 

Name: CFP

Exposure

Time: 1000 ms

Mode: Auto Fixed Camera

Hardware Settings

During acquisition ...

Workgroup: CFP


After acquisition ...

Workgroup: all off

Focus: Current Autofocus

Channel Definition

Dye: GFP

Color: 

Name: GFP

Exposure

Time: 800 ms

Mode: Auto Fixed Camera

Hardware Settings

During acquisition ...


Workgroup: GFP

After acquisition ...

Workgroup: all off

Focus: Current Autofocus

CFP



1000

Channel settings

100

Use camera setting

Workgroup: CFP

Workgroup: all off

Snap

CFP

Black/white value autom. to min/max

1

0

0

0

100

100

2


0

0

<none>

ZStack

GFP



800

Channel settings

100

Use camera setting

Workgroup: GFP

Workgroup: all off

GFP

Black/white value autom. to min/max

1

0

0

0

100

100

0

0

0

<none>

ZStack

c. Sondes fluorescentes utilisées

Colorants	Cibles	λ excitation max (nm)	λ émission max (nm)
Hoechst	Molécules d'ADN (nucléaire et mitochondriale)	350	461
Mitotracker deep Red	Mitochondries	644	665
FM4-64	Membrane plasmique (1-2 min)	515	640
	Endosomes (5 min)		
	Membrane vacuolaire (15 min)		

d. Marquage des compartiments

Lors de la coloration des noyaux, les cellules à une DO_{600nm} comprise entre 0,5 et 1 sont incubées 15-20 min à 30 °C en présence de 1 μ L de Hoechst (10 mg/mL). Suite au marquage, les cellules sont culottées par centrifugation à 3500 g pendant 5 min et lavées deux à trois fois avec de l'eau stérile ou du PBS. Le culot est repris dans 50 μ L d'eau ou de PBS et 2 μ L de la suspension sont déposés entre lame et lamelle. Les échantillons sont alors observés par microscopie à fluorescence.

Les mitochondries sont, quant à elles, marquées par du Mitotracker *Deep Red*. Les cellules à une DO_{600nm} inférieure à 1 sont incubées 15 min à 30 °C en présence de 1 μ L de Mitotracker (200 nM). Elles sont ensuite culottées à 3500 g pendant 5 min, puis lavées trois fois avec de l'eau stérile. Le culot est finalement re-suspendu dans 50 à 100 μ L d'eau stérile et 2 μ L de la suspension sont disposés entre lame et lamelle. Les cellules sont finalement visualisées par microscopie à fluorescence. Si le bruit de fond est trop important, des lavages supplémentaires peuvent être effectués. L'utilisation du Mitotracker *Deep red* n'est pas recommandée pour les souches *ade2* (présence d'un pigment rouge vacuolaire).

Finalement, le compartiment sondé par le marquage au FM4-64 est différentiel et le temps d'incubation permet la visualisation de la membrane plasmique (1-2 min), des

endosomes (5 min) ou de la membrane vacuolaire (15 min). Les cellules (1-2 unités de DO_{600nm}) sont culottées par centrifugation à 3500 *g* pendant 5 min et re-suspendues dans 50 μ L de milieu riche (YPD). Le marquage est alors effectué à 30 °C pendant 1 à 15 min selon le compartiment ciblé, à l'aide de 2 μ L de FM4-64 (8 μ M). Les cellules sont ensuite lavées avec 500 μ L de YPD, remises en suspension dans 300 μ L de ce même milieu et incubées 10 min à 30 °C. Les cellules sont finalement culottées, lavées avec du YPD et déposées entre lame et lamelle au vu de leur observation microscopique.

VI. Obtention et études des extraits protéiques

1. Préparations d'extraits protéiques totaux

Pour l'obtention d'extraits totaux, les souches sont mises en culture dans le milieu permettant le maintien du plasmide d'expression si nécessaire, ou dans un milieu riche. Le lendemain, les souches sont diluées et mises en culture jusqu'à atteindre une DO_{600nm} d'environ 1. L'équivalent d'une unité de DO est alors prélevé et centrifugé à 3500 *g* pendant 5 min. Le culot est repris dans 500 μ L d'eau stérile et incubé 10 min dans la glace en présence de 50 μ L de NaOH 1,85 M. Les protéines sont précipitées par ajout de 50 μ L de TCA 100 % et une incubation 10 min minimum sur glace. Les protéines sont alors culottées par centrifugation de 10 min à 12000 *g* et 4 °C. Le culot est finalement repris dans 50 μ L de tampon de charge contenant 125 mM de Tris-HCl 1M pH 6,8, 2 % de SDS (p/v), 10 % de glycérol (v/v), 2 % de β -mercaptoéthanol (v/v) et 0.0125 % de bleu de bromophénol. Si le pH de la solution est trop acide, il est neutralisé par ajout de quelques μ L de Tris Base 1 M.

2. Immuno-détections de protéines

Les échantillons, mélangés au bleu de charge, sont préalablement dénaturés à 95 °C pendant 10 min. Pour les extraits totaux, 10 μ L sont déposées sur gel de polyacrylamide dénaturant 10 ou 12 % selon la taille des protéines ciblées. Après migration à 200 V pendant 30 à 45 min dans du tampon TG-SDS 1X, les protéines séparées sont alors transférées sur une membrane PVDF à l'aide du Trans-Blot Turbo (Bio-Rad) et du programme mixed MW : 1,3 A jusqu'à 25 A, pendant 7 min. La membrane contenant les protéines est alors activée par un bain d'éthanol, suivi d'un blocage avec du TBS-Tween20 0,3 % additionné de 5 % de lait (p/v) durant 30 min. Les anticorps primaire et secondaire sont dilués dans du tampon de blocage

selon les informations fournisseurs. Les protéines sont marquées avec l'anticorps primaire sous agitation à 4 °C durant la nuit. Après trois lavages de 10 min avec du TBS-Tween20 0,3 %, l'anticorps secondaire est ajouté et la membrane est incubée 2 à 3 heures sous agitation. Si l'incubation est plus longue, celle-ci est réalisée à 4 °C. Après deux lavages de 10 min au TBS-Tween20 0,3 %, la membrane est lavée une dernière fois 10 min avec du TBS. La présence des protéines d'intérêts est finalement visualisée suite à une incubation de 5 min dans la solution de révélation kit *Bio-Rad ECL clarity* et révélation au Chemidoc touch (Bio-Rad). L'image est finalement retravaillée avec le logiciel fournisseur Imagelab.

Résultats et discussions

Partie I : Dynamique du complexe AME
dans la régulation de la transition
diauxique chez la levure *S. cerevisiae*

1. Contexte de la recherche et participation

Mon laboratoire d'accueil s'intéresse au MSC de levure, le complexe AME. Lors de sa thèse, le Dr. Mathieu Fréchin a débuté l'étude de ce MSC et a pu montrer son importance lors du passage de fermentation à respiration (transition diauxique). Lors de ce changement métabolique, la levure remanie son protéome en exprimant ou surexprimant certains gènes indispensables au métabolisme respiratoire. De plus, la stœchiométrie du complexe AME change et la quantité d'Arc1 diminue fortement, tandis que celle des deux aaRS reste inchangée. Suite à la diminution de l'ancre cytosolique du MSC, les deux aaRS sont relâchées et peuvent se relocaliser de manière synchrone du cytosol vers de nouveaux compartiments subcellulaires. Ainsi, la MetRS se relocalise dans le noyau, alors que la GluRS est importée dans la mitochondrie afin de participer à la formation du Gln-mARNt^{Gln} par la voie de transamidation faisant intervenir l'AdT GatFAB (Araiso et al., 2014; Frechin et al., 2009b). Cet aa-ARNt va ensuite être pris en charge par le facteur d'élongation mitochondrial EF-Tu, pour permettre l'incorporation de Gln lors de la synthèse protéique. De manière intéressante, la GluRS cytosolique est la seule enzyme capable de former le Glu-mARNt^{Gln}, puisque la GluRS mitochondriale n'est pas capable de mésacyler l'mARNt^{Gln} (Rinehart et al., 2005) et que la délétion du signal d'import putatif de l'aaRS cytosolique induit une déficience respiratoire (Frechin et al., 2009b). La relocalisation de la GluRS cytosolique au niveau de la mitochondrie est donc cruciale pour la croissance de la levure en respiration. Durant ma thèse, j'ai participé à cette étude, qui à mon arrivée au laboratoire, se concentrait essentiellement sur deux questions : (i) comment se fait la régulation de la quantité d'Arc1 et (ii) la MetRS se relocalise-t-elle dans un autre compartiment et si oui, quelle y est sa fonction.

La diminution de la quantité d'Arc1 peut dépendre de deux types de régulation, transcriptionnelle ou traductionnelle. Dans un premier temps, nous avons pu montrer qu'Arc1 est régulée transcriptionnellement. En effet, lorsque l'on suit l'expression du gène *ARC1* lors de la transition diauxique, on observe une forte diminution de la quantité de l'ARNm. Chez la levure, les facteurs de transcription impliqués dans l'adaptation au changement de sources carbonés sont connus et les souches délétées pour un grand nombre des gènes codant ces protéines sont capables de croître en respiration. Pour déterminer quels sont les facteurs impliqués dans la modulation de la transcription d'*ARC1*, la quantité de l'ARNm a été testée dans ces délétants. Nous avons ainsi pu montrer que les facteurs de transcription Mig1/Hxk2

régulent l'expression d'*ARC1*. En respiration, l'activité transcriptionnelle de Mig1/Hxk2 est réprimée *via* phosphorylation par la kinase Snf1, ce qui induit notamment une forte diminution de la transcription du gène *ARC1*. La quantité d'Arc1 étant réduite, les deux aaRS sont libres dans le cytosol. La relocalisation mitochondriale de la GluRS ayant déjà été prouvée (Frechin et al., 2009b), le devenir de la MetRS devait encore être élucidé. Pour déterminer la possible relocalisation de la MetRS dans un autre compartiment, la version génomique du gène codant cette aaRS (*MES1*) a été modifiée pour permettre son expression en fusion avec la GFP. La version étiquetée de cette enzyme a ensuite été visualisée par microscopie confocale en fermentation et en respiration. La localisation de la MetRS observée par microscopie a été validée par immuno-détection sur des organites purifiés et confirme que la MetRS, libre dans le cytosol, se relocalise dans le noyau en respiration. Ces résultats valident également qu'Arc1 sert d'ancre cytosolique pour les aaRS se relocalisant dans d'autres compartiments cellulaires.

La relocalisation de la GluRS est essentielle pour la croissance respiratoire et nous nous demandions s'il en était de même pour la MetRS et surtout si la relocalisation simultanée des deux aaRS était importante. Pour tester cette hypothèse, quatre souches ont été construites pour permettre (i) la rétention cytosolique des deux aaRS (souche ARC), (ii) la relocalisation organellaire constitutive, uniquement de la GluRS ou de la MetRS (souche nE et nM respectivement) ou (iii) la relocalisation organellaire simultanée et constitutive des deux aaRS (souche nE/nM). Ces trois dernières souches ont été construites à partir de la première (souche ARC). Pour les souches nE et nM, une des deux aaRS est donc constitutivement retenue dans le cytosol et l'autre importée dans son compartiment de destination. Ce système a été validé par des analyses biochimiques. Pour cela, en collaboration avec Ludovic Enkler, alors doctorant au laboratoire, j'ai réalisé des purifications de mitochondries sur les souches ARC, nE, nM et nE/nM. La présence d'une plus grande quantité de GluRS organellaire dans les souches nE et nE/nM par rapport aux souches ARC et sauvage a ensuite été validée par immuno-détection. En parallèle, la relocalisation constitutive de la MetRS a été validée par purifications de noyaux sur les souches nM et nE/nM et immuno-détection de cette protéine. Afin d'étudier le phénotype issu de la désynchronisation de la relocalisation organellaire des deux aaRS, les différentes souches ont été cultivées en fermentation et en respiration. De manière intéressante, lorsque cette relocalisation est perturbée (souche nE et nM), une

déficience respiratoire est observée, suggérant que ce mécanisme est indispensable à la croissance en respiration. La croissance est davantage impactée dans la souche nM qui est incapable de se développer en respiration plus de deux générations. Ce résultat suggère que la fonction nucléaire de la MetRS est indispensable à la transition diauxique, mais qu'elle nécessite l'import concomitant de la GluRS dans la mitochondrie, sans quoi cette fonction devient néfaste pour la croissance respiratoire.

Nous avons ensuite étudié la fonction nucléaire de la MetRS et plus particulièrement l'impact de l'import nucléaire forcé de cette enzyme (souche nM) sur la transcription de gènes nucléaires. De manière intéressante, dans la souche nM, la transcription de la totalité des gènes codants les protéines du complexe III de la CR est augmentée. De plus, il y a également répression de l'expression du gène *ATP1*, codant une s-u de la tête catalytique F₁ de l'ATP synthase mitochondriale. L'absence de cette protéine induit d'ailleurs un défaut d'assemblage des autres s-u de la tête F₁. La MetRS pourrait donc jouer un rôle primordial dans la régulation de l'expression des gènes codant les protéines de la CR mitochondriale et plus particulièrement au niveau transcriptionnel, soit par une action directe, soit par activation ou recrutement d'un facteur de transcription. Néanmoins, cette hypothèse doit encore être validée par l'étude de la capacité de la MetRS à se fixer aux régions promotrices des différents gènes ciblés et la détermination des séquences promotrices pouvant être reconnues par la MetRS seule, ou en complexe avec d'autres facteurs transcriptionnels.

En parallèle, dans la souche nM, la GluRS est constitutivement ancrée dans le cytosol et ne se relocalise que très faiblement dans la mitochondrie. Cette rétention cytosolique engendre une diminution de la quantité d'Atp9, la s-u majoritaire du rotor F₀ de l'ATP synthase mitochondriale. En effet, l'assemblage du rotor F₀ de l'ATP synthase requiert dix copies d'Atp9. Sachant que le gène codant cette s-u contient le plus de codon Gln, la traduction de cette protéine est fortement consommatrice et dépendante de Gln_{-m}ARNt^{Gln}. La synthèse de cet aa-ARNt *via* la voie de transamidation étant dépendante de l'import de la GluRS cytosolique, la séquestration cytosolique de cette enzyme dans la souche nM impacte prioritairement l'accumulation d'Atp9.

Ces différentes observations, issues de la désynchronisation du complexe AME, suggèrent donc que ce MSC participe à la régulation de l'assemblage des complexes de la CR

et qu'un défaut dans la stœchiométrie des s-u impacte directement la production d'ATP par phosphorylation oxydative et surtout la capacité de la levure de se développer en respiration. De plus, la régulation ne fait pas intervenir le transport rétrograde ou antérograde, mais est directement liée aux conditions de croissance. En effet, lorsque la quantité de glucose disponible vient à diminuer, l'activation de la voie Snf1/4 permet de réduire fortement la synthèse de l'ancre cytosolique Arc1. Les deux enzymes se relocalisent alors dans leurs compartiments respectifs pour réguler, d'une part, la transcription des gènes nucléaires codant les s-u de la CR, et de l'autre, la traduction des s-u codées par le génome organellaire. Ce dialogue favorise ainsi l'assemblage stœchiométrique des différents complexes de la CR et plus particulièrement l'ATP synthase qui peut alors utiliser le gradient de protons pour produire de l'énergie.

2. Article 1: Expression of nuclear and mitochondrial genes encoding ATP synthase is synchronized by disassembly of a multisynthetase complex

Expression of Nuclear and Mitochondrial Genes Encoding ATP Synthase Is Synchronized by Disassembly of a Multisynthetase Complex

Mathieu Frechin,^{1,7,8,9,*} Ludovic Enkler,^{1,8} Emmanuel Tetaud,^{2,3,8} Daphné Laporte,¹ Bruno Senger,¹ Corinne Blancard,^{2,3} Philippe Hamann,⁴ Gaétan Bader,¹ Sandra Clauder-Münster,⁵ Lars M. Steinmetz,^{5,6} Robert Pierre Martin,¹ Jean-Paul di Rago,^{2,3,*} and Hubert Dominique Becker^{1,*}

¹Génétique Moléculaire, Génomique, Microbiologie, UMR 7156, CNRS, Université de Strasbourg, 21 Rue René Descartes, 67084 Strasbourg Cedex, France

²Centre National de la Recherche Scientifique, Institut de Biochimie et Génétique Cellulaires, UMR 5095, 1 Rue Camille Saint-Saëns, 33077 Bordeaux, France

³Institut de Biochimie et Génétique Cellulaires, UMR 5095, Université Bordeaux, 146 Rue Léo Saignat, 33076 Bordeaux, France

⁴Plateforme Protéomique de l'Esplanade, Université de Strasbourg, CNRS- FRC1589, IBMC, 15 Rue René Descartes, 67084 Strasbourg Cedex, France

⁵European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany

⁶Stanford Genome Technology Center, Stanford University, Palo Alto, CA 94304, USA and Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA

⁷UPR Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire, Université de Strasbourg, CNRS, 15 Rue René Descartes, 67084 Strasbourg Cedex, France

⁸Co-first author

⁹Present address: Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

*Correspondence: mathieu.frechin@uzh.ch (M.F.), jp.dirago@ibgc.cnrs.fr (J.-P.d.R.), h.becker@unistra.fr (H.D.B.)

<http://dx.doi.org/10.1016/j.molcel.2014.10.015>

SUMMARY

In eukaryotic cells, oxidative phosphorylation involves multisubunit complexes of mixed genetic origin. Assembling these complexes requires an organelle-independent synchronizing system for the proper expression of nuclear and mitochondrial genes. Here we show that proper expression of the F_1F_0 ATP synthase (complex V) depends on a cytosolic complex (AME) made of two aminoacyl-tRNA synthetases (cERS and cMRS) attached to an anchor protein, Arc1p. When yeast cells adapt to respiration the Snf1/4 glucose-sensing pathway inhibits *ARC1* expression triggering simultaneous release of cERS and cMRS. Free cMRS and cERS relocate to the nucleus and mitochondria, respectively, to synchronize nuclear transcription and mitochondrial translation of ATP synthase genes. Strains releasing asynchronously the two aminoacyl-tRNA synthetases display aberrant expression of nuclear and mitochondrial genes encoding subunits of complex V resulting in severe defects of the oxidative phosphorylation mechanism. This work shows that the AME complex coordinates expression of enzymes that require intergenomic control.

INTRODUCTION

Aminoacyl-tRNA synthetases (aaRSs) constitute a family of ubiquitous enzymes essentially known for their primary function in pro-

tein synthesis through tRNA aminoacylation (Söll and Schimmel, 1974). In recent years, many aaRSs have been found to perform additional and equally crucial tasks through recruitment of various functional modules during evolution (Ibba and Söll, 2000; Schimmel and Ribas De Pouplana, 2000; Brown et al., 2010). These non-canonical functions of aaRSs include cellular processes that exploit their primary aa-tRNA synthesizing function like amino acid (aa) or cell wall synthesis (Roy et al., 2003; Stewart et al., 1971). Other unconventional functions have been associated with an additional distinctive feature of aaRSs in higher eukaryotes: their capacity to assemble into large multi-aminoacyl-tRNA synthetase complexes (MSCs) (Negrutskii et al., 1994; Kaminska et al., 2009) and perform additional tasks (Guo and Schimmel, 2013; Guo et al., 2010; Park et al., 2008). The concept has emerged during the last decade that MSCs are reservoirs for releasable and multifunctional aaRSs (Ray et al., 2007; Park et al., 2008; Hausmann and Ibba, 2008). Sequestration inside MSCs confines aaRSs' activities to their primary role of tRNA aminoacylation, while they can carry out alternative regulatory roles upon release (Sampath et al., 2004; Park et al., 2008).

A cytosolic MSC, that we named AME, has been found in the yeast *Saccharomyces cerevisiae*. It is made of a cytosolic glutamyl- (cERS) and a cytosolic methionyl-tRNA synthetase (cMRS) bound to the anchor protein Arc1p, initially described as a cofactor of tRNA nuclear export (Simos et al., 1996). Recently, we discovered that upon switching from fermentation to respiration, expression of Arc1p is strongly decreased, generating free cERS molecules that are then imported into mitochondria. Inside mitochondria, cERS mischarges the mitochondrial tRNA^{Gln} (_mtRNA^Q), which is then converted into glutamyl-tRNA^{Gln} (Q-_mtRNA^Q) by the GatFAB, the mitochondrial tRNA-dependent amidotransferase (AdT) (Frechin et al., 2009, 2010).

As there is no other source of Q_mtRNA^Q in mitochondria, cERS is essential to translate proteins encoded by the mitochondrial genome (mtDNA).

In every eukaryote, the main function of the mtDNA is the production of a few proteins involved in mitochondrial ATP production through the process of oxidative phosphorylation (OXPHOS). These proteins assembled with nuclear-encoded proteins into four complexes (I, III, IV, V) embedded within the mitochondrial inner membrane. Complexes I–III, together with complex II that has a pure nuclear origin, transfer electrons to oxygen coupled to proton translocation across the inner membrane. The resulting transmembrane proton gradient is used by the ATP synthase (complex V) to synthesize ATP from ADP and inorganic phosphate. In the yeast *S. cerevisiae*, seven OXPHOS subunits are encoded by the mtDNA: the complex III cytochrome *b* subunit, three complex IV subunits (Cox1p, Cox2p, Cox3p), and three ATP synthase subunits (Atp6p, Atp8p, Atp9p) (summarized in Figure S3A available online).

Due to the dual genetic origin of the OXPHOS system, a coordinated expression of its nuclear DNA (nDNA)- and mtDNA-encoded components is of crucial importance to avoid accumulation of unassembled OXPHOS subunits or partial assembly intermediates, which may have detrimental consequences. For instance, if the F_O proton-translocating domain of ATP synthase is not properly connected to the F₁ sector of this enzyme where the ATP is synthesized, this will result in proton leaks through the inner membrane (Duvezin-Caubet et al., 2003). Although the assembly of the mitochondrial OXPHOS system has been extensively studied (Ackerman and Tzagoloff, 2005; Fox, 2012), mechanisms that ensure a synchronized expression of its nDNA- and mtDNA-encoded subunits have thus far not been described.

Here we show that the Snf1/4 glucose-sensing pathway regulates the assembly/disassembly of the AME complex in response to the respiratory needs of the cell by shutting off expression of the Arc1p anchor upon shift to respiratory metabolism. This is the first example of an MSC able to release multiple aaRSs in a synchronized fashion. We further show that the release of cMRS and cERS from Arc1p is followed by the relocation of the two aaRSs into the nucleus and mitochondria, respectively. Interestingly, strains unable to release synchronously the two aaRSs display aberrant expression of nuclear and mitochondrial genes encoding subunits of the F₁ and F_O domains of ATP synthase. These results indicate that the AME system allows the cells to coordinate the expression of genes located in the nucleus and mitochondria.

RESULTS

Diauxic Shift Triggers Arc1p Downregulation and Relocation of cERS and cMRS to the Mitochondria and the Nucleus, Respectively

S. cerevisiae ferments glucose into ethanol and carbon dioxide (first exponential phase, Figure 1A) (Otterstedt et al., 2004). Then low glucose triggers a molecular and cellular remodeling program called diauxic shift (DS) allowing cells to metabolize nonfermentable carbon sources like ethanol, glycerol, and lactate through OXPHOS (second growth phase, Figure 1A) (Zaman et al., 2008). Quantitative real-time PCR shows that

amounts of *ARC1* transcript decrease during the DS, while those of cERS (*GUS1*) and cMRS (*MES1*) remain unchanged (Figure 1B). As a result, Arc1p accumulation is decreased by 90% (Figure 1C, inset; Figures S1A–S1F). To know how *ARC1* expression is modulated, we analyzed *ARC1* mRNA levels in strains deleted for genes controlling the three major nutrient-sensing pathways, Snf1/4, TOR-, and PKA (Zaman et al., 2008; Table S1). Relative quantitative real-time PCR quantification showed that the Snf1p kinase, the nuclear-import factor Gal83p, and the Mig1p·Hxk2p transcription factor complex are required to repress *ARC1* transcription upon switch to respiration (Figures 1D and S1G; Table S2; Hedbacker and Carlson, 2008). Fitting our data into the well-established Snf1/4 network (Figure S1G; McCartney and Schmidt, 2001; Zaman et al., 2008) suggests that transcription of *ARC1* is activated by Mig1·Hxk2 (Santangelo, 2006; Schüller, 2003; Figure 1E), whereas repression results from activation of the Snf1/4 complex and nuclear exclusion of Mig1p·Hxk2p (Figure 1F).

Out of the three AME complex components, only Arc1p is subjected to a major downregulation at the DS (Figures 1B–1D and S1F; Table S2; Frechin et al., 2009). As a consequence, the proportion of free cERS and cMRS, i.e., unbound to Arc1p, increases drastically. We previously showed that free cERS largely relocates to mitochondria (Figure 2A; Frechin et al., 2009). To determine the fate of free cMRS, we compared by confocal microscopy the distribution of a cMRS-GFP fusion under its endogenous promoter in fermenting and respiring yeast cells. The fusion protein colocalized with Hoechst-stained nuclei in respiring cells, but not in fermenting cells (Figure 2B). Using anti-cMRS antibodies, nuclear accumulation of cMRS was estimated as 15-fold more important in respiring versus fermenting cells (Figure 2C). The first 55 aa of cMRS harbor a bipartite nuclear localization signal responsible for this nuclear relocation (Figure 2D). Indeed, when GFP was fused to this sequence it entirely relocated to the nucleus (Figure 2D). Interestingly, the RNA Pol II regulatory components Rpb1p and -7p coeluted with cMRS in tandem affinity purification of nuclear extracts (Figures 2E, S2A, and S2B), indicating that the purpose of the nuclear relocation of cMRS in respiring yeast cells might be to modulate the expression of nuclear genes.

Construction of Strains with a Desynchronized Assembly/Disassembly of AME

To evaluate the biological role of the synchronized disassembly of AME at the DS, we constructed the following strains: (i) ARC, in which both cMRS and cERS remain largely anchored to Arc1p regardless of the metabolic state of the cell, (ii) nE/nM, in which most cERS and cMRS are always free, (iii) nE, in which release of cERS is constitutive, while cMRS remains largely bound to Arc1p, and (iv) nM, in which most of cERS is retained in the cytosol by Arc1p, while cMRS is permanently released (Figure 2F; see Experimental Procedures). These four strains can be split in two sets of strains, one composed of the ARC and nE/nM strains, in which we can evaluate the effect of the simple release or sequestration of both aaRSs, and one composed of the nE and nM strains, in which we can assess the effect of the desynchronized release of either one or the other aaRS. They were obtained by rendering the expression of Arc1p constitutive

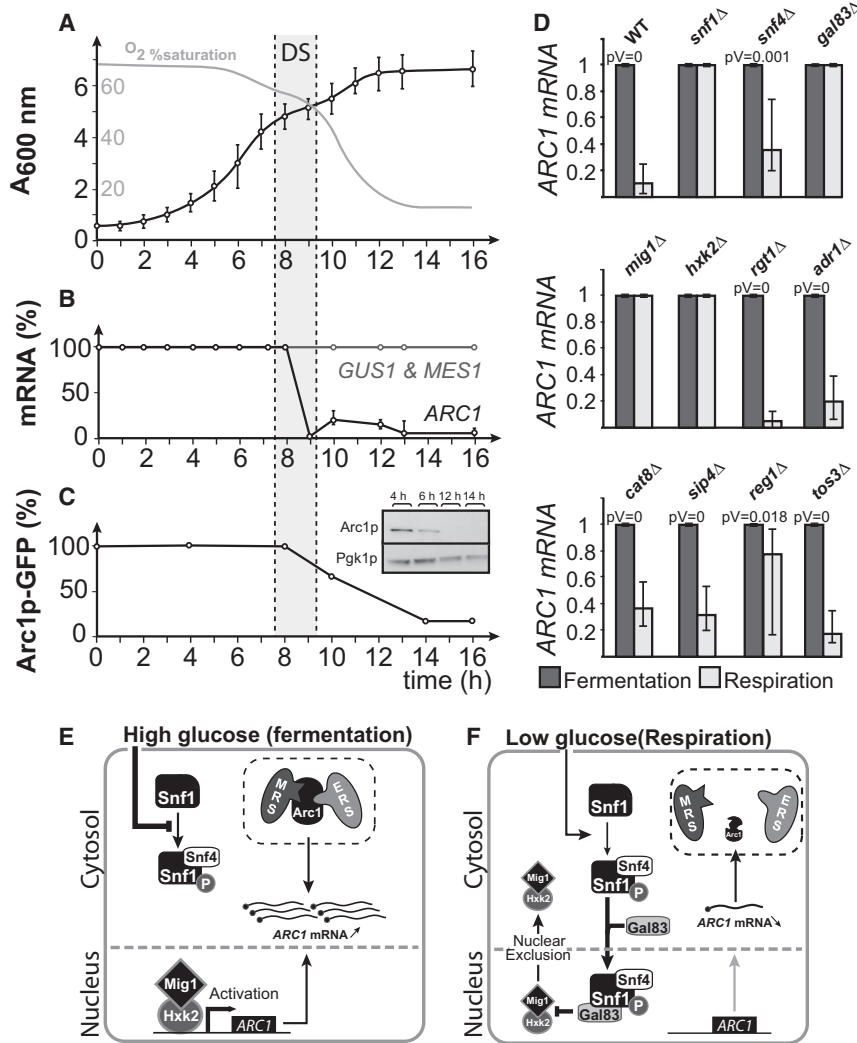


Figure 1. AME Complex Disassembly Is Induced by Snf1/4 Kinase-Mediated Transcription Inhibition of ARC1 at the Diauxic Shift

(A) Wild-type *S. cerevisiae* growth curve on SCGlc medium (n = 3). The diauxic shift (DS) marks the transition from fermentation (first exponential phase) to respiration (second exponential phase). (B) Quantitative real-time PCR quantification of *ARC1*, *GUS1* (cERS), and *MES1* (cMRS) mRNAs (n = 3). Before the DS the p values of each mRNA quantification compared to their quantifications at t = 0 are > 0.05 (thus considered as not different); after the DS, the p values are > 0.05 for *GUS1* and *MES1* and < 10⁻³ for *ARC1*. (C) FACS and immunoblot (inset) quantification of Arc1-GFP (under control of its own promoter) and Arc1p (with anti-Arc1p antibodies), respectively. For the FACS quantification only the mean Arc1-GFP signal value is shown; see Figure S2A for the full distribution of single-cell Arc1-GFP values. (D) Quantitative real-time PCR quantification of *ARC1* mRNA in fermentation (dark gray) and respiration (light gray) in strains deleted from genes of the Snf1/4 pathway components (SEM, p values, see Table S2, n = 4). (E and F) Implementation of our data in (D) into the well-documented Snf1/4 pathway (Zaman et al., 2008).

inability to consume ethanol produced during growth on galactose (Figures 3C and 3D). Upon additional overexpression of the N-terminal Arc1p binding domain of cERS in the nM strain (yielding the nE/nM strain), efficient respiratory growth was restored (Figure 3B), suggesting that resynchronizing the release of both aaRSs is vital when yeast cells switch to

(Figure S2C), and using plasmids expressing peptides that correspond to the N-terminal Arc1p-binding domain of cERS or cMRS (Simader et al., 2006) to prevent the full-length aaRSs from binding to Arc1p (see Experimental Procedures). The expected distribution of cMRS and cERS in cells of these strains was confirmed by the following: (i) size-exclusion chromatography of total cellular extracts coupled to measurements of tRNA aminoacylation activities mediated by cMRS and cERS (Figure 2G), and (ii) western blotting of nucleus- and mitochondria-enriched fractions, respectively (Figures 2H, 2I, and S2D).

Influence of Desynchronized AME Disassembly on Mitochondrial Function

The ARC, nE/nM, nE, and nM strains grew well on fermentable carbon sources (glucose and galactose; Figures 3A and 3B, left). On respiratory carbon sources (glycerol, glycerol/ethanol, lactate), growth of the nM strain was severely compromised, whereas the three others grew relatively well with respect to WT (Figures 3A and 3B, right; Figures S3B and S3C). The respiratory deficiency of the nM strain was further confirmed by its

the respiratory metabolism. Moreover, the overexpression of the N-terminal Arc1p binding domain of cMRS alone in a WT or *arc1Δ* strain did not alter growth in respiration (Figure S3D). Therefore, when both aaRSs are retained by Arc1p or permanently released from it, the respiratory growth is not significantly impacted. The same holds when cMRS is retained on Arc1p while cERS is free. The most severe impact was observed when cMRS alone was allowed to reach the nucleus while cERS cannot relocate to mitochondria (nM strain, Figure 3A).

Biochemical analysis of mitochondria was required to better evaluate the consequences on mitochondrial function of disturbing the AME complex disassembly. Mitochondria were prepared from cells grown in galactose, which is a fermentable substrate that does not elicit repression of mitochondrial function (Polakis et al., 1965, Ostergaard et al., 2000). We first measured oxygen consumption of the mitochondria using NADH as an electron donor, at state 3 (i.e., in the presence of ADP, phosphorylating conditions), state 4 (i.e., without addition of ADP, basal respiration), and in the presence of the membrane potential uncoupler CCCP (i.e., conditions at which

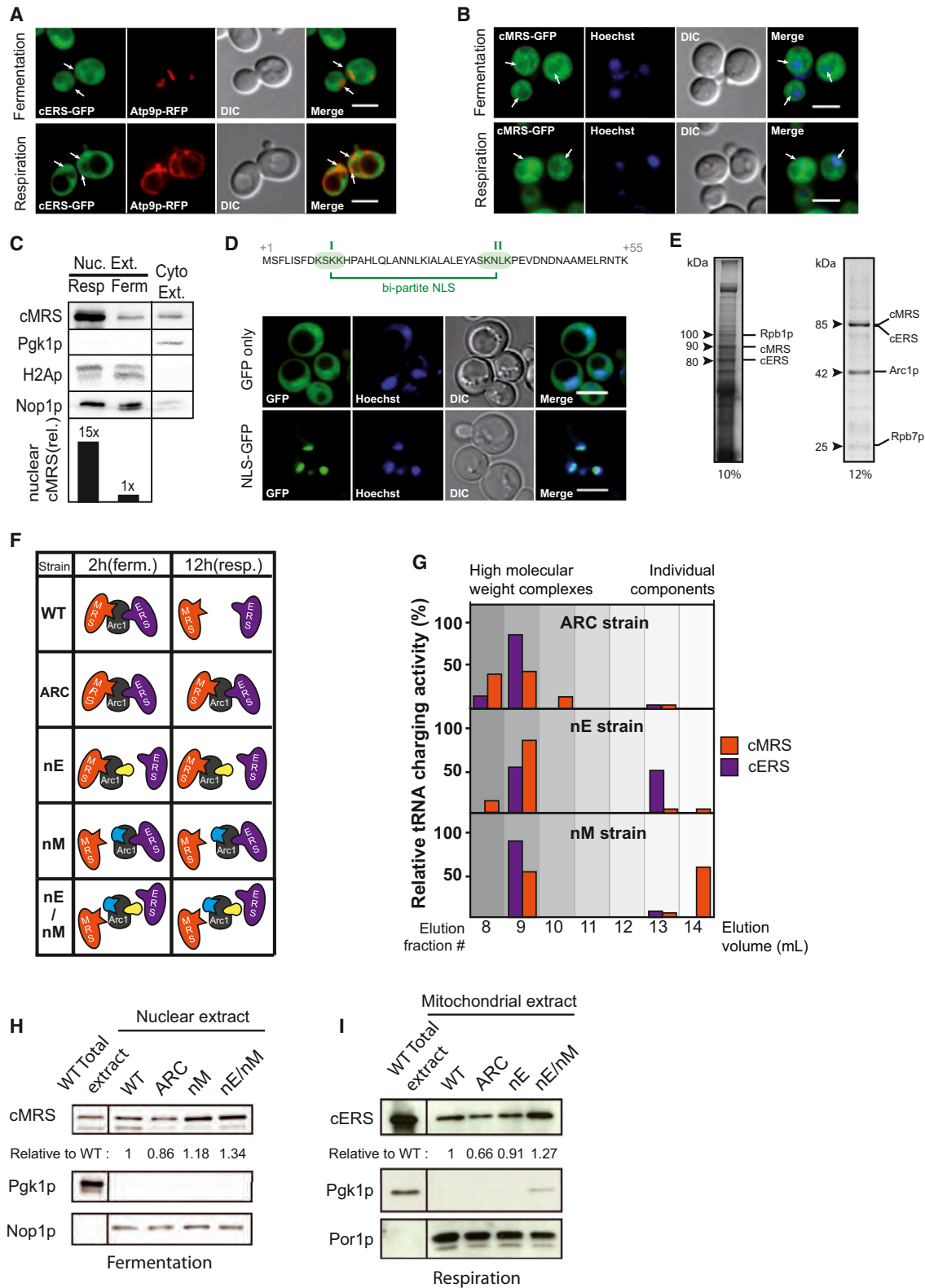


Figure 2. Reprogramming Disassembly of the AME Complex and Subcellular Relocation of cMRS and cERS by Expressing Competing Arc1p-Binding Domains

(A) Colocalization (arrowheads) of GFP-tagged cERS with Atp9p-RFP.

(legend continued on next page)

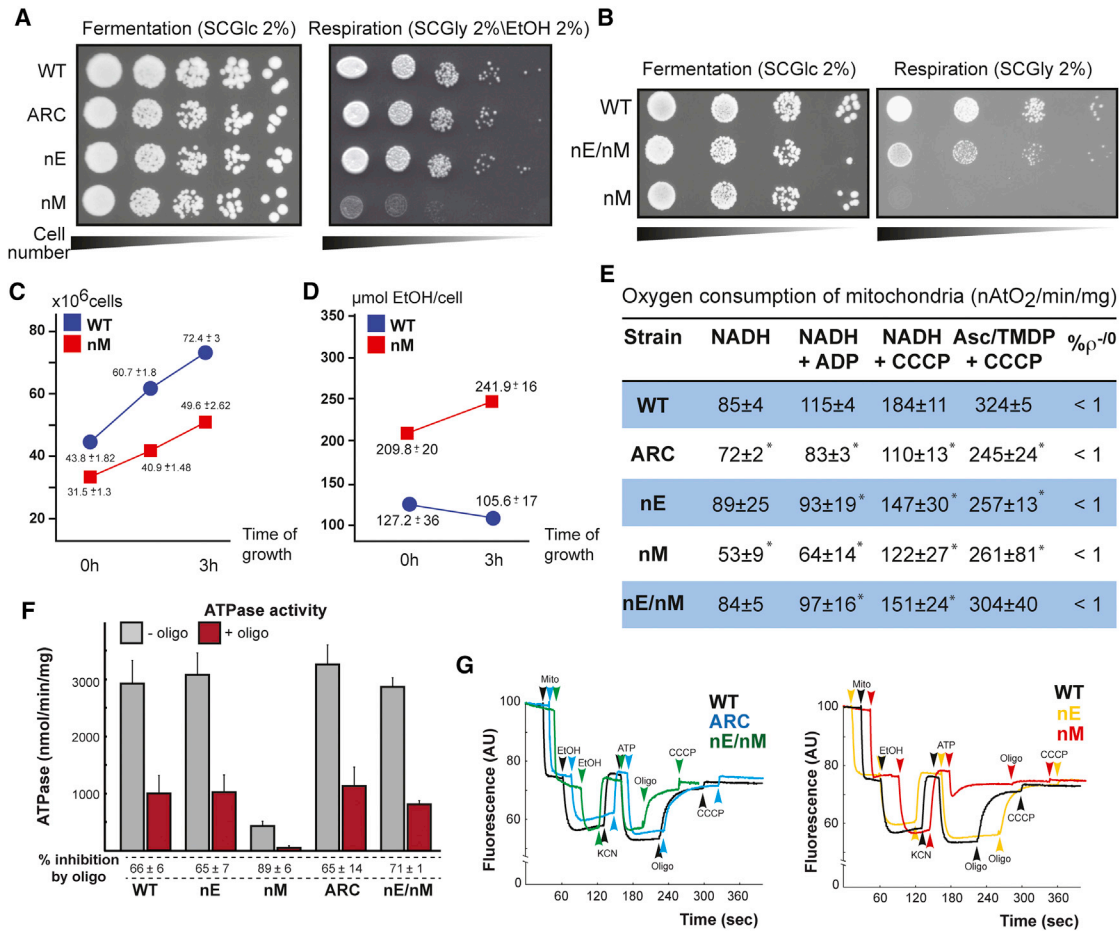


Figure 3. Desynchronized Disassembly of the AME Complex Impairs Growth on Respiratory Media Due to a Severely Diminished ATPase Activity

(A and B) Drop tests of WT, ARC, nE, nM, and nE/nM strains on SCGlc 2% and SCGly 2%/EtOH 2% (A) or SCGly 2% (B).

(C and D) (C) Growth rate comparisons on SCGal 2% and (D) ethanol titration between the WT (blue circle) and nM (red square) strains (n = 3).

(E) Oxygen consumption was measured on isolated mitochondria; percentage of *petite* cells (% ρ^{-0}) in cultures is shown (n = 3). p values were calculated with a Student's z test, *p < 0.001.

(F) Measurements of oligomycin-dependent (+) and -independent (-) ATPase activities on purified mitochondria (SEM, n = 3).

(G) Rhodamine 123 fluorescence quenching on isolated mitochondria. All enzymatic assays (C, D, E, F, and G) were done on extracts prepared from cells grown on SCGal 2%.

respiration is at its maximum). Oxygen consumption was also measured using ascorbate/TMPD that delivers electrons directly at the level of cytochrome c oxidase (complex IV). With respect to the WT, the respiratory activity was diminished

by 15% in the ARC strain and increased by 5%–10% in the nE and nE/nM strains, whereas a 40%–50% deficit was measured in the nM strain (Figure 3E). In all strains, respiration was efficiently (3- to 4-fold) stimulated by CCCP, indicating a normal

(B) Colocalization (arrowheads) of GFP-tagged cMRS with Hoechst-stained nuclei; 24% of yeast cells in respiration conditions show this relocalization (n = 200).

(C) Immunoblot quantification of cMRS levels in normalized (Nop1p, H2Ap) purified (Pgk1p) nuclear extracts from fermenting (ferm) and respiring (resp) cells.

(D) Confocal colocalization with Hoechst-stained nuclei of GFP alone or GFP-tagged nuclear localization sequence of cMRS (NLS-GFP). The NLS (top) is contained within the first 55 amino acids; all cells show the same pattern.

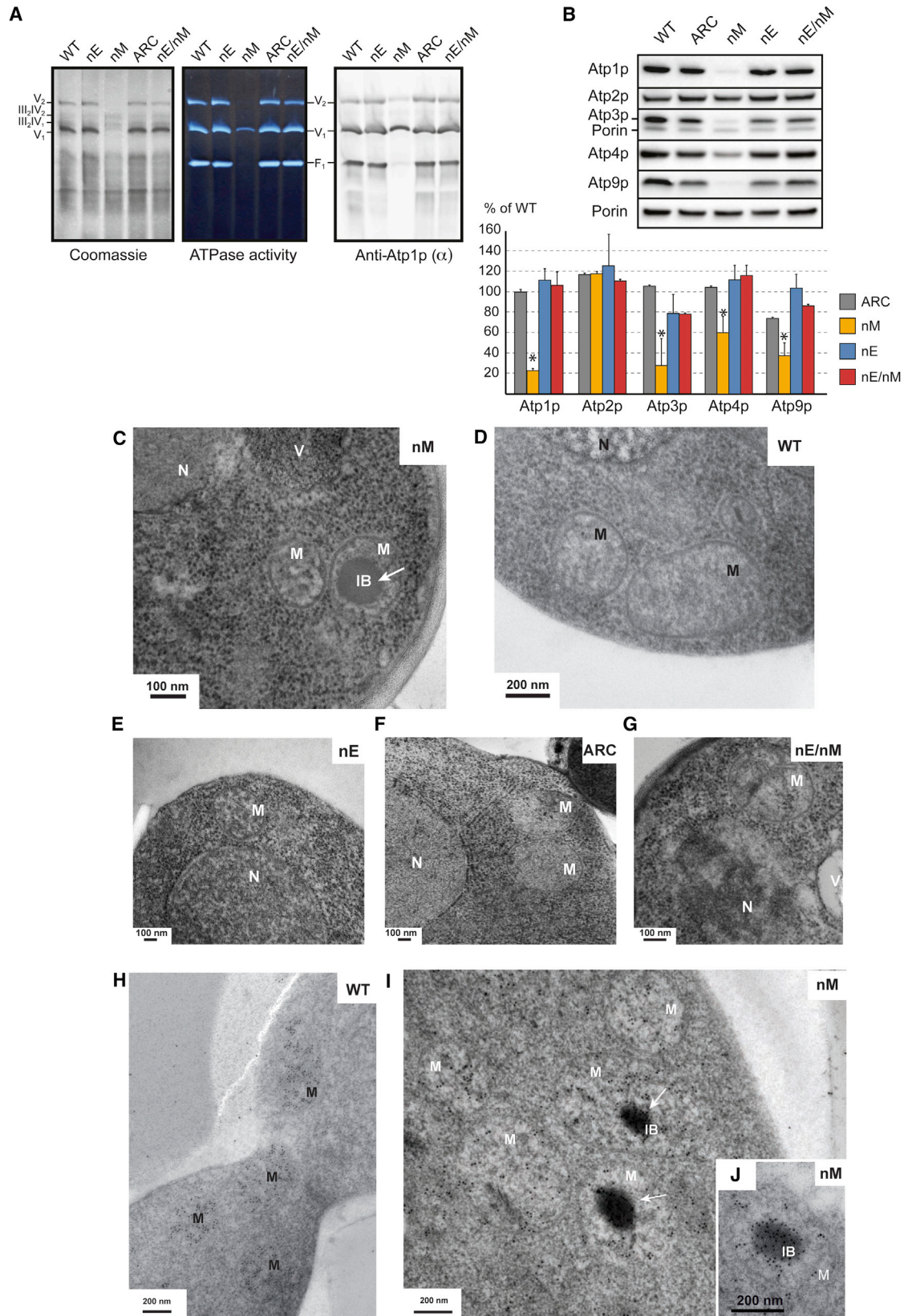
(E) 10% (right) and 12% (left) PAGE of cMRS-TAP interactants pulled down using tandem affinity purification on nuclear protein extracts and identified using mass spectrometry. For spectrum count see Figure S2.

(F) Schematic view of the released or sequestered state of cERS and cMRS in fermenting (ferm) and respiring (resp) wild-type and reprogrammed (ARC, nE, nM, and nE/nM) strains.

(G) Size-exclusion chromatography of total protein extracts from the ARC, nE, and nM strains. Presence of cMRS or cERS in an Arc1p-bound (high molecular weight) or free state (low molecular weight) is monitored by the tRNA^{Met} (for cMRS) or tRNA^{Glu} (for cERS) aminoacylation activity of each fraction.

(H) Immunoblot analysis of the nuclear relocation of artificially released cMRS in the nM and nE/nM strains compared to WT and ARC strains in fermentation (SCGlc 2%).

(I) Same analysis for artificially released cERS relocalization in the mitochondria in nE and nE/nM strains in respiration (SCGly 2%). Scale bars in (A), (B), and (D), 5 μm .



(legend on next page)

passive permeability to protons of the mitochondrial inner membrane.

We measured the rate of ATP hydrolysis in permeabilized mitochondria (Figure 3F). Most (70%–90%) of this activity is normally sustained by the F_1 component of the ATP synthase. Mitochondrial ATPase activity was normal in the ARC, nE, and nE/nM strains, and was efficiently inhibited by oligomycin, a specific inhibitor of the proton-translocating domain of ATP synthase, from which it can be inferred that F_1 and F_0 are correctly coupled in these strains (Figure 3F). In the nM strain, mitochondrial ATPase activity was also sensitive to oligomycin, but was strongly reduced by 85% with respect to the WT (Figure 3F). We next analyzed the mitochondrial membrane potential ($\Delta\Psi$) by fluorescence quenching of rhodamine 123. $\Delta\Psi$ was efficiently maintained by the proton-pumping activity of the respiratory chain (RC) in all strains (Figure 3G). A correct energization of the inner membrane by the proton-pumping activity of the ATP synthase coupled to F_1 -mediated ATP hydrolysis was also observed, but only in the ARC, nE, and nE/nM strains. The inability of the nM strain to energize the membrane with ATP is consistent with its low F_1 -ATPase activity.

Finally, we probed the capacity of the different strains to assemble the OXPHOS complexes by BN- and SDS-PAGE analyses. These complexes were present in normal amounts in ARC, nE, and nE/nM strains, whereas the nM strain had a strongly reduced content in fully assembled ATP synthase (Figure 4A). This defect correlates with low steady-state levels of some ATP synthase subunits, like Atp1p (α -F1) and Atp9p (Figures 4B and S4A). The protein partner of Atp1p, the β - F_1 subunit of ATP synthase (Atp2p), showed a normal accumulation. Since unassembled Atp2p has a strong propensity to aggregate (Lefebvre-Legendre et al., 2005), it was expected to find aggregates of this protein in the nM strain because of its reduced content in Atp1p. Indeed, large matrix-localized inclusion bodies that strongly reacted to gold-conjugated antibodies directed against Atp2p (Figures 4C–4J and S4B, right and left) and that were not labeled with antibodies against Atp1p (Figure S4C) were observed in the nM strain.

Influence of Desynchronized AME Disassembly on Mitochondrial and Nuclear Gene Expression

[35 S]-L-methionine pulse labeling on cycloheximide-treated cells grown on galactose was performed to evaluate the capacity of the AME modified strains to synthesize the proteins encoded by the mtDNA. A striking feature of the nM strain was a strong decrease in Atp9p synthesis (Figure 5A), whereas the other strains did not show any defect in mitochondrial translation (Fig-

ure 5A). These results indicated that the low content of Atp9p in the nM strain is possibly due to a poor synthesis of this protein. We also investigated mitochondrial translation in pure respiratory conditions, 6 hr after switching the cells in a glycerol/ethanol medium (Figures 5B and S5A). In these conditions, the nM strain showed a strong and general decrease in mitochondrial protein synthesis, which is possibly due to the very poor capacity of this strain to produce ATP by OXPHOS (when grown in galactose they can produce ATP by fermentation; Figure S5B). Interestingly, with respect to the WT, the rate of mitochondrial translation was increased by 40% and 54% in respiring nE and nE/nM cells, respectively, which may be due to their higher mitochondrial content in cERS (Figure 5B).

As for Atp9p, the reduced accumulation of Atp1p in the nM strain was apparently due to a decreased rate of synthesis of this protein. Indeed, *ATP1* transcripts were much less (5-fold) abundant in this strain with respect to the WT (Figure 5C; Table S3). Although accumulation of the Atp3p subunit of ATP synthase was reduced in the nM strain, *ATP3* transcripts accumulated normally. This indicates that Atp3p is more susceptible to proteolytic degradation, as was observed in other mutants unable to assemble the F_1 (Lefebvre-Legendre et al., 2001). Consistent with the normal accumulation of Atp2p in the nM strain, accumulation of the *ATP2* transcripts was unaltered with respect to WT.

A strong F_1 defect like the one observed in the nM strain normally leads to low contents in complexes III and IV (Couplan et al., 2011), which is believed to reflect a regulatory mechanism that balances the expression of ATP synthase and respiratory enzymes. In one such mutant (*fmc1* Δ), we found that most of the nuclear-encoded subunits of these complexes are poorly expressed (Couplan et al., 2011). Intriguingly, many of the nuclear transcripts encoding complex III subunits were not down-, but instead upregulated in the nM strain compared to the ARC strain (up to 10-fold, see Figure 5D), while they are downregulated in the nE or ARC strains as compared to WT (Figure 5E), and those of complex IV subunits were not affected (Figure S5C; Tables S4–S6). These data may explain why the nM strain retains a good content in respiratory enzymes despite its major defect in F_1 .

Ectopic Expression of ATP1 Alleviates the Respiratory Phenotype of the nM Strain

Since the most striking effects in the nM strain were compromised expression of Atp1p and Atp9p, we aimed to know if ectopic versions of their genes, i.e., no longer under their normal regulatory sequences, could restore the respiratory function in

Figure 4. Desynchronized AME Complex Disassembly Impairs Expression of Nuclear- and Mitochondrial-Encoded ATP Synthase Subunits and Assembly of the ATP Synthase Enzyme and Promotes Formation of Atp2-Containing Inclusion Bodies

(A) Digitonin-solubilized mitochondrial extracts from WT, ARC, and nE/nM strains and desynchronized strains (nE and nM) were subjected to BN/PAGE prior to Coomassie blue and immunoblot analysis using anti-Atp1p antibodies staining. ATPase complexes (V_1 , monomers; V_2 , dimers) and free F_1 were monitored by ATPase activity. The complexes III and IV (III_2 – IV_2 ; III_2 – IV_1) were monitored by BN-PAGE (Coomassie). Cells were grown on SCGal 2%.

(B) Immunoblot and related quantifications (SEM, * $p < 10^{-4}$ compared to WT, $n = 4$) of ATP synthase subunits Atp1p, -2p, -3p, -4p, and -9p in total cell extracts normalized against porin abundance of strains grown in SCGal 2%.

(C–G) Electron microscopy images of WT and desynchronized strains grown in SCGal 2%. M, mitochondria; N, nucleus; V, vacuole; arrow indicates presence of inclusion bodies (IB).

(H–J) Immunogold detection of Atp2p (β - F_1 subunit) carried out on electron microscopy slides of the WT (H) and nM (I and J) cells grown in SCGal 2% shows that inclusion bodies (IB, arrows) within nM mitochondria (M) contain aggregated Atp2p subunits.

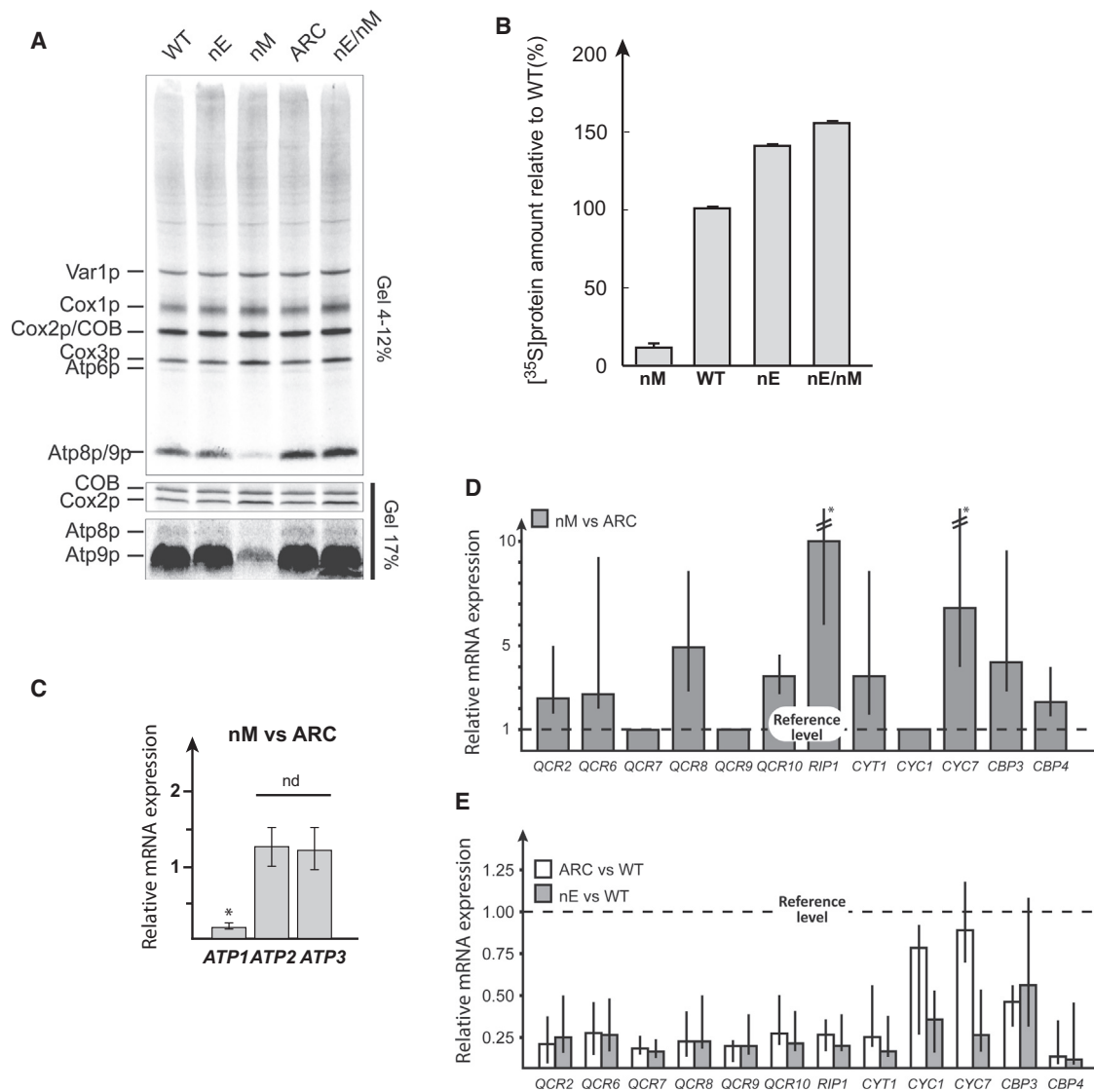


Figure 5. Influence of a Defective Disassembly of the AME Complex on the Transcription of *ATP1* and Nuclear-Encoded Complex III Genes and Mitochondrial Translation Rate

(A) [³⁵S] methionine pulse labeling of proteins translated in mitochondria (30,000 cpm per lane) shows relative variation of single-protein translation within each strain grown in SCGal 2% independent of global mitochondrial translation.

(B) Relative to WT rate of mitochondrial protein synthesis measured in strain constitutively sequestering (nM) or releasing (nE) cERS or both cMRS and cERS (nE/nM). Total protein extracts were prepared from cells incubated in the presence of [³⁵S] methionine and cycloheximide to inhibit cytosolic protein synthesis prior to 12% SDS-PAGE. Radiolabeled protein synthesized in the ARC and nE strains were compared to the WT (grown in SCGly 2%/EtOH 2%) after 20 min of incubation (SEM, $p < 10^{-3}$ compared to WT, $n = 3$).

(C) Quantitative real-time PCR comparative quantification of *ATP1*, -2, and -3 mRNAs in the nM compared to ARC strain grown in SCGal 2% ($*p < 10^{-3}$; nd, not different, $p > 0.05$; SEM, $n = 6$).

(D and E) (D) Quantitative real-time PCR quantification of mRNAs of complex III components when cMRS is artificially released (nM strain) versus confined in the cytoplasm (ARC strain) or (E) when cMRS is artificially confined in the cytoplasm (ARC or nE strain) versus naturally released (WT) in respiration (normal release of cMRS) (SEM, $*p$ value, see Table S3, $n = 4$).

this strain. For these experiments, we used an *ATP1* gene under the dependence of the *PGK* promoter and a nuclear version of the mitochondrial *ATP9* that can complement a strain lacking this gene (Bietenhader et al., 2012). Ectopic expression of *Atp1p* restored the ability of strain nM to grow on respiratory substrates (Figure 6A), whereas the nuclear *ATP9* gene was un-

able to do so (Figures 6A and S6A). The ATP synthase subunits that were in decreased amounts in the nM strain became more abundant with the ectopic *ATP1* gene (Figure 6B), the synthesis of *Atp9p* in mitochondria was improved (Figure 6C), and less *Atp2p* inclusion bodies were found in cells (Figures 6D and S6B). These results indicate that the main factor responsible

for the respiratory deficiency of the nM strain is a defective expression of Atp1p. Around 70% of wild-type (WT) Atp1p amounts in the nM strain are able to restore equivalent levels of Atp3p, -4p, and -9p, while allotopic Atp9p expression fails to do so (Figure 6B). These results are further confirmed using [³⁵S]-L-methionine pulse labeling of mitochondrial proteins from cycloheximide-treated cells. Equivalent amounts of radiolabeled proteins from each strain grown in SCGal 2% were used in order to be independent of global mitochondrial translation, separated by SDS-PAGE, and subjected to autoradiography (Figure 6C). In this experiment we observe that expression of Atp1p restores Atp6p, -8p, and -9p levels, while allotopically expressed Atp9p does, by far, not accumulate to the WT level. Finally, Atp1p ectopic expression reduces the number of Atp2p-containing inclusion bodies found in the nM mitochondria in agreement with the restoration of the required stoichiometry between the α -F₁ (Atp1p) and β -F₁ (Atp2p) subunits of the F₁ domain of ATP synthase (Figure 6D).

Strikingly, resynchronizing the mitochondrial import of cERS with the nuclear import of cMRS (from nM to nE/nM strain) results in a suppression of the nM respiratory defects comparable to that obtained by ectopic expression of Atp1p (Figures 4B, 4G, and 3B–3G) and restores global mitochondrial translation of the nM strain (Figures 5A, 5B, and S5B). A recent study demonstrated that F₁ promotes F₀ mitochondrial translation (Rak and Tzagoloff, 2009). Thus, the very low mitochondrial translation rate originating from cytoplasmic confinement of cERS and decreased levels of Arc1p results in insufficient synthesis of F₀ to allow adequate ATP synthase assembly and activity. Consequently, there are two ways to complement the respiratory growth phenotype of the nM strain, first, re-expressing Atp1p and as a consequence F₀ expression, or second, a general increase of mitochondrial translation by releasing cERS (nE/nM strain) to restore sufficient F₀ levels to sustain ATP synthase assembly and function.

DISCUSSION

MSCs have been found in all eukaryotes studied so far including fungi (Simos et al., 1996), unicellular parasites (Cestari et al., 2013; van Rooyen et al., 2014), worms (Havrylenko et al., 2011), and mammals (Bandyopadhyay and Deutscher, 1971). Investigations carried out to find a functional rationale for the accretion of aaRSs and auxiliary proteins into MSCs were so far only restricted to the mammals, and although the rationale for their existence is still unclear, they seem to be dynamical edifices releasing aaRSs able to exert functions other than tRNA charging upon release.

The present study reveals that the yeast MSC named the AME complex is also dynamic; at the diauxic transition, its two aaRS components (cERS and cMRS) are released due to a strong downregulation of the anchor protein Arc1p by the glucose-sensing Snf1/4 pathway (Figures 1B–1F). We show that released cERS and cMRS partially relocate into mitochondria and the nucleus, respectively (Figures 2A–2D). We further showed that in the nucleus cMRS regulates transcription of some genes of the OXPHOS, and that cERS activates mitochondrial translation upon import (Figure 7). This is an example of a synchronized

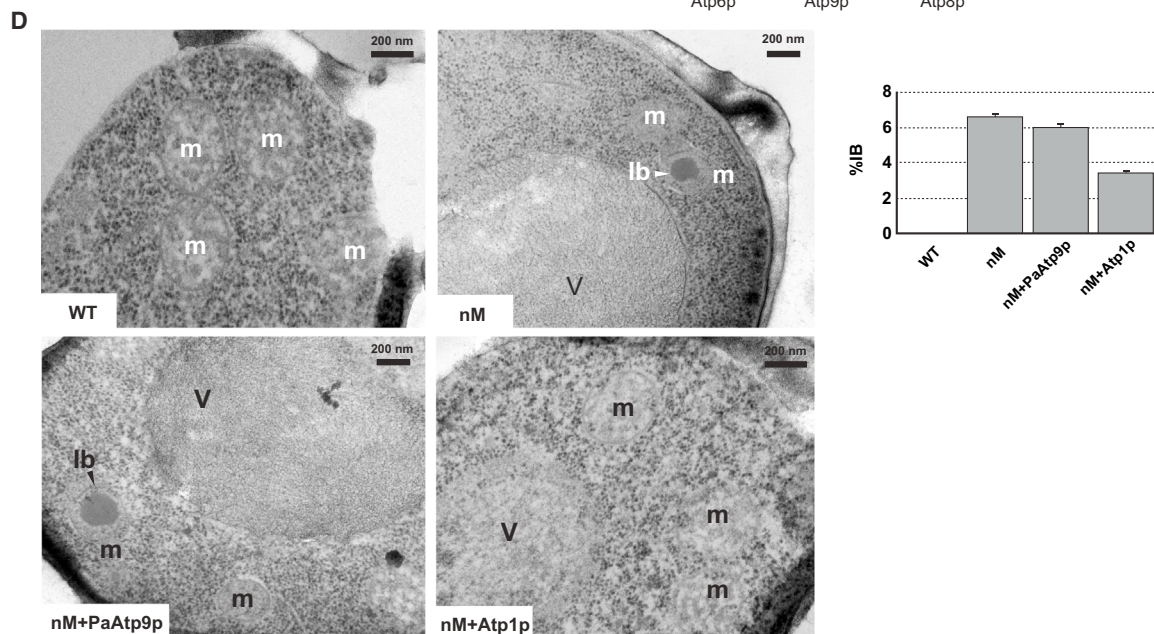
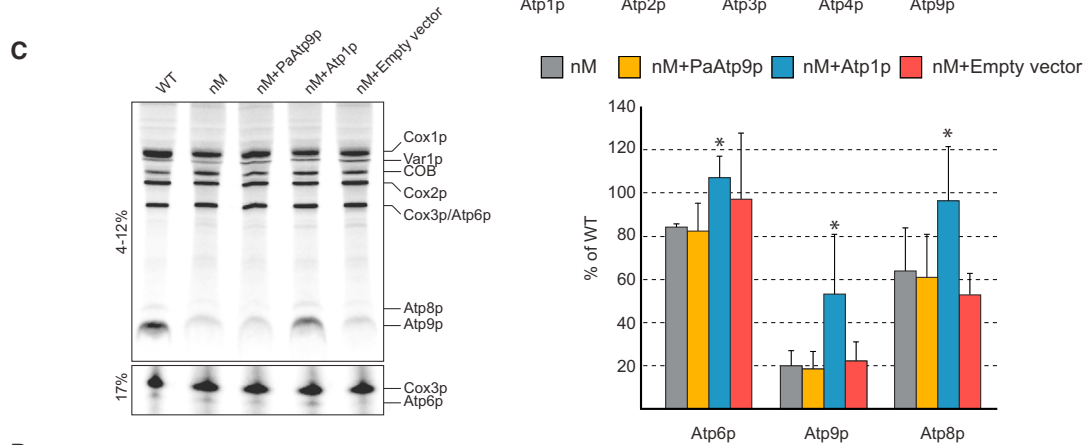
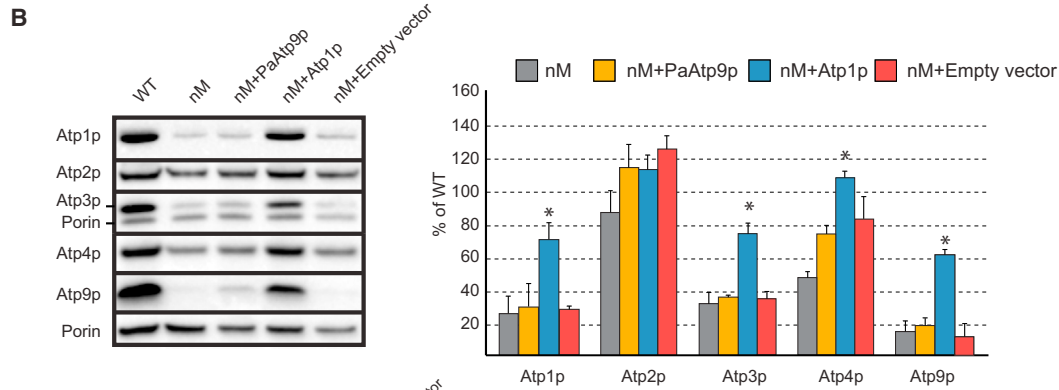
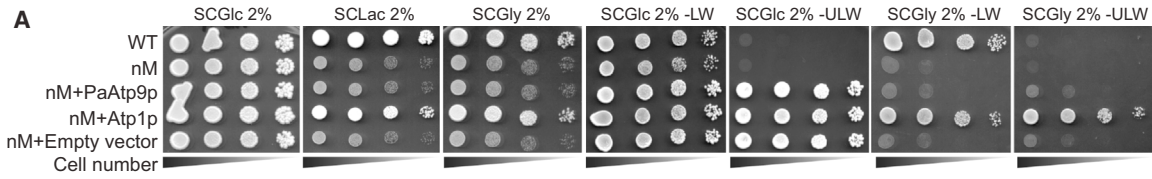
release of multiple aaRSs from an MSC and subsequent relocation to two distinct cellular compartments, nucleus and mitochondria.

In mammals, the accretion of aaRSs into MSCs has no effect on their tRNA charging efficiency and specificity (Mirande et al., 1983). On the opposite, *S. cerevisiae* cMRS and cERS are, respectively, 500- and 10-fold more efficient in tRNA charging when assembled into the AME complex than in free forms (Simos et al., 1996; Graindorge et al., 2005). Since generation of the initiator Met-tRNA^{Met} by cMRS is a rate-limiting step in protein synthesis, the disappearance of Arc1p at the DS should decrease the rate of cytosolic protein synthesis. Since respiring yeast cells grow much (3-fold) slower than fermenting cells, the former necessarily synthesize proteins at a slower rate than the latter. Whether this adaptation of the rate of protein synthesis to the metabolic state of the cell involves a reduction in the tRNA charging activities of cMRS and/or cERS caused by the disassembly of the AME complex is an interesting hypothesis. It was shown that the rapid inhibition of protein synthesis following glucose depletion is controlled by several transcription factors among which Hxk2p, which is, with Mig1p, one of the two transcription factors also inhibiting *ARC1* transcription (Ashe et al., 2000). This suggests that the AME complex dynamics participate in the intimate connection existing between the metabolic state of the yeast cell and its translational activity.

A key feature of the DS is the induction of a profound remodeling of the cellular proteome with a strong enrichment in hundreds of mitochondrial proteins that become essential to provide energy (Haurie et al., 2003). We hypothesized that the AME complex might be involved in this cellular reprogramming. Indeed, upon release from Arc1p, cERS is partly imported into mitochondria where, together with the GatFAB AdT, it generates the mitochondrial Q_m-tRNA^Q necessary for mitochondrial protein synthesis (Frechin et al., 2009). Given that Q_m-tRNA^Q can only be formed by this pathway, the glucose-dependent expression of *ARC1* might thus provide a means to regulate the rate of mitochondrial translation according to the metabolic state of the cell by modulating the amounts of cERS in mitochondria. Consistent with this, the rate of mitochondrial translation was significantly increased (by 50%) when cERS was prevented from interacting with Arc1p by the competitive binding of peptides that correspond to the N-terminal Arc1-binding domain of cERS (Figure 5B).

Stimulating mitochondrial function at the DS requires also nuclear gene expression modifications, which are notably mediated by the yeast HAP transcriptional factor that modulates the expression of dozens of mitochondrial proteins encoded in the nuclear genome (Buschlen et al., 2003). Considering that cMRS and cERS are synchronously released from Arc1p at the DS, and that released cMRS partly relocates to the nucleus, we suspected that cMRS could also play a role in adapting nuclear gene expression when yeast cells rely on their mitochondria to meet their ATP requirements.

To test this hypothesis we engineered strains allowing us to modify the binding and release of cMRS and cERS to/from Arc1p (Figures 2F–2I) and determined how such modifications impact mitochondrial function (Figures 3–5). When release of both aaRSs was blocked by rendering Arc1p's expression



(legend on next page)

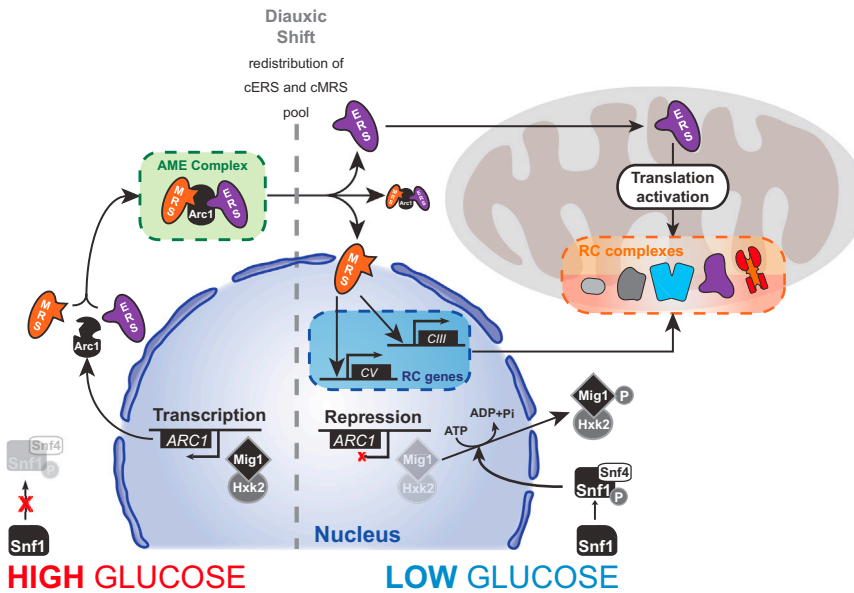


Figure 7. Schematic of AME Regulation and Disassembly upon Shift to Respiration and Its Link with OXPHOS Expression and Assembly

Model of the coordinated intergenomic expression of the OXPHOS subunits mediated by the synchronized relocation of cERS and cMRS upon Snf1/4 kinase-mediated disassembly of the AME complex.

The ATP synthase defect in strain nM apparently resulted from a concomitant failure to express two of its subunits, Atp9p (Figures 4B and 5A) and Atp1p (Figure 4B), which belong, respectively, to the F₀ proton-translocating domain and the F₁ catalytic sector of the enzyme. Atp9p is mitochondrially encoded (Figure S3A) and is present in ten copies (eight in mammals) per F₀ domain. It forms a ring in the inner membrane that

is involved in proton translocation coupled to the generation of a mechanical force that is used to promote ATP synthesis in the F₁. Atp1p is nuclear encoded (Figure S3A) and is present in three copies that together with three Atp2 subunits form the catalytic core of F₁. All the other ATP synthase subunits are in only one copy.

Somewhat remarkably, expression of an ectopic *ATP1* gene under the control of the promoter of the constitutively expressed *PGK1* gene was able to restore substantially Atp9p synthesis and respiratory competence in strain nM (Figures 6A–6C). These results, together with our transcriptomic data (Figure 5C; Table S6), indicate that *ATP1* is a main and quite specific target controlled by AME. The restoration of Atp9p synthesis indicates that the downregulation of this protein in the nM strain may simply be a secondary consequence of the failure to express *ATP1*. However, previous studies have shown that yeast strains lacking *ATP1* (*atp1Δ*) are able to synthesize Atp9p. Instead, *atp1Δ* strains fail to express the two other mitochondrial ATP synthase genes, *ATP8* and *ATP6*, due to a translational block (Rak et al., 2009). The proteins encoded by *ATP6* (Atp6p) and *ATP8* (Atp8p), like Atp9p, are essential for the proton-translocating function of F₀ (Rak et al., 2009). It has been proposed that their F₁-mediated translational control provides a means to avoid accumulation of free F₀ particles, which could otherwise have dramatic consequences by dissipating the mitochondrial membrane potential (Rak et al., 2009). It seems thus unlikely that

Figure 6. Re-Expression of Normal Atp1p Levels in the nM Strain Partly Rescues the Defective Respiration Phenotypes

- (A) Growth test of WT and nM strains rescued for *ATP1* or *-9* expression on fermentative (Glc) or respiration (Gly, Lac) media.
 (B) Immunoblot and related quantifications (SEM, * $p < 10^{-3}$ compared to WT, $n = 3$) of ATP synthase subunits Atp1p, -2p, -3p, -4p, and -9p in total cell extract of strains grown in SCGal 2% shows that the decreased synthesis of Atp1p, -3p, -4p, and -9p subunits in the nM strain can be restored by Atp1p expression.
 (C) [³⁵S] methionine pulse labeling of proteins translated in mitochondria (30,000 cpm per lane) and related quantifications (SEM, * $p < 10^{-3}$ compared to WT, $n = 4$) in WT and nM strains grown on SCGal 2% and rescued for *ATP1* or *-9* expression shows relative variation of single-protein translation within each strain independent of the global mitochondrial translation.
 (D) Electron microscopy and related quantifications of mitochondria with IB in the nM strain, or nM + allotropic expression of Atp9p or expression of Atp1p compared to WT.

the block in Atp9p synthesis in the nM strain is simply due to the failure to express *ATP1*, but is rather a direct consequence of the desynchronized disassembly of AME.

It is interesting to note that the enzymatic partner of cERS in mitochondria, the GatFAB AdT, seems to have also a specific action in the expression of ATP synthase in addition to its conventional role in mitochondrial translation. Indeed, mutations in the GatF subunit of GatFAB can impair respiratory growth by specifically blocking Atp8p translation, a defect that could be suppressed by a nuclear version of the *ATP8* gene (Barros et al., 2011). This defective expression of Atp8p was ascribed to misincorporation of glutamate residues at glutamine codons followed by a rapid degradation of the aberrant protein. The *ATP8* ORF has a high content in glutamine codons (6.3%; Table S7), one of which is in third codon position of the mRNA. The early placement in mRNAs of codons for which the amounts of corresponding aa-tRNA species is insufficient is especially detrimental to ribosomal elongation (Cannarozzi et al., 2010; Tuller et al., 2010), which may explain the strong dependency of *ATP8* mRNA translation toward the cERS-AdT pathway. While Atp9p has a moderate content in glutamine (1.3%; Table S7), it is to be noted that a glutamine codon immediately follows the initiator methionine, and that this protein is in ten copies in the ATP synthase, whereas Atp8 and Atp6p are in only one copy. A proper expression of the Atp9p oligomer should thus also be highly demanding in Q_{-m}tRNA^Q.

The low levels of *ATP1* transcripts in the nM strain indicate that the defective expression of this subunit results from a decreased transcriptional activity or a reduced stability of these transcripts (Figure 5C; Table S6). Like yeast cMRS, human cMRS is sequestered within an MSC and can be released from it by a variety of stimuli, after which it relocates to the nucleolus to activate gene transcription (Ko et al., 2000). It is thus a reasonable hypothesis that a proper expression of *ATP1* in yeast is controlled by the nuclear relocation of cMRS. However, intriguingly, *ATP1* expression was never compromised in strains where cMRS was prevented from reaching the nucleus. This expression was compromised only when cMRS was released from Arc1p, while cERS remained mostly attached to Arc1p. This indicates that the cMRS-mediated expression of *ATP1* is activated by some retrograde signal induced by cERS in mitochondria, possibly as a means to equilibrate the production of the F₁ and F_O sectors of ATP synthase.

Further work will be required to understand mechanistically how ATP synthase expression is influenced by the AME complex. However, it can already be inferred from the data presented here that this MSC has a specific action in the biogenesis of this major mitochondrial enzymatic system. A key feature of this control is a concerted modulation of nDNA- and mtDNA-encoded subunits of the ATP synthase. Our work highlights the existence of an additional degree of complexity in the dynamics of MSCs: the simultaneous release of several response factors. Synchronously releasing several aaRSs might give the cell the ability to coordinate responses mediated by several regulatory networks within one or several compartments of the cell. The increased number of aaRSs and MSC assembly factors in higher eukaryotes may be viewed as a means to expand the diversity of the response repertoire to stimuli.

EXPERIMENTAL PROCEDURES

Engineering of the ARC, nE, nM, and nE/nM Strains

The *ARC1*, *N_ERS*, and *N_MRS* DNA were subcloned into p425GPD (*ARC1*) and p424GPD. An *arc1*Δ strain was transformed with the p425GPD-*ARC1* recombinant plasmid and subsequently transformed either with the empty p424GPD, p424GPD-*N_ERS*, or p424GPD-*N_MRS* plasmids yielding, respectively, the ARC, nE, and nM strains. For p424GPD-*N_ERS*, the 190 first aa do not contain the putative mitochondrial targeting sequence located directly downstream in the native N terminus of cERS. To generate the nE/nM strain, the *TRP1* auxotrophic marker of the p424GPD-*N_MRS* was changed to *HIS3*, and the resulting p423GPD-*N_MRS* was expressed in an nE strain. The control WT strain was WT cotransformed with the p425GPD and p424GPD empty plasmids. For more detailed experimental procedures see Supplemental Experimental Procedures.

Freezing and Freeze Substitution for Ultrastructural Studies

The yeast pellets were placed on the surface of a copper EM grid (400 mesh) coated with formvar. Each grid was very quickly submerged in liquid propane precooled and held at -180°C using liquid nitrogen. Loops were transferred in a precooled solution of 4% osmium tetroxide in dry acetone in a 1.8 ml polypropylene vial at -82°C for 72 hr (substitution), warmed gradually to room temperature, and followed by three washes in dry acetone. The samples were infiltrated progressively with araldite (epoxy resin [Fluka]). Ultrathin sections were contrasted with lead citrate. Specimens were observed with a Hitachi 7650 (80 kV) electron microscope (Electronic Imaging Pole of Bordeaux Imaging Center). The percentage of inclusion bodies was calculated from the observation of 150–460 mitochondria per strain. The percentage of inclusion bodies found in the WT strain was less than 0.1%.

Immunogold Electron Microscopy

Yeasts cells were cryofixed and freeze substituted with acetone plus 0.1% uranyl acetate for 3 days at -82°C. Samples were rinsed with acetone at -20°C and embedded progressively at -20°C in LR Gold resin (EMS). Resin polymerization was carried out at -20°C for 3 days under UV illumination. Ultrathin LR Gold sections were collected on formvar-coated nickel grids. Sections were first incubated in TBS 20 mM (pH 7.8) for 5 min and then 15 min in NH₄Cl 500 mM in TBS 20 mM (pH 7.8). After rinsing, the grids were incubated with antibodies against the ATP synthase α or β subunit (1:2,500) for 1 hr at room temperature and then with an anti-rabbit conjugated to 10 nm gold particles (1:20) (Biocell) for another hour at room temperature. The sections were rinsed with TBS 20 mM (pH 7.8), incubated 15 min in glutaraldehyde 1%, rinsed and contrasted for 5 min with 2% uranyl acetate in water, followed by 1% lead citrate for 1 min. Specimens were observed with a Hitachi 7650 electron microscope (Electronic Imaging Pole of Bordeaux Imaging Center).

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, seven tables, and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2014.10.015>.

AUTHOR CONTRIBUTIONS

M.F., L.E., E.T., B.S., and D.L. performed cloning experiments, complementation studies, growth curves, and genetics. L.E. and D.L. performed the fluorescence microscopy analyses. M.F., L.E., E.T., D.L., and G.B. performed the subcellular fractionations and immunoblot analyses. M.F. performed the aminoacylation activity measurements and FACS experiments. L.E. made the ethanol consumption analyses. E.T. and M.F. made the mitochondrial de novo protein synthesis analyses. E.T. performed the BN/PAGE and the measurements of OXPHOS activities. M.F., D.L., L.E., B.S., and P.H. performed the TAP experiments and mass spectrometry studies. M.F., L.E., S.C.-M., and L.M.S. performed the transcriptomic analyses. C.B. and E.T. performed the electron microscopy studies. M.F., L.E., E.T., B.S., J.-P.d.R., and H.D.B. designed the study. M.F., L.E., E.T., B.S., R.P.M., L.M.S., J.-P.d.R., and

H.D.B. analyzed the data. R.P.M. and L.M.S. edited the manuscript. M.F., L.E., E.T., J.-P.d.R., and H.D.B. wrote the paper.

ACKNOWLEDGMENTS

The work was supported by the French National Program Investissement d'Avenir administered by the "Agence Nationale de la Recherche" (ANR), "MitoCross" Laboratory of Excellence (Labex), funded as ANR-10-IDEX-0002-02 (to H.B.), the ANR-09-BLAN-0091-02 (to H.B.), the ANR Chloro/Mito CES (to J.-P.d.R.), the University of Strasbourg (to H.B.), the JST-CNRS Japanese-French Cooperative Program on "Structure and Function of Biomolecules" (to H.B.), and the CNRS. M.F. was supported by a fellowship from the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie and by the Association Française contre les Myopathies. L.E. and G.B. are supported by the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie. D.L. is supported by the ANR-10-IDEX-0002-02 "Investissements d'Avenir" program.

Received: December 2, 2013
Revised: September 4, 2014
Accepted: October 16, 2014
Published: November 20, 2014

REFERENCES

- Ackerman, S.H., and Tzagoloff, A. (2005). Function, structure, and biogenesis of mitochondrial ATP synthase. *Prog. Nucleic Acid Res. Mol. Biol.* **80**, 95–133.
- Ashe, M.P., De Long, S.K., and Sachs, A.B. (2000). Glucose depletion rapidly inhibits translation initiation in yeast. *Mol. Biol. Cell* **11**, 833–848.
- Bandyopadhyay, A.K., and Deutscher, M.P. (1971). Complex of aminoacyl-transfer RNA synthetases. *J. Mol. Biol.* **60**, 113–122.
- Barros, M.H., Rak, M., Paulela, J.A., and Tzagoloff, A. (2011). Characterization of Gtf1p, the connector subunit of yeast mitochondrial tRNA-dependent amidotransferase. *J. Biol. Chem.* **286**, 32937–32947.
- Bietenhader, M., Martos, A., Tetaud, E., Aiyar, R.S., Sellem, C.H., Kucharczyk, R., Clauder-Münster, S., Giraud, M.F., Godard, F., Salin, B., et al. (2012). Experimental relocation of the mitochondrial ATP9 gene to the nucleus reveals forces underlying mitochondrial genome evolution. *PLoS Genet.* **8**, e1002876.
- Brown, M.V., Reader, J.S., and Tzima, E. (2010). Mammalian aminoacyl-tRNA synthetases: cell signaling functions of the protein translation machinery. *Vascul. Pharmacol.* **52**, 21–26.
- Buschlen, S., Amillet, J.M., Guiard, B., Fournier, A., Marcireau, C., and Bolotin-Fukuhara, M. (2003). The *S. cerevisiae* HAP complex, a key regulator of mitochondrial function, coordinates nuclear and mitochondrial gene expression. *Comp. Funct. Genomics* **4**, 37–46.
- Cannarozzi, G., Schraudolph, N.N., Faty, M., von Rohr, P., Friberg, M.T., Roth, A.C., Gonnet, P., Gonnet, G., and Barral, Y. (2010). A role for codon order in translation dynamics. *Cell* **141**, 355–367.
- Cestari, I., Kalidas, S., Monnerat, S., Anupama, A., Phillips, M.A., and Stuart, K. (2013). A multiple aminoacyl-tRNA synthetase complex that enhances tRNA-aminoacylation in African trypanosomes. *Mol. Cell. Biol.* **33**, 4872–4888.
- Couplan, E., Aiyar, R.S., Kucharczyk, R., Kabala, A., Ezkurdia, N., Gagneur, J., St Onge, R.P., Salin, B., Soubigou, F., Le Cann, M., et al. (2011). A yeast-based assay identifies drugs active against human mitochondrial disorders. *Proc. Natl. Acad. Sci. USA* **108**, 11989–11994.
- Duvezin-Caubet, S., Caron, M., Giraud, M.-F., Velours, J., and di Rago, J.-P. (2003). The two rotor components of yeast mitochondrial ATP synthase are mechanically coupled by subunit δ . *Proc. Natl. Acad. Sci. USA* **100**, 13235–13240.
- Fox, T.D. (2012). Mitochondrial protein synthesis, import, and assembly. *Genetics* **192**, 1203–1234.
- Frechin, M., Senger, B., Brayé, M., Kern, D., Martin, R.P., and Becker, H.D. (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.
- Frechin, M., Kern, D., Martin, R.P., Becker, H.D., and Senger, B. (2010). Arc1p: anchoring, routing, coordinating. *FEBS Lett.* **584**, 427–433.
- Graindorge, J.S., Senger, B., Tritch, D., Simos, G., and Fasiolo, F. (2005). Role of Arc1p in the modulation of yeast glutamyl-tRNA synthetase activity. *Biochemistry* **44**, 1344–1352.
- 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.
- Haurie, V., Boucherie, H., and Sagliocco, F. (2003). The Snf1 protein kinase controls the induction of genes of the iron uptake pathway at the diauxic shift in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **278**, 45391–45396.
- Hausmann, C.D., and Ibba, M. (2008). Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. *FEMS Microbiol. Rev.* **32**, 705–721.
- Havrylenko, S., Legouis, R., Negrutskii, B., and Mirande, M. (2011). *Caenorhabditis elegans* evolves a new architecture for the multi-aminoacyl-tRNA synthetase complex. *J. Biol. Chem.* **286**, 28476–28487.
- Hedbacker, K., and Carlson, M. (2008). SNF1/AMPK pathways in yeast. *Front. Biosci.* **13**, 2408–2420.
- Ibba, M., and Söll, D. (2000). Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* **69**, 617–650.
- 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.
- Ko, Y.G., Kang, Y.S., Kim, E.K., Park, S.G., and Kim, S. (2000). Nucleolar localization of human methionyl-tRNA synthetase and its role in ribosomal RNA synthesis. *J. Cell Biol.* **149**, 567–574.
- Lefebvre-Legendre, L., Vaillier, J., Benabdelhak, H., Velours, J., Slonimski, P.P., and di Rago, J.-P. (2001). Identification of a nuclear gene (FMC1) required for the assembly/stability of yeast mitochondrial F1-ATPase in heat stress conditions. *J. Biol. Chem.* **276**, 6789–6796.
- Lefebvre-Legendre, L., Salin, B., Schaëffer, J., Brêthes, D., Dautant, A., Ackerman, S.H., and di Rago, J.-P. (2005). Failure to assemble the alpha 3 beta 3 subcomplex of the ATP synthase leads to accumulation of the alpha and beta subunits within inclusion bodies and the loss of mitochondrial cristae in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **280**, 18386–18392.
- McCartney, R.R., and Schmidt, M.C. (2001). Regulation of Snf1 kinase. Activation requires phosphorylation of threonine 210 by an upstream kinase as well as a distinct step mediated by the Snf4 subunit. *J. Biol. Chem.* **276**, 36460–36466.
- Mirande, M., Cirakoğlu, B., and Waller, J.-P. (1983). Seven mammalian aminoacyl-tRNA synthetases associated within the same complex are functionally independent. *Eur. J. Biochem.* **131**, 163–170.
- Negrutskii, B.S., Stapulionis, R., and Deutscher, M.P. (1994). Supramolecular organization of the mammalian translation system. *Proc. Natl. Acad. Sci. USA* **91**, 964–968.
- Ostergaard, S., Olsson, L., Johnston, M., and Nielsen, J. (2000). Increasing galactose consumption by *Saccharomyces cerevisiae* through metabolic engineering of the GAL gene regulatory network. *Nat. Biotechnol.* **18**, 1283–1286.
- Otterstedt, K., Larsson, C., Bill, R.M., Ståhlberg, A., Boles, E., Hohmann, S., and Gustafsson, L. (2004). Switching the mode of metabolism in the yeast *Saccharomyces cerevisiae*. *EMBO Rep.* **5**, 532–537.
- Park, S.G., Schimmel, P., and Kim, S. (2008). Aminoacyl tRNA synthetases and their connections to disease. *Proc. Natl. Acad. Sci. USA* **105**, 11043–11049.
- Polakis, E.S., Bartley, W., and Meek, G.A. (1965). Changes in the activities of respiratory enzymes during the aerobic growth of yeast on different carbon sources. *Biochem. J.* **97**, 298–302.
- Rak, M., and Tzagoloff, A. (2009). F1-dependent translation of mitochondrially encoded Atp6p and Atp8p subunits of yeast ATP synthase. *Proc. Natl. Acad. Sci. USA* **106**, 18509–18514.

- Rak, M., Zeng, X., Brière, J.J., and Tzagoloff, A. (2009). Assembly of F₀ in *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* 1793, 108–116.
- Ray, P.S., Arif, A., and Fox, P.L. (2007). Macromolecular complexes as depots for releasable regulatory proteins. *Trends Biochem. Sci.* 32, 158–164.
- Roy, H., Becker, H.D., Reinbolt, J., and Kern, D. (2003). When contemporary aminoacyl-tRNA synthetases invent their cognate amino acid metabolism. *Proc. Natl. Acad. Sci. USA* 100, 9837–9842.
- Sampath, P., Mazumder, B., Seshadri, V., Gerber, C.A., Chavatte, L., Kinter, M., Ting, S.M., Dignam, J.D., Kim, S., Driscoll, D.M., and Fox, P.L. (2004). Noncanonical function of glutamyl-prolyl-tRNA synthetase: gene-specific silencing of translation. *Cell* 119, 195–208.
- Santangelo, G.M. (2006). Glucose signaling in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* 70, 253–282.
- Schimmel, P., and Ribas De Pouplana, L. (2000). Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem. Sci.* 25, 207–209.
- Schüller, H.J. (2003). Transcriptional control of nonfermentative metabolism in the yeast *Saccharomyces cerevisiae*. *Curr. Genet.* 43, 139–160.
- Simader, H., Hothorn, M., Köhler, C., Basquin, J., Simos, G., and Suck, D. (2006). Structural basis of yeast aminoacyl-tRNA synthetase complex formation revealed by crystal structures of two binary sub-complexes. *Nucleic Acids Res.* 34, 3968–3979.
- Simos, G., Segref, A., Fasiolo, F., Hellmuth, K., Shevchenko, A., Mann, M., and Hurt, E.C. (1996). The yeast protein Arc1p binds to tRNA and functions as a cofactor for the methionyl- and glutamyl-tRNA synthetases. *EMBO J.* 15, 5437–5448.
- Söll, D., and Schimmel, P.R. (1974). Aminoacyl-tRNA synthetases. *Enzyme* 10, 489–538.
- Stewart, T.S., Roberts, R.J., and Strominger, J.L. (1971). Novel species of tRNA. *Nature* 230, 36–38.
- Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., and Pilpel, Y. (2010). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* 141, 344–354.
- van Rooyen, J.M., Murat, J.B., Hammoudi, P.M., Kieffer-Jaquinod, S., Coute, Y., Sharma, A., Pelloux, H., Belrhali, H., and Hakimi, M.A. (2014). Assembly of the novel five-component apicomplexan multi-aminoacyl-tRNA synthetase complex is driven by the hybrid scaffold protein Tg-p43. *PLoS One* 9, e89487.
- Zaman, S., Lippman, S.I., Zhao, X., and Broach, J.R. (2008). How *Saccharomyces* responds to nutrients. *Annu. Rev. Genet.* 42, 27–81.

Partie II : La Split-GFP, un outil pour
l'étude des protéines à localisations
subcellulaires multiples chez *S.*
cerevisiae

1. Problématique du projet

Lors de l'étude du complexe AME de levure, nous avons pu mettre en évidence la capacité de deux de ses membres à se relocaliser dans d'autres compartiments cellulaires afin d'assurer des fonctions non conventionnelles, indispensables à la croissance de la levure en respiration. En effet, les deux aaRS du MSC, la MetRS et la GluRS, sont multi-localisées et se retrouvent à la fois dans le cytosol et dans le noyau (MetRS) ou la mitochondrie (GluRS). De plus, Arc1, la 3^{ème} protéine du complexe, est également présente dans plusieurs compartiments au même moment et est détectée aussi bien dans le cytosol que dans le noyau ou encore associée à des membranes (voir introduction § VI.2.c). Nous nous retrouvons donc dans la situation où toutes les protéines du complexe AME sont présentes dans plusieurs compartiments au même moment et la validation de ces localisations multiples est d'autant plus compliquée qu'il faut visualiser le(s) compartiment(s) additionnel(s). Lors de l'étude de la relocalisation de la GluRS dans la mitochondrie, cette protéine a été fusionnée à la GFP pour permettre son observation par microscopie à fluorescence. Cependant, lorsqu'une protéine se trouve dans le cytosol et dans un autre compartiment au même moment, il est très difficile de distinguer la fluorescence issue de la fraction cytosolique de celle issue des fractions organellaires. Pour visualiser le compartiment cellulaire dans lequel pourrait se trouver la protéine d'intérêt, il est indispensable de réaliser un co-marquage du compartiment avec une molécule fluorescente ou un composé chimique. Pour étudier la localisation mitochondriale de la GluRS-GFP, ce compartiment a été marqué au Mitotracker. La colocalisation de la fluorescence verte avec le marquage des mitochondries, observée par microscopie confocale, a permis de visualiser la localisation organellaire de la GluRS (Frechin et al., 2009b). Cependant, la simple utilisation de la fluorescence n'a pas suffi pour valider cette localisation et des techniques biochimiques ont été nécessaires pour prouver les résultats obtenus par microscopie. C'est uniquement la purification de mitochondries, suivie d'immuno-détections avec des témoins de pureté des fractions subcellulaires (Pgk1 (cytosol), mHsp70 (mitochondrie)), ainsi qu'un anticorps dirigé contre la GluRS (@-GluRS) qui a permis de confirmer la localisation mitochondriale de cette protéine.

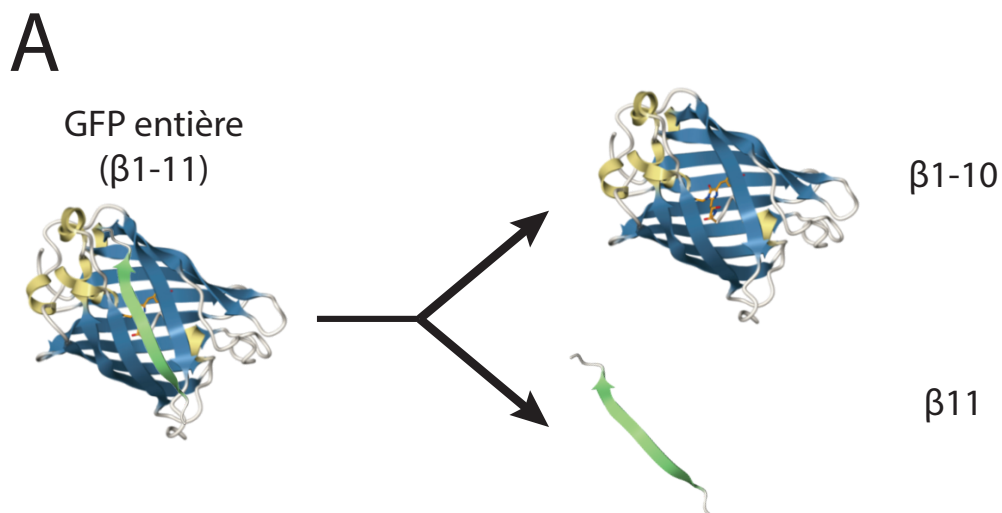
Concernant la MetRS, les mêmes difficultés ont été rencontrées et la visualisation de cette protéine fusionnée à la GFP, combinée au marquage nucléaire au Hoechst, n'a pas suffi pour valider la localisation additionnelle de cette protéine. C'est également la purification du

compartiment d'intérêt, suivie d'immuno-détections de la MetRS qui ont permis de corroborer l'hypothèse. Ces difficultés ne sont pas spécifiques au complexe AME et à notre problématique de recherche. Ainsi, de nombreuses études suggèrent que dans des conditions particulières ou suite à des stimuli/stress spécifiques, certaines protéines essentielles peuvent se relocaliser. En effet, chez la levure, certaines protéines nucléaires se relocalisent dans le cytosol. C'est notamment le cas de la Cycline C, dont la relocalisation est dépendante du stress oxydatif (Cooper et al., 2012) ou de la protéine Kss1 qui se relocalise suite à une stimulation par des phéromones (Pelet, 2017). Par ailleurs, la problématique de la multi-localisation n'est pas restreinte à la levure. Chez l'Homme, les *PolyA binding protein 1* et 4 (PABP), localisées au niveau du cytosol, sont capables de se relocaliser dans le noyau en réponse à une irradiation UV et sont retenues dans ce compartiment. Ces deux protéines se fixent alors au niveau de certains ARNm et favorisent l'accumulation des complexes ARNm/PABP dans le noyau. Ceci induit alors le ralentissement de la synthèse protéique cytosolique (Burgess et al., 2011).

Du fait de l'importance de la relocalisation des protéines pour la modulation de leurs activités, ce phénomène est de plus en plus étudié et c'est notamment le cas pour les aaRS. D'ailleurs, ces protéines, cruciales pour la synthèse protéique, se relocalisent dans divers compartiments cellulaires pour assurer des fonctions additionnelles tout aussi importantes. Les méthodes de détection de la localisation sont multiples et différentes en fonction des organismes, mais ont toutes un point commun, elles nécessitent leur combinaison pour valider la localisation de la protéine. Cette capacité des aaRS à se relocaliser dans d'autres compartiments subcellulaires, ainsi que les méthodes d'étude de ces relocalisations ont été l'objet d'une revue dont je suis co-premier auteur (**Annexe 2**).

2. Ingénierie de l'outil de localisation basé sur la Split-GFP

L'étude de la localisation subcellulaire des protéines n'est pas une chose aisée et la détection de la localisation par fusion avec une molécule fluorescente n'est pas optimale pour visualiser les localisations additionnelles, notamment lorsqu'il s'agit de fractions minoritaires d'une protéine largement cytosolique. Pour remédier à ces problèmes de détection, le principal projet de ma thèse a été d'élaborer un outil permettant de visualiser uniquement la fraction organellaire des protéines multi-localisées. Pour ce faire, en collaboration avec le Dr. Arais, post-doctorant au laboratoire, nous avons décidé d'utiliser un outil de fluorescence



B

β 11	β 1-10
Leu221His, Phe223Tyr, Thr225Asn	Ser30Arg, Asn39Ile, Thr105Lys, Glu111Val, Ile128Thr, Tyr145Phe, Lys166Thr, Ile167Val, Ile171Val, Ala206Val, Ser205Thr

Figure 26 : Principe de la Split-GFP et mutations

(A) La GFP, comprenant onze brins β , est séparée en deux parties : un fragment β 1-10, composé des dix premiers brins et un fragment β 11, formé uniquement du dernier brin. (B) Les mutations présentées dans ce tableau sont ajoutées aux fragments de la Split-GFP pour permettre d'augmenter leur solubilité et leur capacité d'autoassemblage (Cabantous et al., 2005 ; Pedelacq et al., 2006).

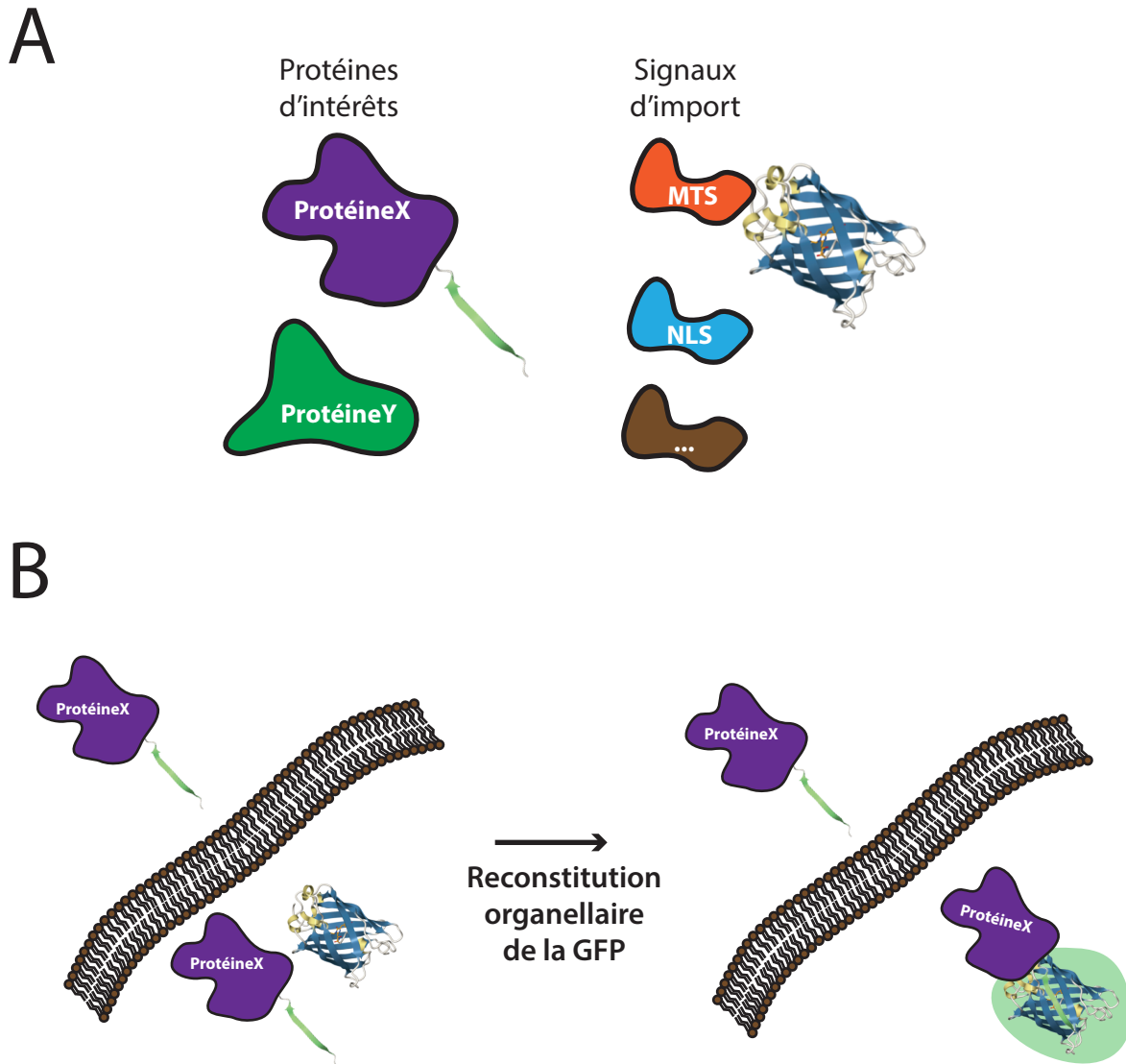


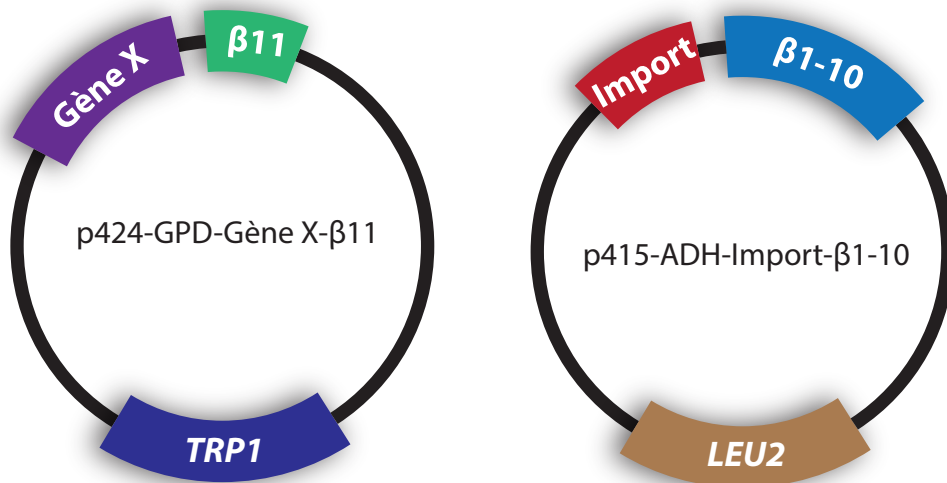
Figure 27 : Etude de la localisation subcellulaire avec la Split-GFP

(A) Les fragments β 1-10 sont exprimés en fusion avec différents signaux d'import induisant leur relocalisation constitutive dans un compartiment donné, tandis que les fragments β 11 sont fusionnés aux protéines dont on souhaite étudier la localisation. (B) Le fragment β 1-10 étant restreint à un compartiment spécifique, c'est uniquement si la protéine d'intérêt, fusionnée au fragment β 11, se relocalise dans ce compartiment, que les deux fragments s'autoassemblent et permettent la reconstitution de la molécule active et la production de fluorescence. Si la protéine est multi-localisée, sa fraction présente dans un autre compartiment ne permet pas la reconstitution de la GFP.

existant, la Split-GFP, et de le détourner pour favoriser l'absence de la fluorescence issue de la fraction cytosolique d'une protéine multi-localisée. Pour cela, nous avons donc utilisé la technique de Split-GFP qui consiste en la séparation de la GFP, contenant onze brins β , en deux fragments distincts : le fragment β 1-10 et le fragment β 11 (**Figure 26A**). Nous avons plus particulièrement utilisé le variant optimisé du fragment β 1-10 comprenant les mutations : Ser30Arg, Asn39Ile, Thr105Lys, Glu111Val, Ile128Thr, Tyr145Phe, Lys166Thr, Ile167Val, Ile171Val, Ala206Val, Ser205Thr. Le fragment β 11 utilisé contient, quant à lui, les mutations : Leu221His, Phe223Tyr, Thr225Asn (**Figure 26B**). Contrairement aux prérequis lors de son utilisation pour l'étude de l'interaction entre deux protéines, la capacité d'autoassemblage des deux fragments est primordiale pour que notre outil soit efficace et permette de visualiser l'apparition d'un signal fluorescent lors de l'import des fragments dans le compartiment cible. En effet, nous avons décidé de fusionner les protéines d'intérêts au fragment β 11, tandis que le fragment β 1-10 est, quant à lui, associé à des signaux d'adressage spécifiques (**Figure 27A**). La localisation de ce dernier est restreinte à un compartiment subcellulaire donné et c'est uniquement si la protéine d'intérêt, fusionnée au fragment β 11, se trouve dans ce même compartiment, que les fragments β 1-10 et β 11 s'autoassemblent, reconstituant ainsi la GFP active et permettant l'émission de fluorescence (**Figure 27B**). Lors de l'étude de protéines multi-localisées, c'est uniquement la fraction organellaire de la protéine qui induit la reconstitution de la molécule active, alors que la fraction cytosolique, qui n'est pas présente dans le même compartiment que le fragment β 1-10, ne permet pas l'émission de fluorescence.

La première partie de l'élaboration de cet outil a été de l'optimiser pour son utilisation chez la levure et notamment au niveau de l'usage des codons qui influe fortement sur l'expression de la protéine et sur l'intensité de fluorescence (Kaishima et al., 2016). De plus, un espaceur (linker) a été ajouté pour éviter une gêne stérique lors de la reconstitution de la GFP active. Les séquences optimisées ont ensuite été clonées dans des vecteurs d'expression. Pour ce faire, nous avons d'abord élaboré un système à deux plasmides : un plasmide centromérique (p415-ADH) et un plasmide 2μ (p424-GPD) possédant tous les deux des promoteurs forts. L'expression de la protéine d'intérêt est l'étape critique pour la détection de sa localisation subcellulaire. En effet, la protéine doit être produite en quantité suffisante pour permettre sa détection. Cependant, l'expression ne doit pas être trop importante pour

A



B

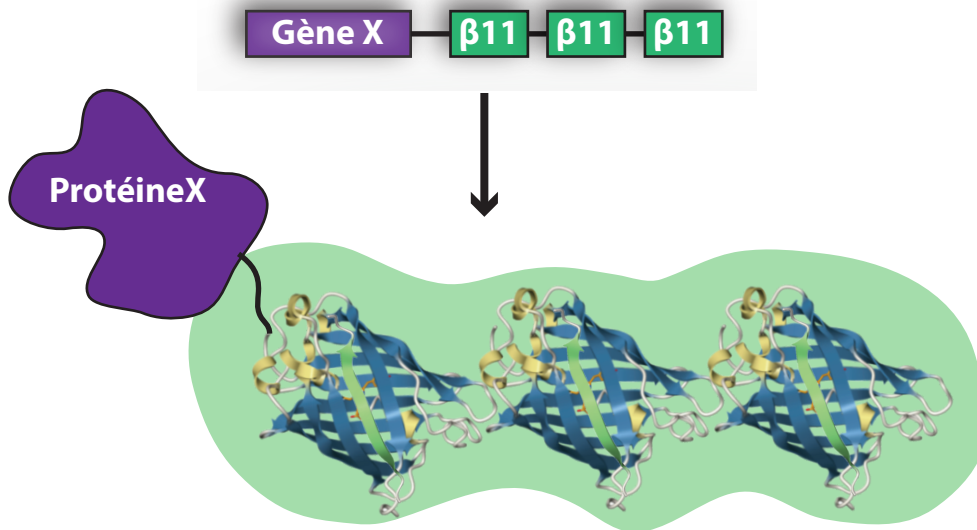


Figure 28 : Stratégies utilisées

(A) Vecteurs utilisés pour l'expression des fragments β1-10 et β11 chez la levure. Les gènes d'intérêts sont clonés en amont de la séquence codant ces fragments. Le vecteur p424-GPD-Gène X-β11 est un vecteur multicopie permettant une expression forte de la protéine d'intérêt fusionnée au fragment β11, tandis que le vecteur p415-ADH-Import-β1-10 est un vecteur centromérique permettant une expression plus modérée du fragment β1-10. (B) Schéma de l'organisation des fragments β11 en chapelet. Cela permettra la reconstitution de trois GFP actives en aval de la protéine d'intérêt.

ne pas perturber sa localisation subcellulaire, sa fonction, mais également la viabilité de l'organisme. Dans l'optique de visualiser la relocalisation des protéines à l'aide de l'outil Split-GFP, le taux de relocalisation est également une notion importante à prendre en compte car la quantité de la protéine d'intérêt présente dans le compartiment additionnel est généralement plus faible que celle dans le compartiment principal. En cas de faible relocalisation de la protéine d'intérêt, il est fort possible que la reconstitution de la GFP active ne soit pas suffisante pour la visualisation d'un signal fluorescent. Pour compenser ce faible taux de relocalisation, nous avons décidé d'exprimer la cassette permettant l'expression de la protéine d'intérêt étiquetée $\beta 11$ à partir d'un vecteur 2μ , présent en grande quantité dans la cellule et induisant une expression importante de la protéine d'intérêt (p424-GPD-Gène X- $\beta 11$) (**Figure 28A**). Le fragment $\beta 1-10$ est, quant à lui, exprimé en fusion avec divers signaux d'import spécifiques à partir d'un vecteur centromérique, présent à deux exemplaires dans la cellule (p415-ADH-Import- $\beta 1-10$) (**Figure 28A**). L'expression de ce fragment est plus faible que celle du fragment $\beta 11$, mais il est fusionné à des signaux d'adressage favorisant sa relocalisation au niveau du compartiment cible. Le fragment $\beta 1-10$ est donc restreint à un seul compartiment, alors que le fragment $\beta 11$, fusionné à la protéine d'intérêt, est potentiellement multi-localisé. La capacité de relocalisation de la protéine d'intérêt est alors déterminée par microscopie à fluorescence. En effet, c'est uniquement si la protéine d'intérêt est présente dans le compartiment cible, que la reconstitution de la GFP active, *via* autoassemblage des fragments $\beta 1-10$ et $\beta 11$, induit la production d'un signal fluorescent. Pour augmenter la sensibilité de l'outil, nous avons décidé d'utiliser la même stratégie utilisée par Kamiyama et al., 2016 et nous avons exprimé le fragment $\beta 11$ en chapelet (trois copies du $\beta 11$) (Kamiyama et al., 2016). La fusion de plusieurs exemplaires du fragment $\beta 11$ favorise le recrutement du même nombre de fragments $\beta 1-10$ et augmente la capacité de reconstitution de la GFP active, accroissant ainsi le signal émis et facilitant sa détection (**Figure 28B**). Les plasmides 2μ et centromérique comportant des marqueurs de sélection différents (*TRP1* et *LEU2* respectivement), sont alors co-transformés dans une souche de levure et la reconstitution de la GFP active est suivie au microscope confocale ou à épifluorescence.

Finalement, lors de l'ingénierie de l'outil, nous avons également construit des plasmides permettant la fusion du fragment $\beta 11$ en N- ou C-ter de la protéine d'intérêt. L'étude du complexe AME nous a permis de montrer la capacité de relocalisation de la MetRS

et de la GluRS. Cette dernière étant impliquée dans la formation du Gln- m ARNt^{Gln}, elle doit probablement se trouver dans la matrice mitochondriale. Les signaux d'import de ce type de protéine étant généralement retrouvés dans la région N-ter, la fusion du fragment β 11 au niveau de cette extrémité pourrait être problématique pour la relocalisation de la protéine et pourrait induire une localisation cytosolique stricte de la GluRS. En parallèle, la MetRS cytosolique est importée dans le noyau et le signal d'import de cette protéine est également retrouvé dans sa région N-ter. Comme pour la GluRS, la fusion du fragment β 11 à cette région n'est, par conséquent, pas appropriée. De plus, l'interaction entre les trois partenaires du complexe AME est également médiée par leurs régions N-ter respectives. Ainsi, il est possible que la fusion du fragment β 11 à l'extrémité N-ter des aaRS ou d'Arc1, empêche l'assemblage du complexe AME. Comme nous avons pu le voir dans la partie I des résultats, la dérégulation de la relocalisation synchrone des aaRS du complexe peut avoir des conséquences dramatiques pour la levure, surtout en respiration. C'est pour cela que nous avons décidé de fusionner préférentiellement la protéine d'intérêt au fragment β 11 en C-ter. Néanmoins, en cas de problèmes d'expression de certaines protéines, la fusion du fragment β 11 au niveau du N-ter pourrait nous être très utile.

3. Fonctionnalité des protéines étiquetées β 11

Après avoir construit les divers vecteurs d'expression des fragments β 11 et β 1-10, j'ai voulu déterminer si l'expression de la protéine de fusion n'est pas délétère pour l'activité de la protéine, mais également pour la viabilité cellulaire. Pour cela, j'ai mis en œuvre la technique de chasse de plasmide au 5FOA. Ainsi, j'ai transformé les souches *gus1 Δ Shuffle* et *dps1 Δ Shuffle* avec les plasmides d'expression des protéines GluRS- β 11 et AspRS- β 11 respectivement. Dans ces souches, le gène chromosomique codant pour la GluRS (*GUS1*) ou l'AspRS (*DPS1*) est délété et compensé par la présence d'un plasmide portant ces mêmes gènes et le marqueur d'auxotrophie *URA3*. J'ai alors testé la fonctionnalité des protéines de fusion par sélection sur milieu minimum contenant de l'acide 5-fluoro-orotique (5FOA), favorisant la perte du plasmide possédant le marqueur *URA3*. En effet, les souches se développent uniquement si elles sont capables d'éliminer ce vecteur et, *de facto*, si la version étiquetée de l'enzyme est active. J'ai déposé quatre colonies isolées de chaque souche transformée et des témoins, tels que des souches sauvages *ura3* (BY4742, RS453) et la souche *gus1 Δ Shuffle* ou *dps1 Δ Shuffle* transformée avec le vecteur vide. Comme attendu, les deux

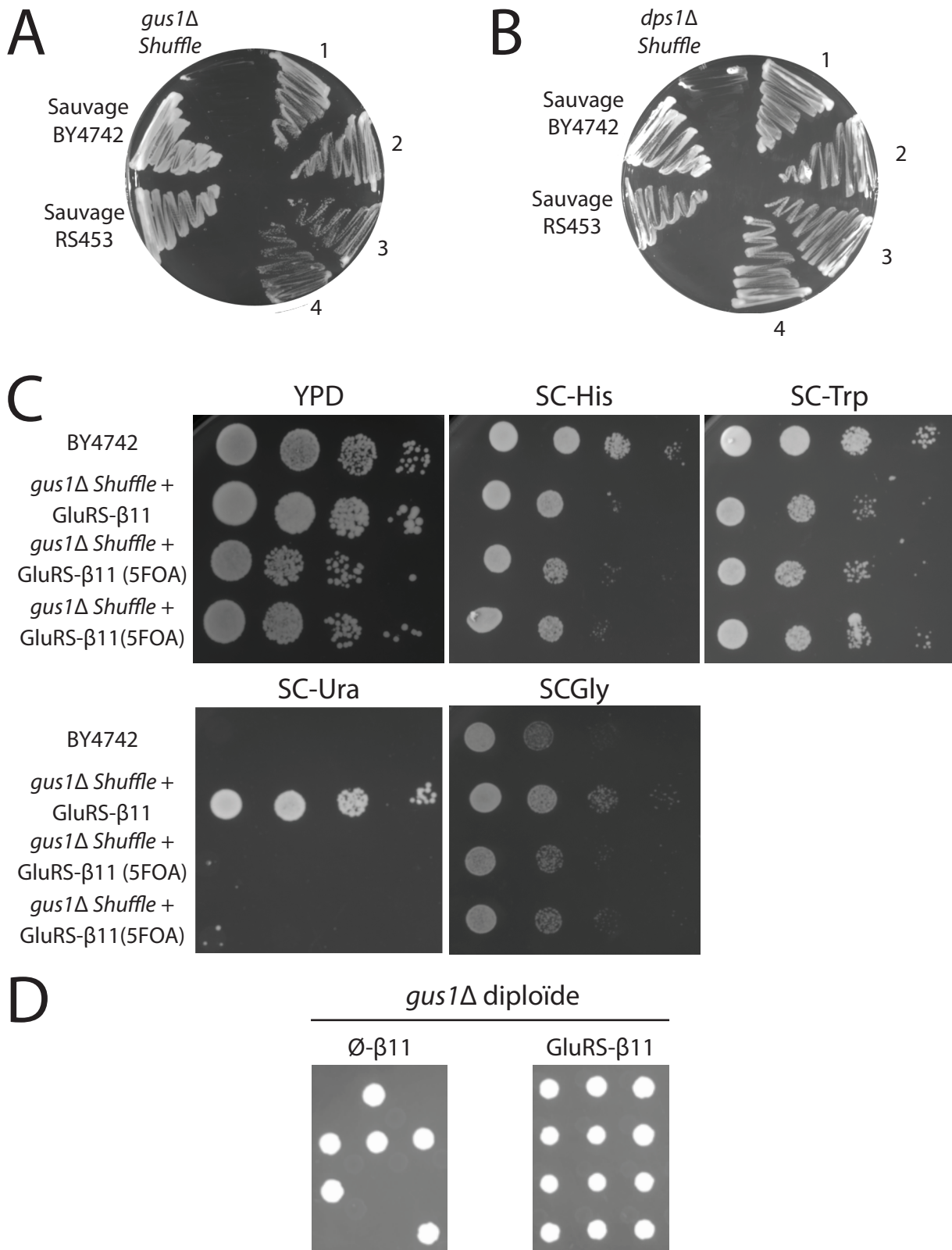


Figure 29 : Fonctionnalité de la GluRS-β11 et de l'AspRS-β11

(A) et (B) Chasse au 5FOA sur des souches *gus1Δ Shuffle* transformées ou non avec le plasmide d'expression de GluRS-β11, ainsi que des souches *dps1Δ Shuffle* transformées ou non avec le plasmide d'expression de l'AspRS-β11. Les souches sauvages sont également déposées (sauvage BY4742 et RS453). (C) Tests en gouttes sur la souche *gus1Δ Shuffle* avant et après chasse au 5FOA. Les souches sont déposées sur milieux fermentaires (YPD, SC-his, SC-Trp, SC-Ura) et respiratoires (SCGly). (D) Dissections de tétrades de souches *gus1Δ* diploïde transformées avec le vecteur vide (∅-β11) ou exprimant la GluRS-β11. Les spores sont déposées sur milieu riche (YPD).

souches *Shuffle* ne se développent pas sur milieu 5FOA contrairement aux souches sauvages qui poussent (**Figure 29A & B**). Une croissance est également observée pour toutes les souches *gus1Δ Shuffle* transformées avec le plasmide d'expression de la GluRS-β11 et toutes les souches *dps1Δ Shuffle* transformées avec le plasmide d'expression de l'AspRS-β11. Ensuite, en collaboration avec Laurence Huck, technicienne au laboratoire, nous avons réalisé un test en gouttes pour vérifier la perte du plasmide *URA3* et la capacité proliférative des souches. Uniquement les résultats pour la chasse de la souche *gus1Δ Shuffle* transformées avec le plasmide d'expression de la GluRS-β11 sont montrés. De plus, les résultats sont identiques pour la souche *dps1Δ shuffle* transformées avec le plasmide d'expression de l'AspRS-β11. Ainsi, nous avons déposé une souche sauvage *TRP1* (BY4742), la souche *gus1Δ Shuffle* avant chasse du plasmide et deux colonies obtenues après chasse. Les résultats indiquent que toutes les souches sont capables de se développer sur milieu riche (YPD) et que la souche avant chasse, possédant le plasmide *URA3*, se développe sur SC-Ura, tandis que les souches après chasse n'en sont plus capables (**Figure 29C**). En parallèle, toutes les souches transformées se développent sur SC-Trp (le marqueur d'auxotrophie du vecteur étant *TRP1*). Tous ces résultats confirment qu'il y a expression des protéines de fusion GluRS-β11 et AspRS-β11 et que la fusion avec le petit fragment de la Split-GFP n'impacte pas la fonction principale de l'enzyme à savoir l'aminocyclation des ARNt homologues. De plus, les souches *gus1Δ Shuffle*, transformées avec le plasmide d'expression de la GluRS-β11, sont capables de se développer en respiration (SCGly) après chasse du plasmide, validant la capacité de la GluRS-β11 à assurer la fonction mitochondriale de la GluRS sauvage (**Figure 29C**). Finalement, comme la fonction mitochondriale est assurée, ceci suggère que la fusion avec le fragment β11 n'empêche pas la relocalisation organellaire de la GluRS-β11.

En parallèle, j'ai également étudié l'activité de la GluRS-β11 par une technique de génétique de levure basée sur l'utilisation d'une souche *gus1Δ* diploïde. Au sein de cette souche diploïde, le gène *GUS1* est représenté par deux allèles, un sauvage (*GUS1*) et un mutant (*gus1*). J'ai testé la fonctionnalité de la protéine de fusion par transformation avec le plasmide vide (\emptyset) ou le plasmide d'expression de la GluRS-β11. La souche utilisée pour créer la banque de diploïde n'étant pas mutée pour *TRP1*, j'ai sous-cloné la cassette d'expression dans un plasmide 2 μ possédant le marqueur *LEU2* (p425-GPD- \emptyset -β11 et p425-GPD-GluRS-β11). Ces souches transformées sont alors mises à sporuler et après dissection, deux spores

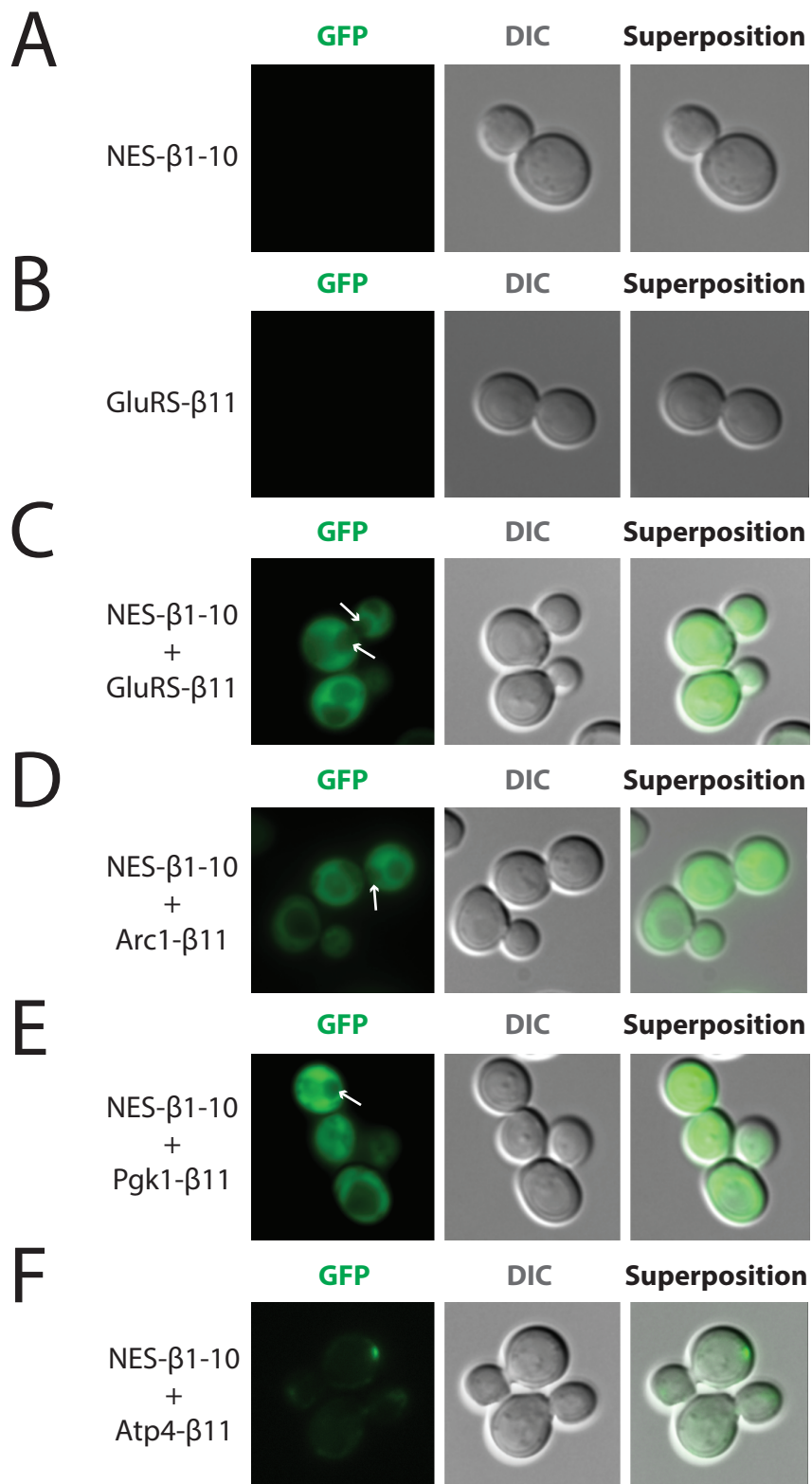


Figure 30 : Autoassemblage cytosolique des fragments β 1-10 et β 11

Une souche sauvage BY4742 est transformée avec les plasmides d'expression de NES- β 1-10 (A) ou GluRS- β 11 (B) ou co-transformée avec les plasmides d'expression de NES- β 1-10 et GluRS- β 11 (C), Arc1- β 11 (D), Pgk1- β 11 (E) ou Atp4- β 11 (F). L'apparition d'un signal est observée par microscopie à épifluorescence. Les flèches blanches indiquent la position du noyau.

possèdent l'allèle sauvage de *GUS1* et deux, l'allèle muté. Uniquement deux spores sont viables (ségrégation 2/2) après dissection de la souche exprimant le plasmide vide, confirmant bien l'incapacité des spores portant l'allèle muté (*gus1*) de se développer (**Figure 29D**). De manière intéressante, lorsque la souche exprime la GluRS- β 11, la ségrégation est différente : 4/0 (**Figure 29D**). Finalement, ces résultats confirment que la GluRS étiquetée β 11 est fonctionnelle et capable d'aminoacyler l' c ARNt^{Glu}. De plus, ces spores sont également capables de se développer en respiration (résultats non montrés), validant également la relocalisation mitochondriale de la protéine de fusion et sa capacité à aminoacyler l' m ARNt^{Gln}.

4. Autoassemblage des fragments β 1-10 et β 11

L'expression de la protéine d'intérêt étiquetée β 11 étant suffisante pour compenser la fonction de la protéine endogène, nous avons voulu tester si la co-expression des deux fragments de la Split-GFP dans le même compartiment favorise la reconstitution de la GFP active. Pour débiter, nous avons décidé d'étudier la reconstitution cytosolique de la GFP. Pour ce faire, nous avons utilisé un signal d'export nucléaire (NES : *Nuclear Export Signal*) fusionné au fragment β 1-10. Ce fragment étant théoriquement restreint au cytosol, le fragment β 11 a été fusionné à des protéines présentes dans ce compartiment (GluRS, Pgc1 et Arc1). Avant de tester la capacité d'autoassemblage des deux fragments, nous avons, tout d'abord, confirmé que leur séparation empêche l'émission de fluorescence. Pour cela, une souche sauvage BY4742 a été transformée, soit avec le plasmide d'expression de la GluRS- β 11, soit avec le plasmide d'expression de NES- β 1-10. L'expression indépendante de la protéine de fusion GluRS- β 11 ou de NES- β 1-10 n'induit pas la production d'un signal de fluorescence (**Figure 30A & B**). Conformément aux indications de la littérature, la séparation des deux fragments inhibe donc bien la formation du chromophore et l'émission de fluorescence. Par la suite, nous avons testé la capacité d'autoassemblage des deux fragments par co-transformation des deux plasmides d'expression (i) de NES- β 1-10 (restreint au cytosol) et (ii) d'une des protéines cytosoliques d'intérêts fusionnées au fragment β 11 (GluRS, Pgc1 ou Arc1). Les deux fragments (β 1-10 et β 11) sont présents dans le même compartiment et s'ils sont capables de s'autoassembler, il y aura reconstruction de la GFP active et production d'un signal fluorescent. De manière intéressante, un signal fluorescent est visualisable au niveau du cytosol pour les trois souches co-exprimant les deux fragments de la Split-GFP (**Figure 30C, D & E**). De plus, il n'y a pas de fluorescence nucléaire confirmant bien l'efficacité du NES utilisé

(**Figure 30, flèches blanches**). Les deux fragments sont capables de reconstituer la molécule active de façon autonome lorsqu'ils sont retrouvés dans le même compartiment. Lorsque la protéine n'est pas cytosolique comme, par exemple, la protéine mitochondriale Atp4, le NES étant libre dans le cytosol, le fragment β 1-10 est capable de s'associer avec le fragment β 11, produisant ainsi un signal fluorescent très faible (**Figure 30F**). Lors de l'ingénierie de l'outil, nous avons ajouté différentes mutations permettant d'augmenter la solubilité et la capacité d'autoassemblage des deux fragments de la Split-GFP. N'ayant aucune idée de l'affinité entre les deux fragments, nous supposons que la présence d'un signal fluorescent est due à l'autoassemblage des fragments NES- β 1-10 et Atp4- β 11, avant import de ce dernier au sein de la mitochondrie. Il serait alors intéressant de déterminer si la reconstitution de la GFP bloque l'import de la pré-protéine au niveau de la membrane de ce compartiment.

En perspective de ces résultats, il serait nécessaire de valider cette hypothèse en purifiant les mitochondries de souches (i) co-exprimant les fragments NES- β 1-10 et Atp4- β 11 ou (ii) exprimant uniquement un des deux fragments. L'immuno-détection des protéines d'intérêts avec un @-GFP validerait la localisation de la GFP active au niveau de la membrane mitochondriale externe (MME) et son absence dans la matrice de ce compartiment. En parallèle, il serait également intéressant de déterminer l'affinité de ces deux fragments. Pour cela, à partir de souches bactériennes surexprimant chacun des deux fragments de la Split-GFP, nous pourrions purifier les deux partenaires et utiliser la technique de Biacore, mais également des techniques de microcalorimétrie (ITC : *Isothermal Titration Calorimetry*) pour mesurer la constante de dissociation (K_D) des fragments. En effet, nous pourrions alors étudier la forte affinité entre les fragments, médiée par l'ajout des mutations de la littérature, et valider la forte capacité d'autoassemblage des fragments β 1-10 et β 11.

5. Autoassemblage mitochondrial des fragments β 1-10 et β 11

Après avoir démontré que la Split-GFP se reconstitue de manière efficace chez la levure, nous avons décidé de modifier le signal d'adressage car la localisation cytosolique des protéines est facile à étudier par simple fusion GFP. L'étude du complexe AME de levure a montré la présence de la GluRS à la fois dans le cytosol et la mitochondrie. De plus, le génome nucléaire ne codant pas toutes les aaRS mitochondriales (Bader et al., 2017), certaines des aaRS cytosoliques sont probablement importées dans ce compartiment pour compenser

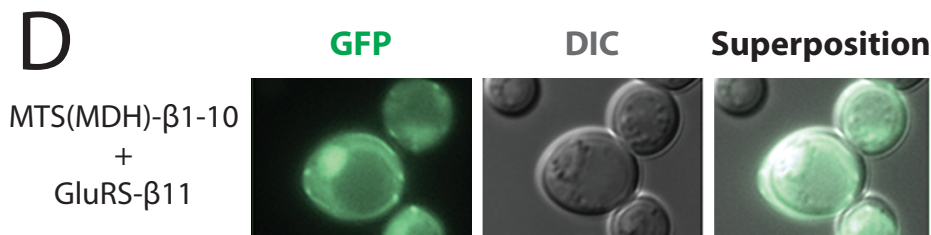
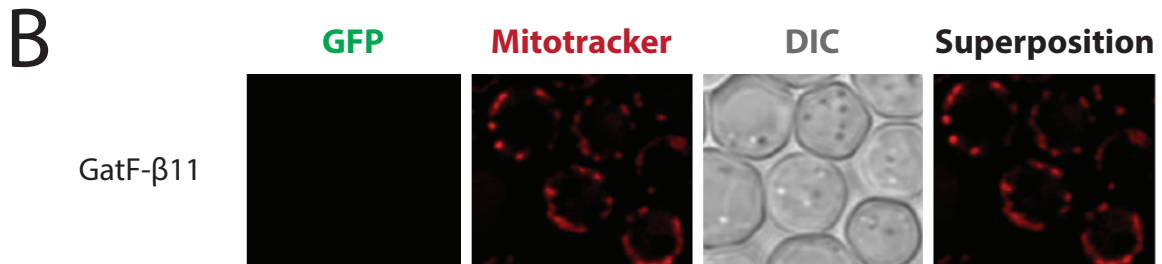
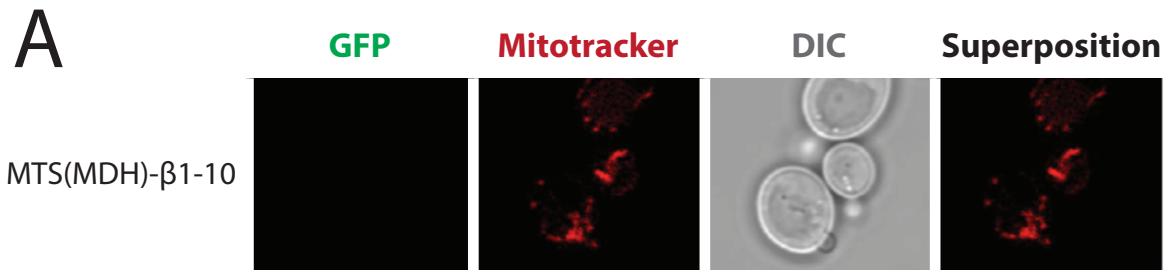


Figure 31 : Expression d'un fragment β 1-10 importé dans la mitochondrie par le MTS de la MDH

Les souches exprimant la protéine MTS(MDH)- β 1-10 (A), GatF- β 11 (B) et co-exprimant les protéines MTS(MDH)- β 1-10 et GatF- β 11 (C) sont observées par microscopie confocale et les mitochondries sont marquées au Mitotracker. (D) La souche co-exprimant MTS(MDH)- β 1-10 et GluRS- β 11 est observée par microscopie à épifluorescence.

l'absence de l'enzyme organellaire. Nous avons décidé de modifier l'outil afin de détecter les protéines se relocalisant dans les mitochondries. Pour ce faire, nous avons sous-cloné la séquence codant le MTS de la malate déshydrogénase (MDH) en amont de celle codant le fragment β 1-10. Par conséquent, la protéine de fusion résultante se relocalisera constitutivement dans la matrice mitochondriale. Ensuite, le plasmide exprimant MTS(MDH)- β 1-10, ainsi que les plasmides permettant d'exprimer les protéines d'intérêts étiquetées β 11 (GatF- β 11, GluRS- β 11) ont été transformés dans une souche de levure sauvage BY4742. Dans un premier temps, nous avons cultivé les souches en respiration et visualisé la production de fluorescence par microscopie confocale, afin d'augmenter le signal et surtout de réduire le bruit de fond issu de la visualisation de plusieurs tranches focales. Pour faciliter la détection d'une potentielle fluorescence organellaire, nous avons également effectué un marquage des mitochondries au Mitotracker. Comme lors de l'étude de l'autoassemblage des deux fragments, nous avons confirmé que les fragments exprimés individuellement ne produisent pas de signal de fluorescence (**Figure 31A & B**). De manière intéressante, un signal fluorescent ponctué est visualisable pour la souche co-exprimant GatF- β 11 et MTS(MDH)- β 1-10. La colocalisation entre le signal de la GFP reconstituée et celui du Mitotracker confirme bien que le signal est présent au niveau de la fraction mitochondriale et que les protéines MTS(MDH)- β 1-10 et GatF- β 11 sont importées dans ce compartiment (**Figure 31C**).

L'objectif de l'outil que nous avons conçu est d'éliminer la fluorescence cytosolique issue d'une protéine doublement localisée. Nous avons voulu savoir si l'outil tel que nous l'avons développé, le permet. Pour ce faire, nous avons étudié la reconstitution de la GFP par microscopie à épifluorescence dans une souche co-exprimant les protéines MTS(MDH)- β 1-10 et GluRS- β 11 ; cette dernière est doublement localisée dans le cytosol et la mitochondrie. De manière surprenante, un signal de fluorescence est visualisable au niveau du cytosol, dans le noyau, mais également sous forme de points correspondant probablement aux mitochondries (**Figure 31D**). La protéine MTS(MDH)- β 1-10 étant de petite taille (30 kDa), elle est capable de diffuser dans le noyau. La fraction nucléaire de cette protéine de fusion favorise probablement la reconstitution de la GFP par autoassemblage avec la GluRS- β 11 présente dans ce compartiment. En effet, cette aaRS possède des signaux d'adressage nucléaires putatifs induisant potentiellement sa présence dans ce compartiment (Bader et al., 2017). La GluRS étant également présente dans le cytosol, le signal dans ce compartiment est également

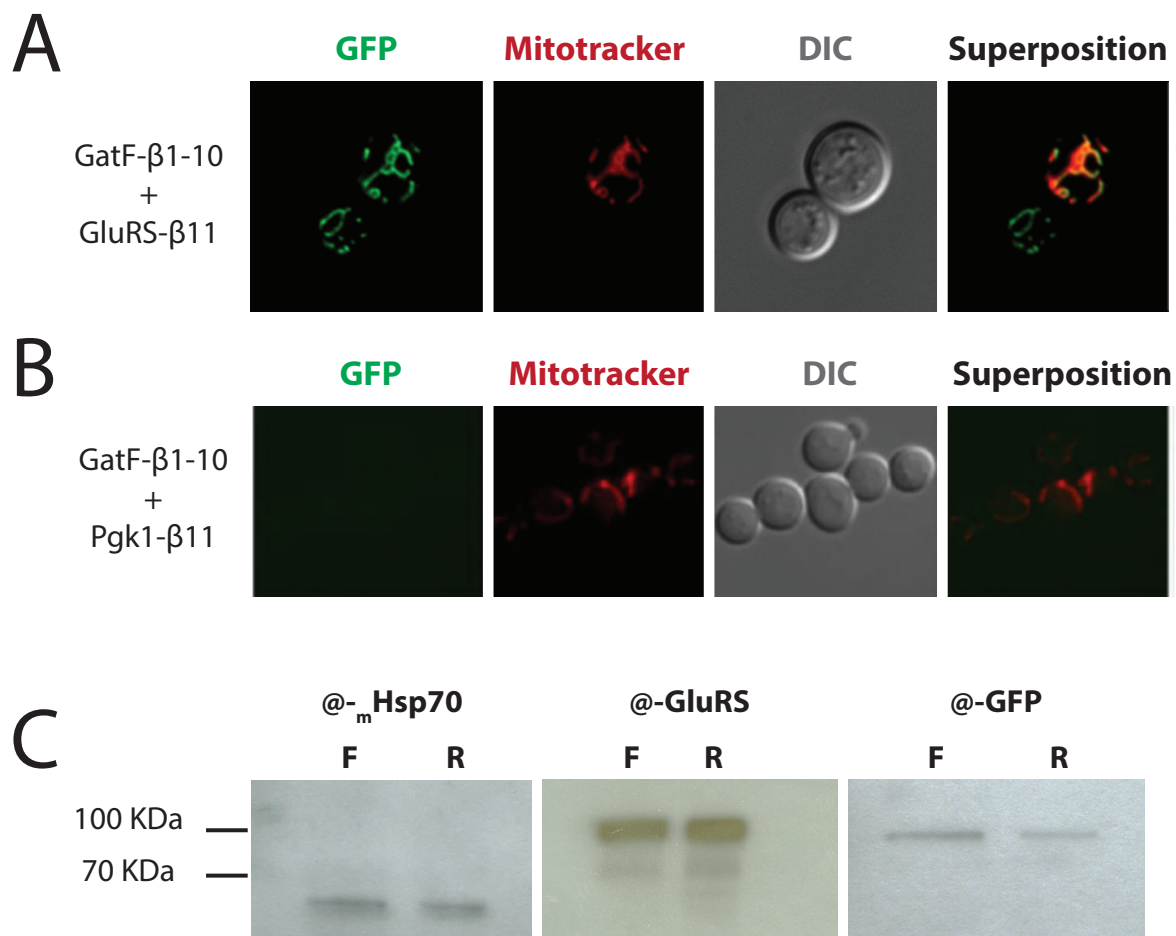


Figure 32 : Utilisation de GatF pour l'import mitochondrial du fragment β 1-10

Les souches transformées avec les plasmides d'expression de GatF- β 1-10 et GluRS- β 11 (A) ou Pgk1- β 11 (B) sont visualisées par microscopie à épifluorescence. (C) La présence des protéines d'intérêts est étudiée sur des extraits totaux mitochondriaux issus d'une souche exprimant GatF- β 1-10 et GluRS- β 11 par immuno-détection et utilisation de l'@-GluRS et -GFP. La présence des mitochondries est révélée par l'anticorps @-_mHsp70. Les extraits ont été réalisés à partir des différentes souches cultivées en fermentation (F) ou respiration (R). La MME a été éliminée par traitement à la protéinase K.

cohérent. L'apparition d'un signal fluorescent dans l'ensemble de la cellule est témoin de la relocalisation partielle du fragment MTS(MDH)- β 1-10 dans la mitochondrie ou de sa rétention dans le cytosol ou le noyau, lors de sa liaison aux fractions cytosolique et nucléaire de la GluRS- β 11. En résumé, deux scénarios sont possibles : (i) l'efficacité d'import médiée par le MTS de la MDH n'est pas optimale ou (ii) la force d'interaction entre les fragments β 1-10 et β 11 est telle, qu'elle conduit à la rétention de la GFP autoassemblée par les fractions nucléaire et cytosolique de la GluRS étiquetée. Cependant, il est également possible que ce soit la combinaison de ces deux hypothèses qui induise ce type de fluorescence. Le signal n'est cependant pas comparable à la fusion de la GFP entière à la GluRS. Effectivement, malgré un signal cytosolique, les localisations nucléaire et mitochondriale (à confirmer par co-marquage) de cette protéine sont révélées. Ces résultats suggèrent néanmoins que l'utilisation du signal d'import de la MDH n'est pas suffisante pour éliminer la fluorescence issue de la fraction cytosolique de la GluRS.

Pour remédier à ces problèmes, nous avons donc décidé de changer le signal d'import et nous avons opté pour la fusion de la protéine GatF au fragment β 1-10. Cette protéine est de petite taille (20 kDa) et primordiale pour la mitochondrie car elle participe à la formation du Gln_mARNt^{Gln}, indispensable à la synthèse protéique organellaire (Frechin et al., 2009b). L'import de cette protéine dans la mitochondrie doit être constitutif car en son absence, la souche n'est plus capable de croître en métabolisme respiratoire (Araiso et al., 2014). Pour déterminer si l'utilisation de GatF favorise l'élimination de la fluorescence cytosolique issue de la GluRS, nous avons transformé la souche exprimant le fragment GluRS- β 11 avec le plasmide d'expression de GatF- β 1-10. La fluorescence a ensuite été visualisée par microscopie à épifluorescence. De manière intéressante, un signal de fluorescence sous forme de réseau est observable. Celui-ci est parfaitement colocalisé avec le marquage des mitochondries au Mitotracker, confirmant la localisation organellaire de la fluorescence (**Figure 32A**). De plus, il n'y a pas de fluorescence dans l'ensemble de la cellule, la GFP se reconstitue donc uniquement au niveau des mitochondries. En parallèle, lorsque la protéine Pgk1- β 11 est exprimée avec GatF- β 1-10, uniquement un faible signal fluorescent est détectable (**Figure 32B**). La présence d'un signal résiduel indique qu'une petite proportion de GatF- β 1-10 s'associe au fragment β 11 avant import ou que l'import de ce fragment n'est que partiel et induit la reconstitution de la GFP active au niveau de la MME. Cependant, ces résultats confirment que GatF se relocalise

totalemment au niveau de la mitochondrie et que le signal fluorescent se trouve bien au niveau de ce compartiment. Enfin, ces expériences valident l'utilisation de GatF comme l'équivalent d'un MTS très efficace pour l'import d'une protéine de fusion dans la mitochondrie et l'outil Split-GFP pour l'étude de la localisation mitochondriale des protéines doublement localisées. En effet, la fluorescence issue de la fraction cytosolique de la GluRS est totalement éliminée.

Pour confirmer que la GFP reconstituée se situe bien au sein des mitochondries, j'ai réalisé des purifications de ces organelles à partir de la souche co-exprimant la GluRS- β 11 et GatF- β 1-10 et j'ai digéré la MME avec la protéinase K. J'ai alors réalisé des immuno-détections sur des extraits obtenus à partir de ces mitoplastes pour détecter la présence des deux protéines de fusion. La détection avec l'anticorps @-mHsp70 confirme bien que la fraction récupérée contient des mitochondries (**Figure 32C**). La GluRS- β 11 est également détectée dans ces extraits mitochondriaux (@-GluRS) (**Figure 32C**). De manière très intéressante, l'anticorps dirigé contre la GFP (@-GFP) détecte la protéine GluRS- β 11 (**Figure 32C**), mais ne détecte pas la protéine GatF- β 1-10. Cet anticorps reconnaît donc uniquement le polypeptide de seize aa formant le fragment β 11 de la GFP (présent en trois copies). La reconnaissance est néanmoins beaucoup plus faible que celle par l'anticorps @-GluRS, mais pourrait être prometteuse pour l'utilisation du chapelet β 11 pour la purification de la protéine d'intérêt ou l'étude des interactants par co-immunoprécipitation (CoIP) à partir d'extraits mitochondriaux. Finalement, les purifications de mitochondries ont été réalisées sur des cellules cultivées en fermentation et respiration et je n'ai pas détecté de différence significative dans la quantité de GluRS cytosolique importée. Cependant, cela ne veut pas dire qu'il n'y a pas de différence dans le contenu mitochondrial en GluRS- β 11 en fonction de métabolisme utilisé. En effet, les plasmides codant GatF- β 1-10 et GluRS- β 11 induisent la surexpression des deux fragments. L'assemblage du complexe AME étant stœchiométrique, Arc1 n'est pas capable de prendre en charge la quantité de GluRS- β 11 supplémentaire et cette fraction libre dans le cytosol se relocalise constitutivement dans la mitochondrie. Ceci explique sa présence mitochondriale en grande quantité dans les deux conditions testées. Le signal fluorescent et la quantité de GluRS- β 11 dans la mitochondrie étant importante, une expression moindre des protéines étiquetées β 11 pourrait être suffisante pour visualiser la reconstitution de la GFP. J'ai décidé

de sous-cloner les différentes constructions dans un plasmide centromérique afin de réduire l'expression des protéines étiquetées $\beta 11$ (p414-GPD-Gène X- $\beta 11$).

Après quelques modifications de l'outil, nous avons réussi à élaborer un outil de Split-GFP pour l'étude des protéines mitochondriales. Celui-ci est spécifique, car il permet l'élimination de la fluorescence issue de la fraction cytosolique d'une protéine doublement localisée (GluRS). De manière intéressante, l'anticorps @-GFP reconnaît le fragment $\beta 11$, suggérant sa possible utilisation pour des études d'interactomique ou de purification des protéines d'intérêts par CoIP. De plus, ces résultats suggèrent un import fort de la GluRS- $\beta 11$ au niveau de la mitochondrie. A court terme, nous pourrions donc tenter de diminuer la quantité de GluRS cytosolique se relocalisant dans ce compartiment. Pour cela, nous envisageons de transformer la souche co-exprimant les deux fragments de la Split-GFP (GatF- $\beta 1-10$ et GluRS- $\beta 11$) avec un plasmide de surexpression d'*Arc1* (p423-GPD-*ARC1* par exemple) et déterminer, par immuno-détection sur des mitochondries purifiées, si la relocalisation mitochondriale de la GluRS cytosolique est plus faible. En parallèle, la version sauvage de la GluRS pourrait également être remplacée par la version étiquetée $\beta 11$. Nous avons déjà montré que cette dernière est active et nous pourrions générer une souche *gus1* Δ exprimant la GluRS- $\beta 11$ et GatF- $\beta 1-10$. En absence de la GluRS sauvage (endogène), la rétention cytosolique de la GluRS- $\beta 11$ devrait être plus importante et la quantité de GluRS- $\beta 11$ importée dans la mitochondrie plus faible, induisant probablement une diminution de l'intensité de fluorescence émise. Dans cette même souche, la surexpression d'*Arc1* réduirait drastiquement la quantité de GluRS cytosolique importée dans la mitochondrie et une croissance respiratoire lente devrait être observable.

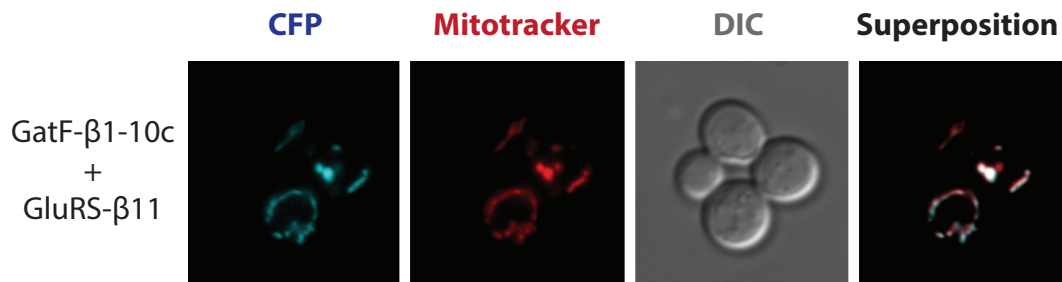
6. Reconstitution mitochondriale d'une Split-CFP et -YFP

La création de la palette de variants de la GFP a été une avancée cruciale pour la détection simultanée de plusieurs protéines d'intérêts. En effet, lorsque deux protéines sont fusionnées à des molécules fluorescentes différentes, il est possible d'étudier la localisation de ces deux protéines, ainsi que leur colocalisation. Plusieurs protéines fluorescentes sont dérivées de la même protéine (GFP) et ne diffèrent que de quelques résidus essentiels permettant de moduler les spectres d'émission et d'absorption de la molécule fluorescente. C'est notamment le cas de la CFP et de l'YFP, dont les mutations sont décrites dans

A

	65	66	67	Mutations
Split-GFP	Thr	Tyr	Gly	
Split-YFP	Gly	Tyr	Gly	Tyr203, Pro64 Leu68, Ala72
Split-CFP	Thr	Trp	Gly	Thr146, Thr153 Ala163

B



C



Figure 33 : Reconstitution mitochondriale d'une Split-CFP et -YFP

Les mutations permettant de modifier les propriétés spectrales de la Split-GFP sont toutes retrouvées dans le fragment β 1-10 (A). La GluRS- β 11 est co-exprimée avec les protéines GatF- β 1-10c (B) ou GatF- β 1-10y (C) et la reconstitution de la CFP et de l'YFP est observée par microscopie à épifluorescence.

l'introduction § VIII.3.b et VIII.3.c et dans la **Figure 33A**. Ces mutations peuvent également être transposées à l'outil de visualisation Split-GFP pour former notamment la Split-CFP ou -YFP. Néanmoins, l'utilisation de ces variants pour étudier la localisation subcellulaire des protéines n'a pas encore été décrite. De manière intéressante, les mutations générant la CFP ou l'YFP sont toutes retrouvées dans le fragment β 1-10. Cela signifie que la reconstitution d'une Split-GFP, Split-CFP ou Split-YFP est possible, sans que nous ayons à modifier la séquence codant le fragment β 11. J'ai réalisé les mutations dans le fragment β 1-10 fusionné à GatF pour obtenir les plasmides d'expression GatF- β 1-10c (cyan) et GatF- β 1-10y (Yellow). Pour tester la capacité de reconstitution de ces variants, mais également pour étudier si l'intensité de fluorescence est suffisante pour être détectée, j'ai co-transformé une souche sauvage BY4742 avec les plasmides d'expression de la GluRS- β 11 et de GatF- β 1-10c ou GatF- β 1-10y. Les images de microscopie à épifluorescence confirment la capacité de la Split-CFP et YFP à s'autoassembler pour produire un signal fluorescent cyan et jaune qui colocalise avec la fraction mitochondriale dévoilée par le marquage au Mitotracker (**Figure 33B & C**). Par conséquent, la modification du fragment β 1-10 par l'ajout des différentes mutations est suffisante pour moduler les spectres d'émission et d'absorption de la molécule fluorescente reconstituée.

7. Autoassemblage nucléaire des fragments β 1-10 et β 11

Le noyau est un compartiment crucial pour l'expression des gènes et la régulation de divers processus vitaux pour la cellule. L'import de protéines dans ce compartiment fait intervenir des signaux spécifiques, pouvant être retrouvés tout au long de la séquence protéique, nommés les *Nuclear Localization Signal* (NLS). Ceux-ci sont pris en charge par des transporteurs spécifiques favorisant l'entrée des protéines dans ce compartiment *via* un transport actif. La relocalisation nucléaire est un phénomène abondant chez les eucaryotes, surtout lors de stress ou de dommages, afin de réguler l'expression des gènes ou favoriser la réparation de l'information génétique lésée (Gorner et al., 1998; Sugiyama et al., 2014). De plus, de nombreuses études suggèrent que certaines aaRS se relocalisent dans le noyau pour assurer des fonctions régulatrices essentielles, mais également pour aminoacyler leurs ARNt homologues, favorisant ainsi leur maturation ou leur export du noyau (Azad et al., 2001; Ko et al., 2000). Durant ma thèse, nous avons démontré que la MetRS se relocalise dans le noyau et

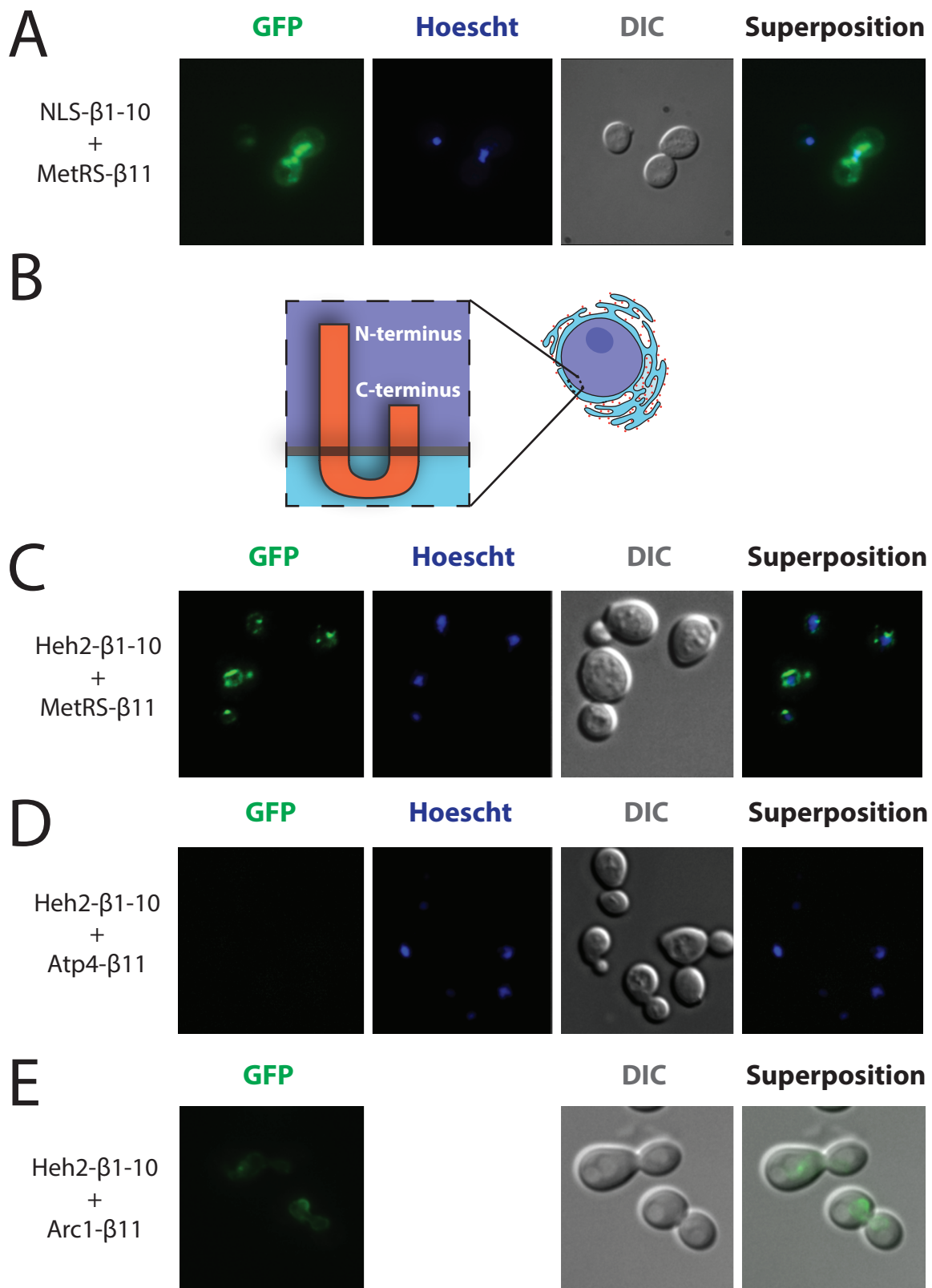


Figure 34 : Expression d'un fragment β 1-10 importé dans le noyau

La souche sauvage BY4742 est co-transformée avec les plasmides d'expression du NLS du SV40 fusionné au fragment β 1-10 (NLS- β 1-10) et de la MetRS- β 11 (A). Le NLS a été modifié et la protéine Heh2 a été sélectionnée pour le remplacer (B). Le gène codant Heh2- β 1-10 est intégré au niveau du locus *HIS3* et cette souche est transformée avec les plasmides d'expression de la MetRS- β 11 (C), d'Atp4- β 11 (D) ou d'Arc1- β 11 (E). La fluorescence est observée par microscopie à épifluorescence et la fraction nucléaire est révélée par marquage au Hoechst.

constaté les difficultés de l'étude de la localisation nucléaire de cette protéine (Frechin et al., 2014).

Pour faciliter la détection de protéines multi-localisées se relocalisant dans le noyau, j'ai donc décidé, sur le même principe, de transposer l'outil de Split-GFP à l'étude de ces protéines "nucléo-cytosoliques" et j'ai fusionné le NLS du virus SV40 (Kalderon et al., 1984) (couramment utilisé en biologie moléculaire) au fragment β 1-10 pour le restreindre à ce compartiment. J'ai alors co-transformé une souche sauvage BY4742 avec les plasmides permettant l'expression des protéines de fusion NLS- β 1-10 et MetRS- β 11. Pour rappel, à partir du moment où je me suis lancé dans l'étude des protéines nucléo-cytosoliques, j'ai systématiquement utilisé un vecteur centromérique (p414-GPD-Gène X- β 11) pour exprimer les protéines d'intérêts fusionnées au fragment β 11. Au sein de la souche co-exprimant les deux protéines de fusion, un signal fluorescent vert est détectable. Celui-ci colocalise partiellement avec le marquage du noyau au Hoechst, suggérant une reconstitution nucléaire de la GFP active et confirmant la présence de la MetRS dans ce compartiment. Néanmoins, un signal faible et diffus est également présent dans le cytosol (**Figure 34A**). La reconstitution de la GFP est probablement nucléaire et la présence d'un signal au niveau du cytosol peut être due : (i) à un import partiel de NLS- β 1-10 dans ce compartiment, induisant son autoassemblage avec la fraction cytosolique de la MetRS- β 11 ou (ii) à l'export d'une fraction de la MetRS-GFP reconstituée et nucléaire vers le cytosol.

Pour remédier à ce problème, il était nécessaire de restreindre le fragment β 1-10 au seul compartiment nucléaire. Pour ce faire, j'ai cherché une protéine ancrée dans la membrane interne du noyau. J'ai trouvé plusieurs candidats potentiels et j'ai décidé d'utiliser la protéine Heh2. Cette protéine est intéressante car elle est ancrée dans la membrane interne du noyau et ses extrémités N- et C-ter sont libres dans le nucléoplasme, permettant potentiellement la fusion du fragment β 1-10 au niveau de chacune d'entre elles (**Figure 34B**). J'ai dans un premier temps testé la capacité d'ancrage d'Heh2 par fusion du fragment β 1-10 en C-ter. Pour éviter la transformation avec deux plasmides, j'ai décidé d'exprimer ce fragment à partir d'une version chromosomique de la construction et je l'ai donc cloné dans un vecteur intégratif. Ceci permet, d'une part de s'affranchir de la co-transformation de deux plasmides, et d'autre part, de réduire l'expression de ce fragment. Ainsi, j'ai linéarisé ce vecteur au niveau de son marqueur d'auxotrophie (*HIS3*) pour favoriser son intégration au

niveau de ce locus et ne pas perturber l'homéostasie cellulaire par modification de la copie endogène d'*HEH2*. J'ai ensuite transformé cette souche (nommée *HEH2*- β 1-10) avec le plasmide d'expression de la MetRS- β 11. De manière intéressante, cette souche présente un signal de fluorescence vert sous forme d'anneau correspondant probablement à l'enveloppe nucléaire (**Figure 34C**). La fraction nucléaire est révélée par co-marquage avec le Hoechst, validant la localisation nucléaire de la fluorescence et de ce fait de la MetRS- β 11. De manière intéressante, la fluorescence cytosolique observée avec le NLS classique n'est plus visualisable. De plus, c'est la première fois que le fragment β 1-10 est exprimé à partir du génome. L'expression de ce dernier est plus faible que lors de son expression sur plasmide, mais néanmoins suffisante pour la visualisation d'un signal fluorescent. La spécificité de la fluorescence est également confirmée par l'étude de la localisation d'une protéine mitochondriale *Atp4*. En effet, lorsque cette protéine est exprimée dans la souche *HEH2*- β 1-10, aucune fluorescence n'est émise, confirmant bien que cette protéine mitochondriale n'est pas détournée vers le noyau par autoassemblage des fragments β 1-10 et β 11 (**Figure 34D**). L'apparition d'un signal est également observée pour la protéine *Arc1*- β 11, dont la localisation nucléaire est très importante, mais qui est fortement réexportée de ce compartiment (Galani et al., 2005). Cette protéine se retrouve alors transitoirement dans le noyau et favorise la reconstitution de la GFP active. Notre outil permet de détecter la localisation de cette protéine et la fluorescence est maintenue dans le noyau par ancrage au niveau de la membrane interne de ce compartiment. Par ailleurs, l'interaction entre le fragment β 1-10 et β 11 est stable et forte car la protéine *Arc1*, fortement exportée du noyau, est maintenue dans ce compartiment (**Figure 34E**).

8. Reconstitution de la CFP à la surface de la vacuole

Chez la levure, la vacuole est un compartiment particulier, impliqué dans la dégradation des protéines, mais également le stockage des aa et la détoxification de la cellule (Li et Kane, 2009). Le transport vers la vacuole fait intervenir des endosomes qui fusionnent avec la membrane vacuolaire pour libérer des protéines dans ce compartiment ou à sa surface (Feyder et al., 2015). Récemment, des études ont montré la capacité de la LeuRS de levure de se relocaliser au niveau de la vacuole pour y assurer le rôle de senseur de leucine pour la kinase TOR (Bonfils et al., 2012). La majorité des aa étant présents dans ce compartiment, il ne serait pas étonnant que certaines, voir toutes les aaRS se relocalisent à la surface de la

vacuole pour jouer le rôle de senseurs de leurs aa homologues et de ce fait, moduler la croissance de la levure. De plus, il semblerait également qu'Arc1, la protéine d'ancrage du complexe AME, soit capable de s'associer aux endosomes et à la membrane vacuolaire de par son interaction avec les phosphatidylinositol-3-phosphate et -3,5-bisphosphate (PI3P et PI3,5P₂) (Fernandez-Murray and McMaster, 2006). La relocalisation vacuolaire est probablement indispensable à l'homéostasie cellulaire et l'étude des protéines se relocalisant dans ou à la surface de ce compartiment pourrait favoriser la compréhension des mécanismes de régulation liés à la disponibilité en aa, à la réponse aux stress, mais également à la dégradation des protéines. Pour transposer mon outil à l'étude des protéines vacuolaires, j'ai étroitement collaboré avec le Dr. Sylvie Friant, qui dirige une équipe au sein de notre UMR. En tant que spécialiste du trafic intracellulaire et plus particulièrement de la vacuole, elle m'a alors proposé plusieurs candidats pour l'ancrage du fragment β 1-10 et j'ai décidé d'utiliser la protéine Vph1, généralement utilisée comme témoin de pureté vacuolaire, pour ancrer le fragment β 1-10 dans cette membrane.

Vph1 est une s-u de l'ATPase vacuolaire et surtout c'est une protéine transmembranaire, dont les deux extrémités sont retrouvées libres dans le cytosol. Par conséquent, le fragment β 1-10 pourra être fusionné à ces deux extrémités. Comme l'outil Split-GFP était déjà au point pour l'étude des protéines nucléo-cytoplasmiques, j'ai voulu élaborer une Split-CFP vacuolaire, de façon à avoir la possibilité ultérieurement de combiner les deux systèmes Split-GFP/CFP dans une même souche. Pour cela, j'ai utilisé le fragment β 1-10c possédant les mutations décrites **Figure 33** et je l'ai exprimé en fusion avec Vph1. L'expression génomique d'Heh2- β 1-10 favorise la production d'une quantité suffisante de ce partenaire pour la visualisation d'un signal, j'ai donc sous-cloné cette construction dans un plasmide intégratif. Après linéarisation dans le marqueur de sélection (ici *URA3*), ce plasmide est transformé dans une souche sauvage BY4742 pour induire l'expression génomique de la protéine Vph1- β 1-10c. Cette souche a finalement été transformée avec les plasmides exprimant Arc1- β 11, His3- β 11 et Atp4- β 11. Lors de la co-expression d'Arc1- β 11 et Vph1- β 1-10c, un signal de fluorescence cyan apparaît sous forme d'anneau et il colocalise avec la vacuole observable en DIC. Pour confirmer cette localisation, la fraction vacuolaire a été marquée avec le FM4-64. Ce composant permet le marquage dynamique des membranes et une incubation de 20 min est nécessaire pour visualiser la membrane vacuolaire. La

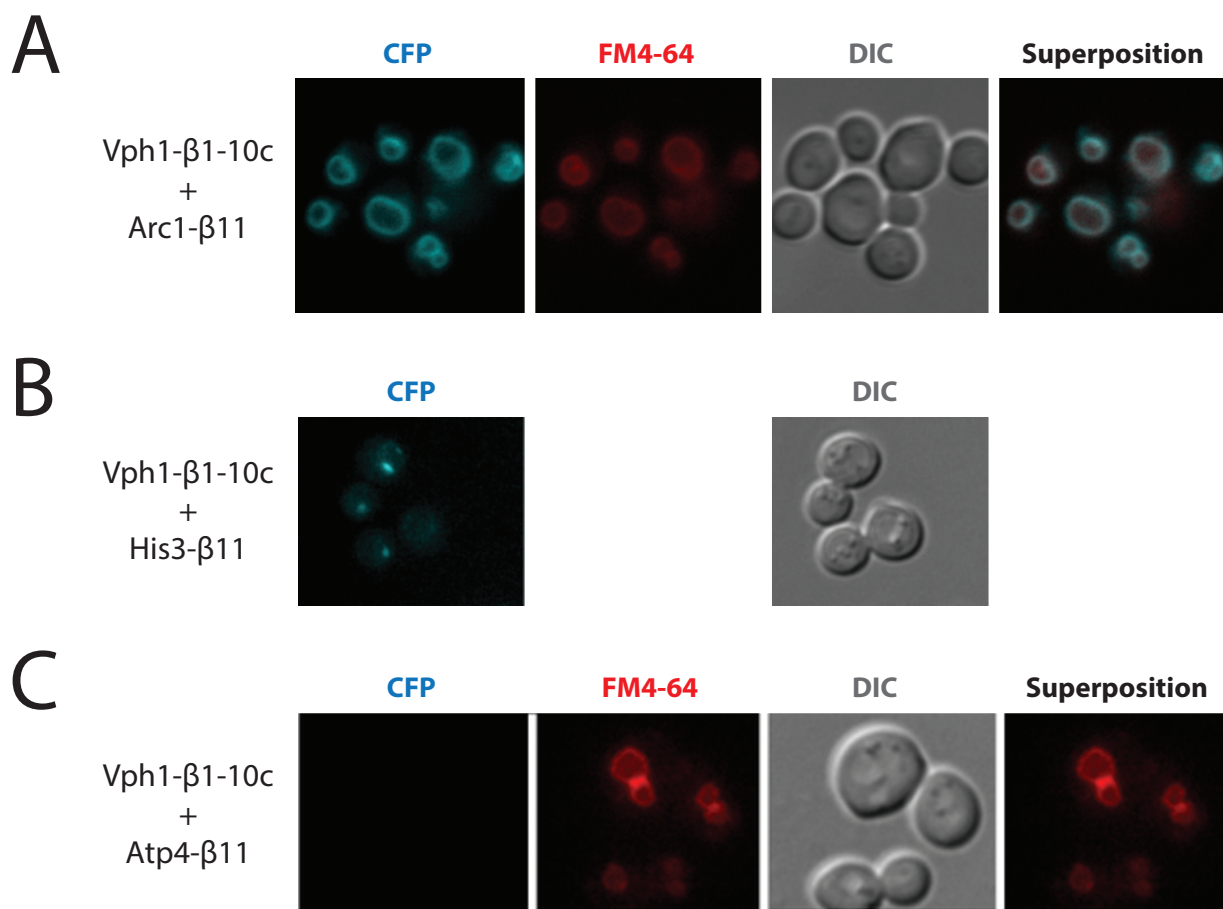


Figure 35 : Expression d'un fragment β 1-10c retrouvé à la surface de la vacuole

Les souches co-exprimant la protéine vacuolaire Vph1- β 1-10c et Arc1- β 11 (A), His3- β 11 (B) ou Atp4- β 11 (C) sont visualisées par microscopie à épifluorescence. La fraction vacuolaire est observée suite à 20 min de marquage au FM4-64.

colocalisation du marquage vacuolaire avec le signal fluorescent cyan confirme bien la localisation vacuolaire de la CFP reconstituée (**Figure 35A**). Comme suggéré dans la littérature, Arc1 est capable de se lier aux fractions membranaires, et plus particulièrement à la vacuole (contenant le PI3,5P₂). La fusion de cette protéine au fragment β 11 induit donc bien le rapprochement des deux fragments et la reconstitution d'une CFP active.

Lors de l'expression d'une protéine cytosolique (His3- β 11), un signal de fluorescence très faible apparaît (**Figure 35B**). Ce signal est néanmoins plus ponctué que le signal pour la protéine Arc1- β 11, suggérant qu'His3 se trouve au niveau de la vacuole, mais que sa relocalisation au niveau de ce compartiment est moins importante et qu'elle forme plutôt des amas (cluster). Il est en effet possible que cette protéine se retrouve transitoirement au niveau de la vacuole et que l'autoassemblage des fragments β 1-10 et β 11 favorise la reconstitution de la CFP active. De plus, le mutant de délétion de cette protéine induit un phénotype de sensibilité à différents métaux comme le cuivre, le cobalt et le nickel (source www.yeastgenome.org/). La vacuole ayant un rôle de détoxification des métaux lourds, cela peut expliquer la localisation vacuolaire d'His3 et l'apparition d'un signal fluorescent ponctué.

Finalement, l'expression de la protéine de fusion Atp4- β 11 n'induit pas l'apparition d'un signal fluorescent, il n'y a donc pas détournement d'Atp4 au niveau de la vacuole par autoassemblage de CFP active au niveau du cytosol (**Figure 35C**). Ainsi, Cet outil de Split-CFP vacuolaire élimine la fluorescence cytosolique d'une protéine multi-localisée (Arc1- β 11). En parallèle, une protéine qui ne se relocalise pas à la surface de ce compartiment n'induit pas la production d'un signal fluorescent (Atp4- β 11). La détection d'un signal fluorescent pour His3- β 11 est plutôt surprenante mais, au vu des données de la littérature, pourrait être explicable de par la fonction de la vacuole dans la détoxification et la sensibilité accrue du mutant de délétion de *his3* à différents métaux. Par conséquent, l'outil peut être utilisé pour détecter la localisation vacuolaire des protéines de levure.

9. Reconstitution nucléaire et vacuolaire de la Split-GFP et -CFP

Après avoir montré la capacité de reconstitution de la Split-GFP nucléaire et de la Split-CFP vacuolaire, nous avons voulu savoir s'il pouvait y avoir reconstitution de la GFP et la CFP au sein de la même cellule. Pour déterminer si l'étude de plusieurs localisations en même

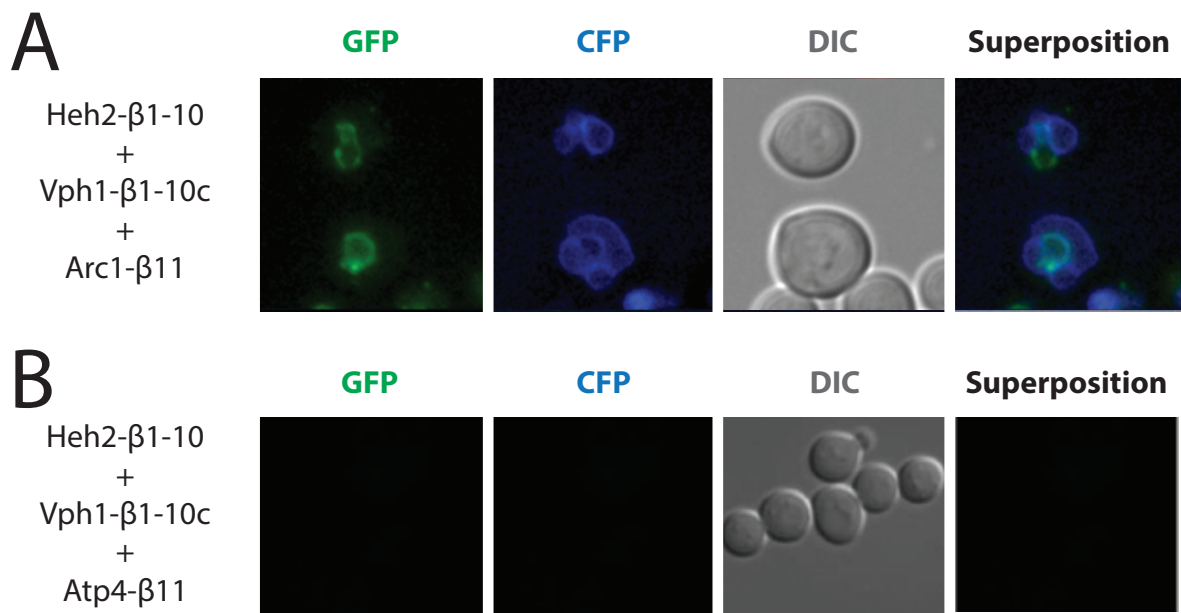


Figure 36 : Co-expression d'un fragment β 1-10 dans le noyau et β 1-10c à la surface de la vacuole

La souche co-exprimant génétiquement Heh2- β 1-10 et Vph1- β 1-10c est transformée avec un plasmide d'expression d'Arc1- β 11 (A) ou d'Atp4- β 11 (B). La reconstitution de la CFP et GFP est observée par microscopie à épifluorescence.

temps est possible avec l'outil de Split-GFP, j'ai utilisé la souche *HEH2-β1-10* dans laquelle j'ai intégré le gène codant *Vph1-β1-10c*. Cette souche exprime deux fragments $\beta1-10$; le premier favorise la reconstitution d'une GFP active dans le noyau et le second permet l'autoassemblage d'une CFP active au niveau de la membrane vacuolaire. Comme je l'ai décrit dans la **Figure 33**, toutes les mutations modulant les spectres d'émission et d'absorption des variants sont retrouvées dans le fragment $\beta1-10$. Par conséquent, afin de déterminer s'il est possible de détecter simultanément la localisation de plusieurs protéines et, *de facto*, la reconstitution de deux molécules fluorescentes, j'ai utilisé une seule et même protéine, *Arc1-β11*. Lors de l'expression d'*Arc1-β11* dans cette souche, il y a apparition d'un signal vert sous forme d'anneau correspondant au noyau et d'un signal cyan correspondant à la vacuole (**Figure 36A**). La relocalisation nucléaire et vacuolaire de la protéine *Arc1-β11* induit donc bien la reconstitution de la GFP et CFP dans le noyau et à la surface de la vacuole respectivement. De plus, de manière intéressante, il n'y a pas colocalisation de ces deux fluorescences, suggérant que ce n'est pas l'activation de la première fluorescence qui stimule la seconde molécule. Finalement, comme précédemment, lorsque la protéine *Atp4-β11* (contrôle négatif) est exprimée dans cette souche, il n'y a pas de signal fluorescent (**Figure 36B**). Ces résultats confirment la capacité de l'outil à produire un signal fluorescent vert et cyan, que la détection de ces fluorescences simultanément révèle la multi-localisation d'*Arc1-β11*, et surtout qu'une protéine ne se relocalisant pas naturellement dans le noyau et à la surface de la vacuole ne favorise pas la reconstitution de la molécule fluorescente active dans ces deux compartiments.

10. Localisation nucléaire et vacuolaire de *NES-β11*

Lors de l'étude de la localisation nucléaire et vacuolaire de la protéine *Atp4-β11*, aucun signal fluorescent n'est observé pour ce contrôle négatif, confirmant la spécificité de l'outil Split-GFP nucléaire et Split-CFP vacuolaire pour la visualisation des protéines cytosoliques pouvant se relocaliser au niveau de ces deux compartiments. Cependant, *Atp4* est une protéine membranaire et aucune évidence ne montre la capacité de l'outil à discriminer une protéine soluble, capable de diffuser au niveau de ces deux compartiments, d'une protéine s'y trouvant constitutivement. Afin de déterminer si une protéine libre dans le cytosol induit la reconstitution de la GFP dans le noyau, ou de la CFP à la surface de la vacuole, j'ai décidé de construire un vecteur permettant l'expression du fragment $\beta11$, fusionné au NES que j'ai utilisé précédemment (*NES-β11*). En effet, comme nous avons pu l'observer dans la partie II.4

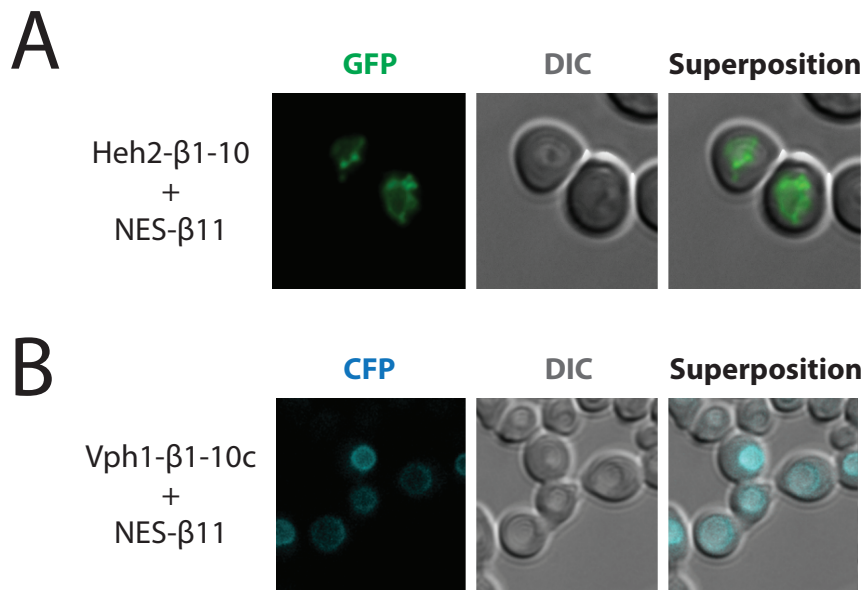


Figure 37 : Expression de NES-β11 dans les systèmes Split-nucléaire et -vacuolaire

La protéine de fusion NES-β11 est exprimée dans la souche exprimant Heh2-β1-10 (A) ou Vph1-β1-10c. L'apparition de fluorescence est visualisée par microscopie à épifluorescence.

des résultats, ce signal d'export induit une localisation cytosolique stricte de la protéine d'intérêt et elle devrait donc uniquement se trouver transitoirement dans le noyau et être absente de la surface de la vacuole.

J'ai, tout d'abord, étudié la capacité de reconstitution nucléaire de la GFP en exprimant NES- β 11 dans la souche *HEH2- β 1-10*. De manière surprenante, un signal nucléaire est observable, ce qui suggère que NES- β 11 est présent dans le noyau et induit la reconstitution de la GFP active à l'intérieur de ce compartiment (**Figure 37A**). Ces résultats sont inattendus, car, en théorie, le peptide NES- β 11 ne devrait pas avoir la capacité de se localiser dans le noyau. Plusieurs possibilités pourraient expliquer ce résultat, notamment la diffusion nucléaire du fragment NES- β 11 de petite taille ou la reconstitution de la protéine Heh2-GFP-NES avant l'import de Heh2- β 1-10 dans le noyau et son intégration dans la membrane interne. Le NES- β 11 ne serait ainsi plus exporté et la fluorescence nucléaire pourrait alors être expliquée.

J'ai ensuite étudié la capacité de reconstitution de la CFP à la surface de la vacuole dans une souche co-exprimant *Vph1- β 1-10c* et NES- β 11. Comme lors de l'étude de la relocalisation nucléaire de NES- β 11, un signal fluorescent est observable (**Figure 37B**). Celui-ci est retrouvé au niveau de la vacuole et confirme qu'une protéine, libre dans le cytosol, peut diffuser vers la membrane vacuolaire et reconstituer la CFP active à la surface de ce compartiment. Néanmoins, l'autre hypothèse qui pourrait expliquer l'apparition d'un signal est la reconstitution de la CFP avant import et ancrage vacuolaire de *Vph1- β 1-10c*.

Ces deux résultats suggèrent que la reconstitution peut avoir lieu avant import du fragment β 1-10 au niveau de compartiment ciblé. Cependant, le peptide NES- β 11 est de petite taille et peut potentiellement diffuser à l'intérieur du noyau ou la surface de la vacuole. Pour éviter que la protéine étiquetée β 11 ne diffuse dans le noyau, il serait nécessaire d'utiliser une protéine cytosolique de taille suffisamment importante pour diminuer ou empêcher sa diffusion à travers des pores nucléaires. Pour obtenir une protéine de ce type et qui serait strictement restreinte au cytosol, nous envisageons de fusionner la DHFR (Dihydrofolate Réductase) au fragment β 11. Cette protéine, faisant une taille de 24 kDa (gène *DFR1*), sera associée concaténée (deux ou trois copies), afin d'augmenter la taille de la protéine de fusion et ainsi, sortir de la limite de diffusion nucléaire. Cette protéine, incapable de se relocaliser

dans ce compartiment, nous permettra de conclure quant à la capacité de reconstitution de la Split-GFP avant l'import nucléaire de Heh2- β 1-10. La protéine (DHFR)_{2/3}- β 11 pourra également être utilisée pour tester la capacité de diffusion à la surface de la vacuole. Ce témoin est donc primordial pour tester la fiabilité de l'outil pour la détection d'une protéine se relocalisant spécifiquement dans le compartiment ciblé par le fragment β 1-10, par opposition à une protéine pouvant diffuser dans le noyau ou se trouver à proximité de la membrane vacuolaire. En effet, pour l'instant, les contrôles que nous avons réalisés sont tous présents à l'intérieur d'un autre compartiment et ne nous renseignent pas quant à l'assemblage aspécifique des deux fragments de la Split-GFP.

Si le témoin (DHFR)_{2/3}- β 11 s'avère positif (observation d'un signal nucléaire et/ou vacuolaire), cela signifiera que l'affinité entre les deux fragments est trop importante et induit une reconstitution aspécifique de la GFP ou CFP (au regard de la compartimentalisation de la Split-GFP ou -CFP). Pour remédier à cela, il serait alors possible d'entamer une collaboration avec l'équipe du Dr. Cabantous ; elle a optimisé les fragments β 1-10 et β 11 afin d'augmenter leur solubilité et leur capacité d'autoassemblage. Nous pourrions ainsi modifier certaines mutations décrites dans la littérature, ou muter certains résidus spécifiques afin de diminuer l'autoassemblage des deux fragments. Il faudrait alors vérifier que la protéine (DHFR)_{2/3}- β 11 (et/ou NES- β 11) ne permette(nt) plus l'apparition d'un signal GFP ou CFP, et surtout qu'une protéine se trouvant au niveau des compartiments ciblés en est toujours capable.

Lorsque ces contrôles seront validés, nous pourrons utiliser l'outil de Split-GFP nucléaire pour étudier la localisation systématique du protéome de *S. cerevisiae* dans ce compartiment, délimiter les signaux d'import de protéines nucléaires, mais aussi étudier et prouver la relocalisation nucléaire de la majorité des aaRS pour finalement caractériser les NLS présents au sein de ces enzymes. A plus long terme, il pourrait être intéressant de muter ces signaux d'import et de valider les fonctions nucléaires essentielles de ces enzymes. De plus, avec l'outil de Split-CFP vacuolaire, il sera possible d'étudier la capacité de relocalisation de la totalité du protéome à la surface de ce compartiment. Nous pourrions également étudier la capacité de relocalisation de l'ensemble des aaRS de levure et de la protéine d'assemblage du complexe AME, Arc1. En outre, ces données seront cruciales pour : (i) comprendre les règles gouvernant la localisation des protéines au niveau de la membrane vacuolaire et (ii) déterminer les signaux consensus de localisation au niveau de cette membrane. Nous

pourrons alors déterminer si l'élimination de la capacité d'ancrage des aaRS au niveau de ce compartiment induit des retards de croissance dans différentes conditions et, par la même occasion, tester si ces enzymes peuvent jouer le rôle de senseurs de leurs aa homologues, comme c'est le cas de la LeuRS par exemple (Bonfils et al., 2012).

11.Expression du fragment β 1-10 à partir du génome mitochondrial

La relocalisation des protéines dans les mitochondries fait intervenir des transporteurs spécifiques au niveau de la membrane mitochondriale externe (MME) et interne (MMI) (**Figure 19**). Les protéines matricielles possèdent, au niveau de leurs régions N-ter, des signaux d'import reconnus spécifiquement par le complexe TOM. En parallèle, l'import des protéines en forme de tonneaux, contenant des signaux d'adressage internes, fait intervenir une autre voie qui favorise leurs incorporations directes dans la MMI ou MME. La GFP, ainsi que ses variants se replie sous forme de tonneau et l'ajout de GatF au niveau de leur région N-ter ne permet peut-être pas la relocalisation de la totalité de GatF- β 1-10 dans la matrice mitochondriale. Il reste potentiellement une fraction de cette protéine de fusion au niveau de la MME et celle-ci peut probablement s'autoassembler avec une protéine étiquetée β 11 qui serait retrouvée au niveau de cette membrane. L'utilisation de la protéine GatF n'est probablement pas le meilleur moyen d'importer le fragment β 1-10 dans la mitochondrie. De plus, lorsque je visualisais la fluorescence issue de la souche co-exprimant GatF- β 1-10 et P_{gk1}- β 11, le signal de fluorescence fluctuait et certaines cellules (minoritaires) présentaient un signal fort, tandis que les autres n'émettaient peu ou pas de signal. Cette fluorescence était donc probablement due à une reconstitution pré-import ou à une saturation de la machinerie d'import mitochondriale, suffisante pour obtenir un signal à la surface de ce compartiment après autoassemblage avec une protéine étiquetée β 11 non importée. Pour s'affranchir des problèmes issus de la présence d'une fraction résiduelle non importée de GatF- β 1-10, nous avons décidé d'exprimer le fragment β 1-10 directement à partir du génome mitochondrial. En effet, chez la levure, le génome mitochondrial peut être modifié et cette technique, appelée transformation biolistique, a déjà fait ses preuves (Bonnefoy and Fox, 2007; Johnston et al., 1988). Ne disposant pas de la technologie à Strasbourg, nous avons collaboré, dans un premier temps, avec le laboratoire du Pr. Jean Paul Di-Rago à Bordeaux et plus précisément avec le Dr. Emmanuel Tetaud. Pour cela, j'ai dû construire des fragments β 1-10 dont l'expression est possible dans la mitochondrie (l'usage des codons est différent) et qui sont flanqués des

séquences 5' et 3' UTR du gène mitochondrial *COX2* ; la souche utilisée pour la transformation biolistique est *cox2* (gène muté). Cette souche ne permet donc pas l'expression de la protéine Cox2 correspondante et n'est donc pas capable de se développer *via* un métabolisme respiratoire. J'ai directement réalisé les constructions pour les trois couleurs (verte, jaune et cyan). Après avoir conçu les séquences, je les ai fait synthétiser par Genscript®. Ces séquences ont ensuite été sous-clonées dans le vecteur de biolistique pPT24. En plus de la séquence codant les fragments β 1-10 (GFP, YFP et CFP), ce plasmide contient la région génomique de *COX2*. Lors de la transformation biolistique, ces plasmides sont bombardés sur des levures et certains d'entre eux se retrouvent dans la mitochondrie. C'est alors par recombinaison homologe, grâce aux 5' et 3' UTR de *COX2*, que la séquence codant le fragment β 1-10, ainsi que le gène *COX2* seront intégrés au niveau du génome mitochondrial. Ceci permet alors de restaurer la capacité de croissance respiratoire (Expression de Cox2) de la souche transformée et d'exprimer le fragment β 1-10 à partir des machineries transcriptionnelle et traductionnelle mitochondriales. Nos collaborateurs ont réalisé la transformation biolistique, vérifié la capacité respiratoire des transformants validant l'intégration au locus. Finalement, quatre à six mois après la transformation biolistique, nous avons pu tester la reconstitution mitochondriale de la molécule fluorescente dans les trois souches exprimant chacune, un fragment β 1-10 différent. Pour valider l'intégration mitochondriale et l'expression des différents fragments β 1-10, j'ai transformé ces souches avec des plasmides d'expression de la GluRS- β 11 et Pgc1- β 11 et testé la capacité de reconstitution mitochondriale de la GFP, CFP ou YFP. Malheureusement, les différentes tentatives ne m'ont jamais permis de visualiser un signal fluorescent vert, jaune ou cyan. En effet, la synthèse de protéines codées directement au niveau du génome mitochondrial n'est pas chose aisée et de nombreux facteurs peuvent entrer en jeu. La transcription de ce gène peut être défaillante, induisant une expression trop faible du fragment β 1-10 mitochondrial pour permettre la détection de la molécule fluorescente reconstituée. De plus, il est possible que le mitoribosome, qui synthétise uniquement des protéines très hydrophobes que l'on qualifie de protéolipides, ne puisse pas synthétiser le fragment β 1-10 ou que la protéine β 1-10, adoptant une structure de tonneau, soit toxique et puisse potentiellement s'agréger. Par conséquent, il y aurait incapacité pour le fragment β 11 importé dans la mitochondrie de s'associer avec le fragment β 1-10 ; aucun signal ne serait visualisable.

Pour éliminer l'hypothèse d'une erreur de notre part, nous avons néanmoins contrôlé toutes les séquences utilisées et les plasmides construits. En vérifiant les séquences synthétisées et optimisées par Genscript®, nous avons observé que le codon stop que nous avons précisé comme devant être UAA, avait été changé en UGA. Or, dans la mitochondrie, celui-ci n'est pas un codon stop, mais code pour un tryptophane et induit l'incorporation de sept aa supplémentaires au niveau de la chaîne polypeptidique des fragments β 1-10. L'ajout de ces résidus influence potentiellement la reconstitution de la GFP, CFP ou YFP. Après avoir muté ce codon Trp (UGA) en codon stop (UAA), les transformations biolistiques ont été effectuées à nouveau avec les trois constructions. Après une nouvelle période de validation des souches, la reconstitution de la GFP active a de nouveau été étudiée par microscopie à fluorescence sur ces nouvelles souches exprimant la GluRS- β 11 ou P_{gk1}- β 11. Cependant, malgré l'élimination de ces sept aa supplémentaires, aucun signal n'a pu être observé et il semblerait que la reconstitution de la GFP active ne soit pas suffisante pour permettre la visualisation d'un signal. Comme je l'ai montré par immuno-détections de la GluRS sur des extraits mitochondriaux, cette protéine est fortement importée dans ce compartiment et peut potentiellement s'associer avec le fragment β 1-10 mitochondrial et reconstituer la molécule fluorescente active. Les seuls facteurs limitants sont donc l'expression de la protéine β 1-10, la capacité de reconstitution de la GFP active et la quantité de signal produit.

Nous n'avons cependant pas renoncé à exprimer ce fragment à partir du génome mitochondrial et nous avons contacté une autre collaboratrice, la Pr. Roza Kucharczyk, qui est également une spécialiste de la transformation biolistique. Après lui avoir expliqué les différents problèmes rencontrés, nous lui avons envoyé les vecteurs biolistiques utilisés lors du second essai. En temps qu'experte de cette technique, elle a soulevé un autre problème dans le plasmide final. En effet, sur la carte du plasmide pPT24 que nous possédons, six nucléotides sont manquants à la fin de la séquence de *COX2* et cela induit l'expression d'une version tronquée de cette protéine qui ne permet peut-être pas une expression optimale de la protéine correspondante. L'intégration au niveau du locus *COX2* n'ayant pas fonctionné à deux reprises, nous nous interrogeons également sur le fait que ce locus puisse être la source de nos problèmes. Comme la Pr. Kucharczyk avait créé une souche dans laquelle le locus *ATP6* était disponible pour une intégration, nous avons choisi d'utiliser ce locus pour nos nouvelles tentatives d'intégration des gènes codant les fragments β 1-10. Cette souche, nommé DFS160,

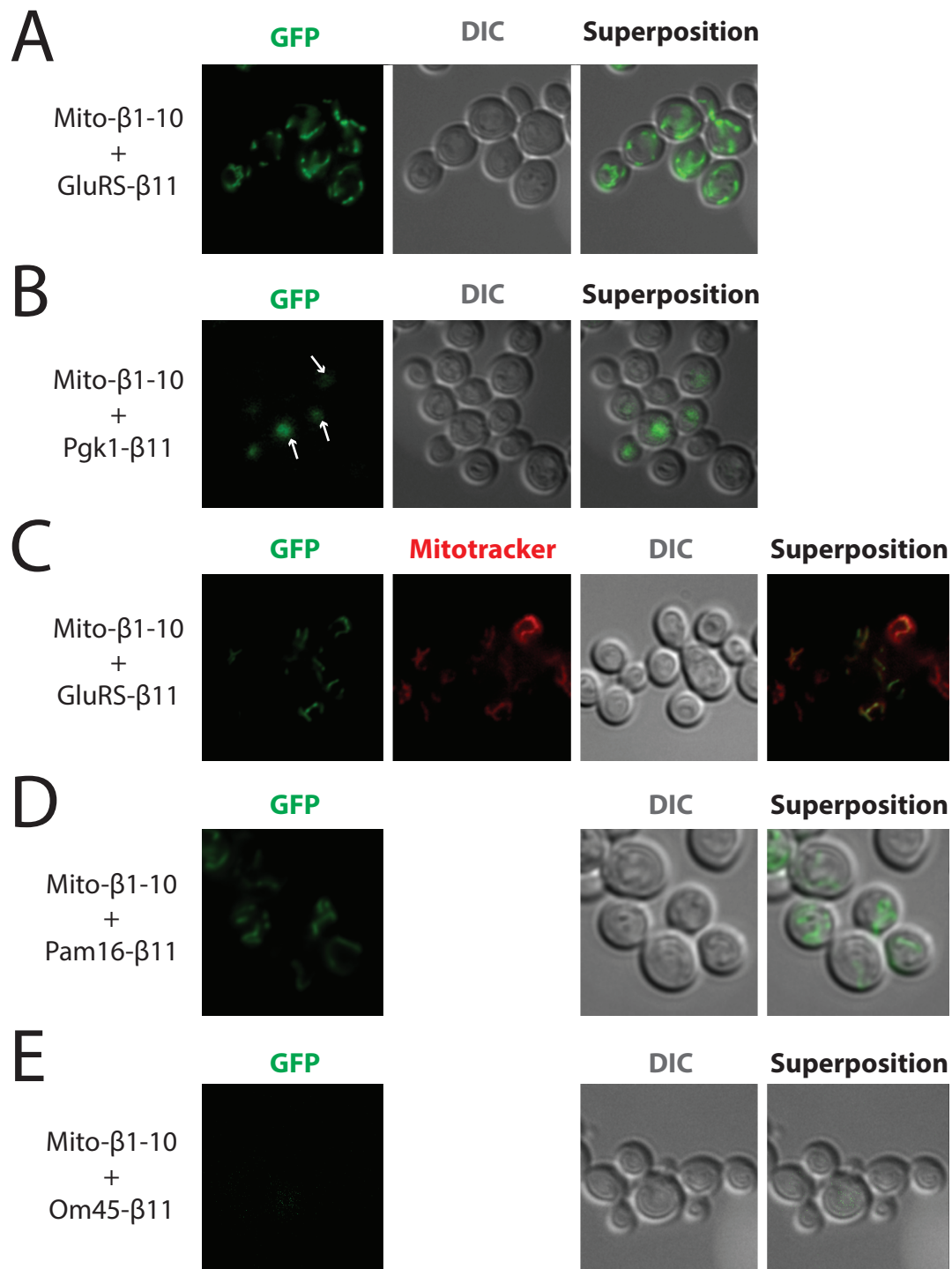


Figure 38 : Expression du fragment β 1-10 à partir du génome mitochondrial

Le fragment β 1-10 de la Split-GFP est exprimé directement à partir du génome mitochondrial suite à une transformation biolistique et intégration au niveau du locus *atp6*. Cette souche (Mito- β 1-10) est transformée avec le plasmide d'expression de la GluRS- β 11 (A) ou Pgc1- β 11 (B) et l'apparition de fluorescence est observée par microscopie à épifluorescence. Les flèches représentent la fluorescence vacuolaire non spécifique issue de l'absence d'Ade2. Après restauration de l'expression d'Ade2, les souches ont été transformées avec les vecteurs d'expression de : GluRS- β 11 (C), Pam16- β 11 (D) et Om45- β 11 (E). La fraction mitochondriale est révélée par marquage au Mitotracker et l'apparition de fluorescence est visualisée par microscopie à épifluorescence.

possède un génome mitochondrial "muté". En effet, le gène *ATP6* a été remplacé par le marqueur d'auxotrophie *ARG8m*. Ce gène, initialement exprimé à partir du génome nucléaire, a été optimisé pour son expression dans la mitochondrie. La protéine Atp6 étant indispensable pour la croissance respiratoire, le gène correspondant est exprimé à partir du locus *COX2* ; Cox2 est également exprimé à partir de ce locus. Les gènes codant le fragment β 1-10 ont donc été flanqués des séquences non traduites d'*ATP6* et clonés dans un plasmide en vue de leur intégration suite à la transformation biolistique. Ces plasmides ont ensuite été bombardés par biolistique dans la souche DFS160, permettant, après suite à des croisements successifs, l'expression du fragment β 1-10 au niveau du locus *ATP6*. Comme pour les autres souches, nos collaborateurs ont réalisé la transformation biolistique et la sélection des intégrants mitochondriaux. Finalement, après toutes les vérifications nécessaires, nous avons pu tester les souches ayant intégré les différentes constructions. Arrivant en fin de thèse, je n'ai cependant pas testé toutes les souches et je me suis focalisé sur celle exprimant le fragment β 1-10, permettant la reconstitution d'une GFP active. En effet, j'ai décidé d'utiliser cette souche car personnellement, je pense que la détection du signal vert est plus sensible que celle des autres couleurs. Je l'ai alors transformée avec les plasmides d'expression de la GluRS- β 11 ou Pkg1- β 11 et l'apparition d'un signal a été visualisée par microscopie à épifluorescence. De manière intéressante, la souche exprimant la GluRS- β 11 produit un signal fluorescent sous forme de réseau (**Figure 38A**). Ces résultats confirment (i) que le fragment β 1-10 est exprimable et exprimé à partir du génome mitochondrial et (ii) que celui-ci est produit en quantité suffisante pour induire la reconstitution de la GFP active et la détection d'un signal fluorescent. Lorsque cette souche exprime la protéine Pkg1- β 11, aucune fluorescence mitochondriale n'est observable, confirmant bien la spécificité de reconstitution de la GFP après import de la GluRS dans la mitochondrie (**Figure 38B**). Cependant, cette souche possède un gène *ade2* muté, induisant la production d'un pigment rouge au niveau de la vacuole qui est détectable lors de l'excitation avec tous les canaux (**Figure 38B, Flèches blanches**). Lors du marquage des mitochondries avec le Mitotracker, le pigment rouge pourrait empêcher la visualisation du réseau mitochondrial et il est nécessaire de rajouter une forte quantité d'adénine (100 mg/L) dans les boîtes de culture, mais également lors des cultures liquides pour empêcher la production de ce pigment. Pour s'affranchir de ce problème, j'ai donc restauré le gène *ADE2* par transformation avec un produit de PCR favorisant la recombinaison homologue de ce fragment au niveau du gène déficient et empêchant la production du

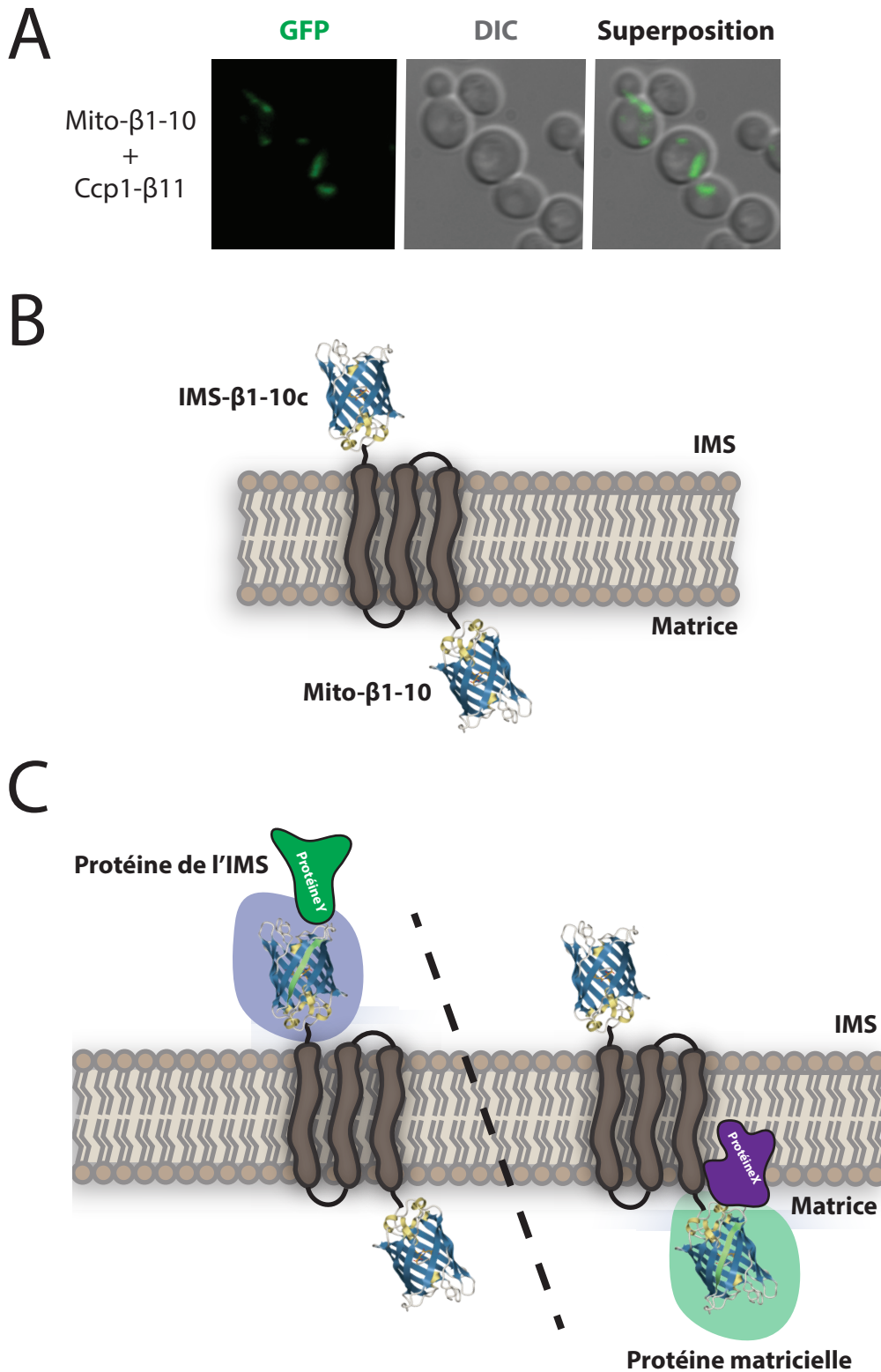


Figure 39 : Expression de la protéine de l'IMS Ccp- β 11 et perspectives

La souche Mito- β 1-10, exprimant le fragment β 1-10 à partir du génome mitochondrial, est transformée avec le plasmide d'expression de la protéine de l'IMS Ccp1- β 11 (A). La reconstitution de la GFP active est visualisée par microscopie à épifluorescence. Afin d'étudier la maturation des protéines de l'IMS, une souche exprimant deux fragments β 1-10c (vers l'IMS) et β 1-10 (vers la matrice) au sein de la MMI va être construite (B). Cette souche permettra notamment l'établissement simultané du protéome de l'IMS et de la matrice par émission d'une fluorescence spécifique de la localisation sub-mitochondriale de la protéine d'intérêt (C).

pigment. Finalement, j'ai testé plusieurs clones par transformation avec le plasmide d'expression de la GluRS- β 11, afin de vérifier l'intégration du gène *ADE2* sauvage et la capacité de la souche à produire un signal fluorescent. Les images de microscopie confirment la production de fluorescence lorsque la GluRS- β 11 est importée dans la mitochondrie (**Figure 38C**). De plus, le signal colocalise parfaitement avec le marquage mitochondrial au Mitotracker, confirmant bien la présence des deux fragments de la Split-GFP, ainsi que la reconstitution de la molécule active dans ce compartiment. Lorsque qu'une protéine mitochondriale est fusionnée au fragment β 11 (Pam16), le signal de fluorescence est identique à celui obtenu pour la GluRS (**Figure 38D**). Par ailleurs, l'expression d'une protéine de la MME (Om45) étiquetée au fragment β 11, ne permet pas la reconstitution de la GFP, confirmant bien que le fragment β 1-10, exprimé à partir du génome mitochondrial, ne traverse pas la MMI et l'IMS pour s'enchâsser dans la MME (**Figure 38E**). Finalement, contrairement à la souche initiale, aucun signal vert ou rouge n'est observé au niveau de la vacuole, ce qui confirme la restauration d'un gène *ADE2* et l'expression de la protéine fonctionnelle. J'ai décidé de nommer cette souche Mito- β 1-10, car elle produit le fragment β 1-10 directement à partir du génome mitochondrial.

De manière plus surprenante, lorsque la protéine de l'IMS Ccp1- β 11 est exprimée dans cette souche, un signal de fluorescence est visualisable (**Figure 39A**). Plusieurs hypothèses peuvent expliquer la reconstitution de la GFP. En effet, il est possible que lors de sa maturation, Ccp1 ou du moins l'extrémité C-ter (fusionnée au fragment β 11), se retrouve transitoirement dans la matrice mitochondriale. Les deux fragments étant dans le même compartiment, leur autoassemblage induit la reconstitution de la molécule active et l'émission de fluorescence. Cependant, cette hypothèse semble peu plausible car le mécanisme de maturation est clairement défini et c'est l'extrémité N-ter qui est clivée lors de son passage dans la matrice (Esser et al., 2002; Tatsuta et al., 2007; van der Blik and Koehler, 2003). De plus, lors de son repliement, le fragment β 1-10 adopte probablement une structure en tonneau et il est envisageable qu'elle puisse éventuellement favoriser son intégration dans la MMI. Suite à l'internalisation membranaire du fragment β 1-10, une protéine de l'IMS fusionnée au fragment β 11 engendrerait alors la reconstitution de la GFP et l'émission de fluorescence. Ceci est néanmoins une hypothèse qui devra être vérifiée expérimentalement.

En guise de perspectives, nous souhaiterions confirmer l'hypothèse du passage de Ccp1- β 11 dans la matrice mitochondriale. Pour ce faire, il faudrait fusionner d'autres protéines de l'IMS au fragment β 11, comme par exemple une protéine ancrée dans la MME dont l'extrémité C-ter est libre dans l'IMS ou une protéine totalement libre ce sous-compartiment. La protéine de la MME serait probablement une protéine reconnue directement par le complexe MIM, capable de l'intégrer sans son passage dans l'IMS. En parallèle, la protéine libre dans l'IMS pourrait être une protéine dont le signal d'import est riche en Cys, et de ce fait, reconnue par le complexe MIA 40. Ces deux témoins nous permettront de valider l'absence du fragment β 1-10 mitochondrial dans l'IMS, mais également de confirmer que la maturation de Ccp1 implique un passage de la protéine, ou du moins de la région C-ter, dans la matrice mitochondriale. Si un signal est détecté uniquement lors de l'expression de la protéine libre dans l'IMS, cela suggérerait une localisation membranaire du fragment β 1-10. Il serait alors intéressant de déterminer comment se déroule son intégration, à savoir si elle est co-translationnelle, comme pour les protéines de la CR par exemple (Barrientos et al., 2004; Mick et al., 2007; Pierrel et al., 2007) ou post-translationnelle.

La transformation biolistique du gène codant le fragment β 1-10 dans la mitochondrie est efficace pour son intégration dans le génome et son expression en quantité suffisante pour que les deux fragments s'autoassemblent et produisent un signal fluorescent visualisable. Nous voudrions donc créer une souche ayant le ou les fragments β 1-10 ancré(s) au niveau de la MMI. Pour cela, nous allons les fusionner à des domaines transmembranaires d'une des protéines de la CR (Atp6 ou 9) de façon à orienter leur insertion. A partir du génome mitochondrial, nous voudrions exprimer deux fragments β 1-10 favorisant la reconstitution de GFP et CFP actives. Afin de déterminer l'orientation membranaire de la protéine, nous allons fusionner le fragment β 1-10c et le fragment β 1-10 au niveau de l'extrémité N-ter et C-ter respectivement (**Figure 39B**). Après synthèse, cette protéine sera incorporée dans la MMI et nous pourrions étudier si l'orientation est toujours la même. En effet, nous allons tester plusieurs protéines libres dans l'IMS et plusieurs protéines matricielles afin de déterminer si l'un des deux signaux est constitutivement émis ou si un mélange de fluorescence verte et cyan est visualisable. Si l'orientation est toujours la même, nous pourrions alors utiliser cette souche pour étudier la relocalisation des protéines de l'IMS et de la matrice mitochondriale

simultanément. Bien sûr, il faudra dans un premier temps montrer qu'une protéine strictement matricielle ne permet pas l'émission du signal cyan, caractéristique de l'IMS (**Figure 39C**) et inversement, que la fluorescence verte ne soit pas visualisable lors de l'expression d'une protéine de l'IMS.

Cette souche ouvrira donc de nombreuses perspectives pour l'étude de la localisation mitochondriale des protéines. En effet, lors de l'expression des protéines de levure étiquetées GFP, plus de 25 % ne produisent pas de fluorescence (Ghaemmaghami et al., 2003). Ceci est probablement dû à la taille ou à la position de l'étiquette. Le fragment $\beta 11$ étant plus petit, il serait intéressant de créer une banque de levure où toutes les ORF sont suivies de la région codant ce fragment. Ces souches pourront alors être croisées avec la souche exprimant les fragments $\beta 1-10$ mitochondriaux et des études de l'ensemble de mitoprotéome et du protéome de l'IMS pourront être réalisées par microscopie à fluorescence. L'étiquette étant plus petite, nous nous attendons à réduire fortement la proportion de protéines non exprimées ou ayant des problèmes de repliement. Pour étudier la relocalisation mitochondriale, il est néanmoins nécessaire que la quantité d'import soit suffisante. Ceci passera peut-être par la modification du promoteur de certains gènes peu exprimés qui permettra alors la surproduction de la protéine correspondante. Une fois les diploïdes générés, la relocalisation mitochondriale de l'ensemble du protéome pourra être étudiée dans différentes conditions physiologiques (variation de température, source de carbone utilisé...) ou en réponse à des stimuli (irradiation aux UV ou tests d'antibiotiques). Finalement, l'expression des versions étiquetées N-ter pourra également être testée pour les études de relocalisation mitochondriale des protéines de levure. Ceci pourrait notamment être intéressant pour les protéines dont la fusion C-ter induit une perte de fonction ou une absence de relocalisation.

Partie III : Applications de l'outil de Split-
GFP pour l'étude des protéines à
localisations subcellulaires multiples
chez *S. cerevisiae*

1. Contexte du projet

Afin de comprendre la fonction d'une protéine, il est indispensable de déterminer sa localisation subcellulaire. Cependant, dans certains cas, l'étude de la localisation est plus complexe car des protéines cytosoliques, dites multi-localisées, peuvent également se trouver dans plusieurs autres compartiments au même moment. Pour de nombreuses protéines cytosoliques multi-localisées, seule la localisation principale (le cytosol) a été caractérisée. En effet, lors de la fusion de ce type de protéines à la GFP, la fraction cytosolique est détectée et elle masque les autres fractions de la protéine se relocalisant notamment dans les organelles. La mise en évidence des compartiments additionnels, dans lesquels se relocalisent ces protéines, est d'autant plus importante que, la plupart du temps, leurs fonctions changent lorsqu'elles rejoignent ces compartiments. Bien souvent ce nouveau rôle est également essentiel pour la cellule. Ainsi, une protéine se relocalisant dans le noyau pourra aussi bien moduler la transcription des gènes ou la réplication de l'ADN, alors que sa fonction cytosolique peut être essentielle pour la régulation de la synthèse protéique et le trafic intracellulaire. La famille d'enzyme que nous étudions illustre parfaitement ce principe. Les aaRS cytosoliques, essentielles pour la synthèse protéique de par leur activité "primaire" d'aminocyclation des ARNt, ont la capacité de se relocaliser dans d'autres compartiments pour assurer des fonctions non canoniques indispensables à l'homéostasie cellulaire. Comme je l'ai décrit dans l'introduction et dans la partie I des résultats, nous avons pu montrer la capacité de la GluRS et la MetRS de levure de se relocaliser dans d'autres compartiments subcellulaires (Frechin et al., 2014; Frechin et al., 2009b). D'ailleurs, la MetRS humaine, tout comme la MetRS de levure, est capable de se relocaliser dans le noyau pour y réguler la transcription des gènes. Il est à noter que les deux homologues ne régulent cependant pas la transcription des mêmes gènes (Frechin et al., 2014; Ko et al., 2000). La LeuRS cytosolique de levure est également capable de rejoindre un autre compartiment cellulaire. En effet, en réponse à la concentration intracellulaire en leucine, cette enzyme se relocalise à la surface de la vacuole pour réguler la voie TOR et la synthèse protéique (Bonfils et al., 2012). Comme l'aaRS de levure, la LeuRS humaine est capable de réguler la synthèse protéique en réponse à la quantité de leucine (Han et al., 2012). Elle se localise alors au niveau du lysosome (vacuole chez la levure) et permet de réguler la voie TOR. Il en est de même pour la TyrRS, où les homologues humain et de levure se relocalisent dans le noyau. L'enzyme de levure va y favoriser la maturation et l'export de

l'ARNt^{Tyr} (Azad et al., 2001), tandis que l'aaRS humaine répond au stress oxydatif et active un facteur de transcription, qui va augmenter l'expression de protéines de réparation des dommages à l'ADN (Wei et al., 2014). La relocalisation des aaRS est primordiale pour leur participation à de nombreux processus biologiques et elle peut être conservée à travers les espèces. L'étude de la capacité des aaRS de levure à se relocaliser pourrait être un bon indicateur de la capacité de leurs homologues humains à se relocaliser dans les mêmes compartiments.

Durant ma thèse, j'ai co-écrit une revue portant notamment sur la présence de signaux d'import putatifs chez les aaRS eucaryotes, nous permettant ainsi d'avoir une première idée de la capacité de relocalisation de ces enzymes. En effet, nous avons effectué une recherche bio-informatique des signaux d'adressage nucléaires putatifs retrouvés au sein de cette famille d'enzyme à l'aide d'outils de prédictions. Cette étude suggère que toutes les aaRS de levure et humaines possèdent un NLS et sont donc potentiellement importées dans le noyau (Bader et al., 2017). De plus, nous avons également étudié la présence de signaux d'adressage mitochondriaux au sein de ces mêmes enzymes cytosoliques à l'aide de logiciels de prédiction. L'analyse de la présence de ce type de signaux chez les aaRS cytosoliques confirme bien que certaines de ces protéines peuvent être importées dans la mitochondrie. Néanmoins, leur nombre est beaucoup plus faible, suggérant que ce phénomène est moins répandu (Bader et al., 2017). Un signal d'import mitochondrial est notamment détecté au sein de la GluRS. De manière intéressante, nous avons déjà validé la présence de cette protéine dans la mitochondrie. Cependant, le signal identifié bio-informatiquement n'est pas présent dans la même région que celui délimité lors d'une précédente étude (Frechin et al., 2009b). Malgré de nombreuses preuves de la présence de signaux d'adressage compartiments-spécifiques au sein des aaRS, la validation de ces localisations doit être réalisée *in vivo*. A l'heure actuelle, aucune étude n'a permis l'identification des localisations additionnelles du jeu complet d'aaRS, notamment dans la mitochondrie. Ainsi, j'ai décidé d'utiliser les outils que j'ai élaboré durant ma thèse pour étudier la localisation de l'ensemble de ces protéines.

2. Amplification et clonage des gènes d'aaRS

J'ai tout d'abord recherché les gènes codant les aaRS cytosoliques sur la SGD (Cherry et al., 2012). J'ai alors voulu sous-cloner les différentes séquences dans des vecteurs

d'expression en amont de la séquence codant le fragment $\beta 11$. Cependant, l'utilisation des techniques de clonage conventionnelles est trop compliquée à mettre en œuvre car pour le clonage orienté de ces gènes, il faut trouver deux sites de restriction qui sont absents du gène d'intérêt et présents dans la cassette de clonage du vecteur. La probabilité qu'au moins deux sites de restriction, présents dans le SCM (Site de Clonage Multiple) et permettant le clonage en phase, soient disponibles pour cloner les 20 gènes d'aaRS est très faible. Nous avons alors décidé de changer de technique de clonage et de passer au système Gateway[®]. Ainsi, j'ai dessiné les 20 paires d'amorces d'amplification de l'ensemble des gènes codant les aaRS cytosoliques, flanqués des séquences AttB1 et AttB2 et favorisant leur clonage par la technique Gateway[®]. La création de la collection des 20 gènes des aaRS cytosoliques compatibles avec le système Gateway[®] a été réalisée en collaboration avec trois étudiants de licence et master que j'ai encadrés (Chloé Groh, Guillaume Grob et Maximilian Geiger). Par la suite, le clonage de ces gènes par Gateway[®] a été laborieux car le fichier expliquant la conception des amorces qui nous avait été fourni était erroné et un codon crucial pour la recombinaison du gène d'intérêt dans le vecteur intermédiaire (pDONR221) était absent. L'absence de ce codon réduit drastiquement l'efficacité de la recombinaison et la plupart des colonies obtenues après transformation ne contenait pas le plasmide recombiné (contenant le gène d'intérêt). Finalement, après avoir changé les amorces et réamplifié chaque gène, 18 des 20 gènes d'aaRS ont pu être sous-clonés dans les vecteurs d'expression contenant la séquence codant le fragment $\beta 11$ (pAG414-GPD-aaRS- $\beta 11$). Seuls deux gènes d'aaRS cytosoliques (CysRS et ProRS) n'ont pas pu être cloné par cette technique. En effet, malgré une quantité suffisante de produits PCR obtenus par amplification, la ligation par la technique classique, mais également par la technique Gateway[®] n'a jamais permis d'obtenir ne serait-ce que quelques colonies positives. Le nombre de colonies était très faible et les plasmides n'avaient jamais les bons profils de restriction ; le gène d'intérêt n'était jamais recombiné dans le vecteur intermédiaire. Le passage par une bactérie ne m'a donc pas permis d'obtenir des clones positifs et cela pourrait être dû à la toxicité de ces deux enzymes lors du passage dans *E. coli*. Cette hypothèse est corroborée par une étude bio-informatique suggérant que plus de 15000 gènes provenant de 384 espèces microbiennes sont toxiques pour la bactérie et, de ce fait, difficile à exprimer (Kimelman et al., 2012). Pour le gène codant la CysRS, nous avons également effectué une synthèse de gène, ainsi qu'un sous-clonage, auprès de la société Genscript[®]. Les techniciens qui étaient en charge de ce projet ont également éprouvé des

difficultés pour le clonage de ce gène et ils ont finalement dû le cloner dans un vecteur à faible nombre de copies contenant un promoteur inductible. Cependant, par manque de temps, je n'ai pas pu utiliser cette construction. Par conséquent, j'ai réussi à exprimer 18 aaRS étiquetées β 11 dans la souche Mito- β 1-10, exprimant le fragment β 1-10 à partir du génome mitochondrial. Comme je l'ai démontré lors de l'élaboration de l'outil de Split-GFP mitochondriale, la GluRS est fortement importée aussi bien fermentation qu'en respiration. J'ai décidé de réaliser mes cultures en milieux fermentaires sans répression des mitochondries (gal, galactose) et permettant le maintien du plasmide d'expression (SCGal-Trp). Les images ont toutes été collectées par microscopie à épifluorescence et retravaillées avec le logiciel Image J.

3. Etude de la relocalisation mitochondriale des aaRS cytosoliques, d'AIM10, la ProRS mitochondriale putative et de la protéine d'ancrage du complexe AME, Arc1

Dans la cellule, il existe un jeu d'aaRS dans chaque compartiment traductionnellement actif. Cependant, chez de nombreux organismes, le jeu d'aaRS d'un des compartiments est incomplet et la cellule doit compenser cette absence par divers mécanismes. Chez la levure, tous les gènes codant les aaRS mitochondriales ne sont pas présents. En effet, certains gènes codent pour la version cytosolique et organellaire grâce à un mécanisme d'initiation de la traduction alternative (la ValRS, l'HisRS, la GlyRS et l'AlaRS). Lors de l'amplification des gènes codant les aaRS cytosoliques, j'ai décidé d'amplifier la totalité de la région codante. Cela signifie que les protéines exprimées à partir de ces gènes contiennent les MTS permettant leurs relocalisations mitochondriales. L'expression plasmidique ne favorisera probablement pas la synthèse des versions cytosolique et organellaire de ces aaRS. Cependant, il serait intéressant d'étudier par des tests de complémentation, mais également par immuno-détections sur des extraits totaux, si les deux versions sont produites à partir du plasmide d'expression. En parallèle, les gènes codant deux autres aaRS mitochondriales sont absents (GlnRS et CysRS). Ainsi, la formation des aa- m ARNt^{aa} correspondants doit être catalysée par une aaRS ND, provenant du jeu mitochondrial ou cytosolique, induisant la formation d'une espèce mésacylée. Celle-ci est alors reconnue par une AdT ARNt-dépendante qui génère l'aa- m ARNt^{aa} correspondant. Au laboratoire, nous avons déjà prouvé que la formation du Gln-

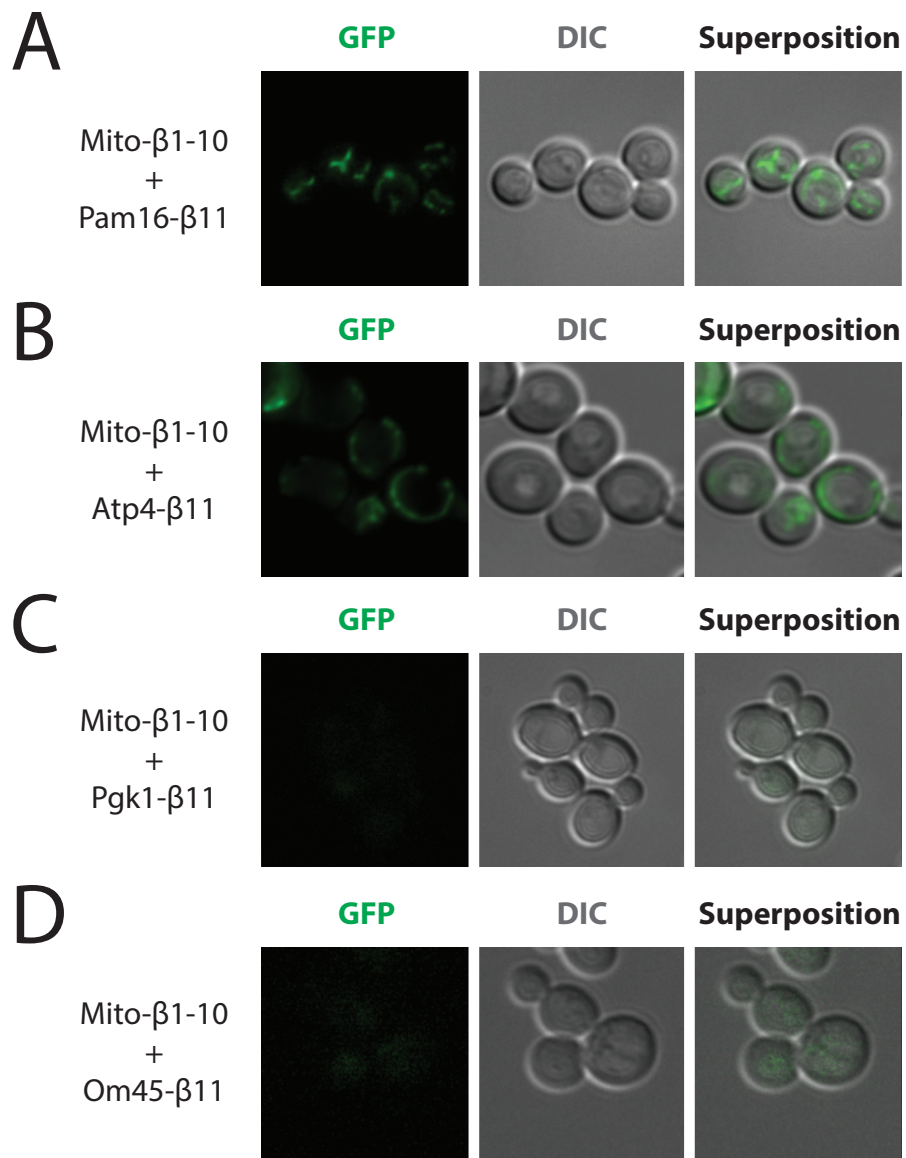


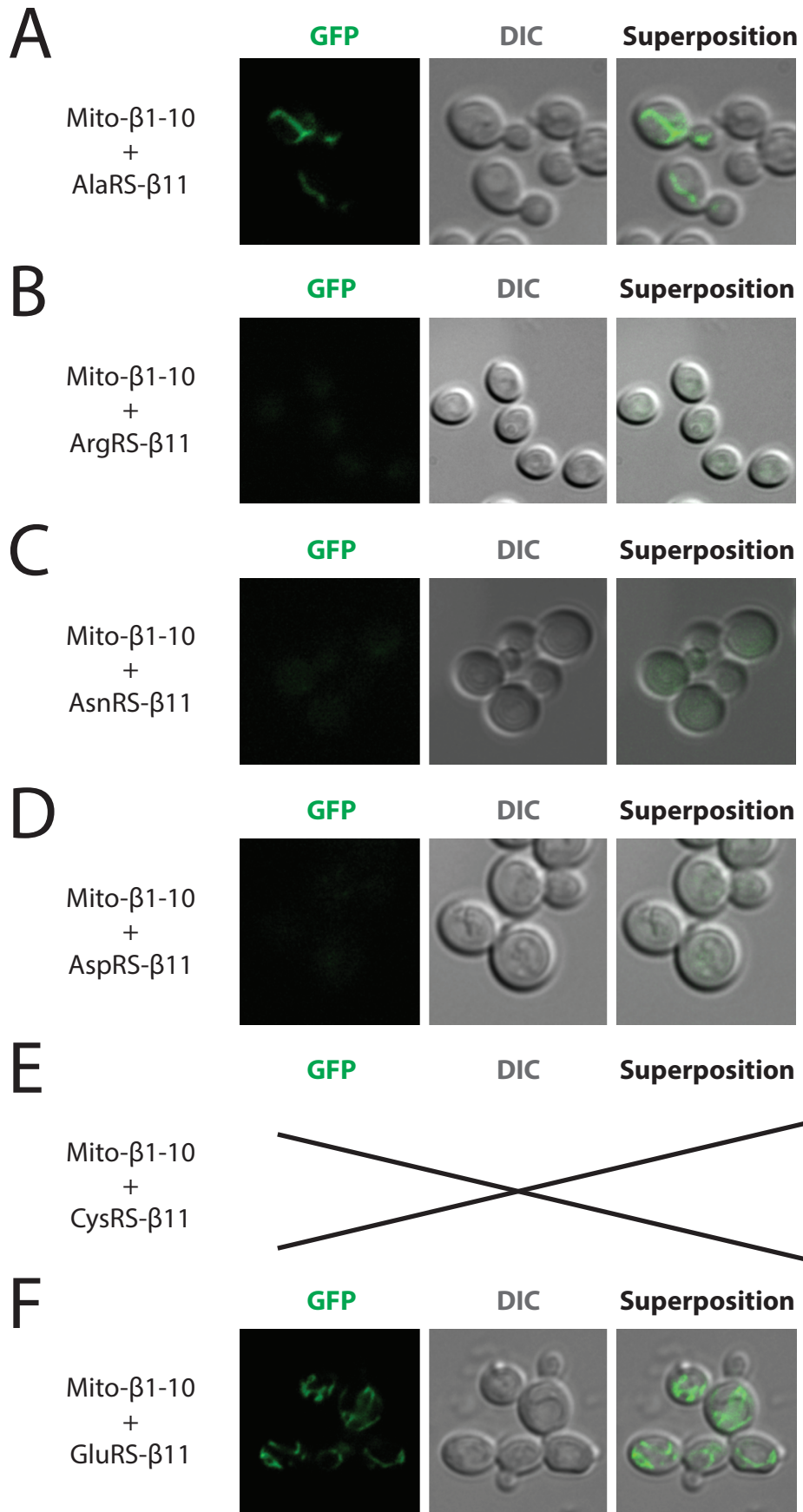
Figure 40 : Validation de témoins mitochondriaux et cytosoliques

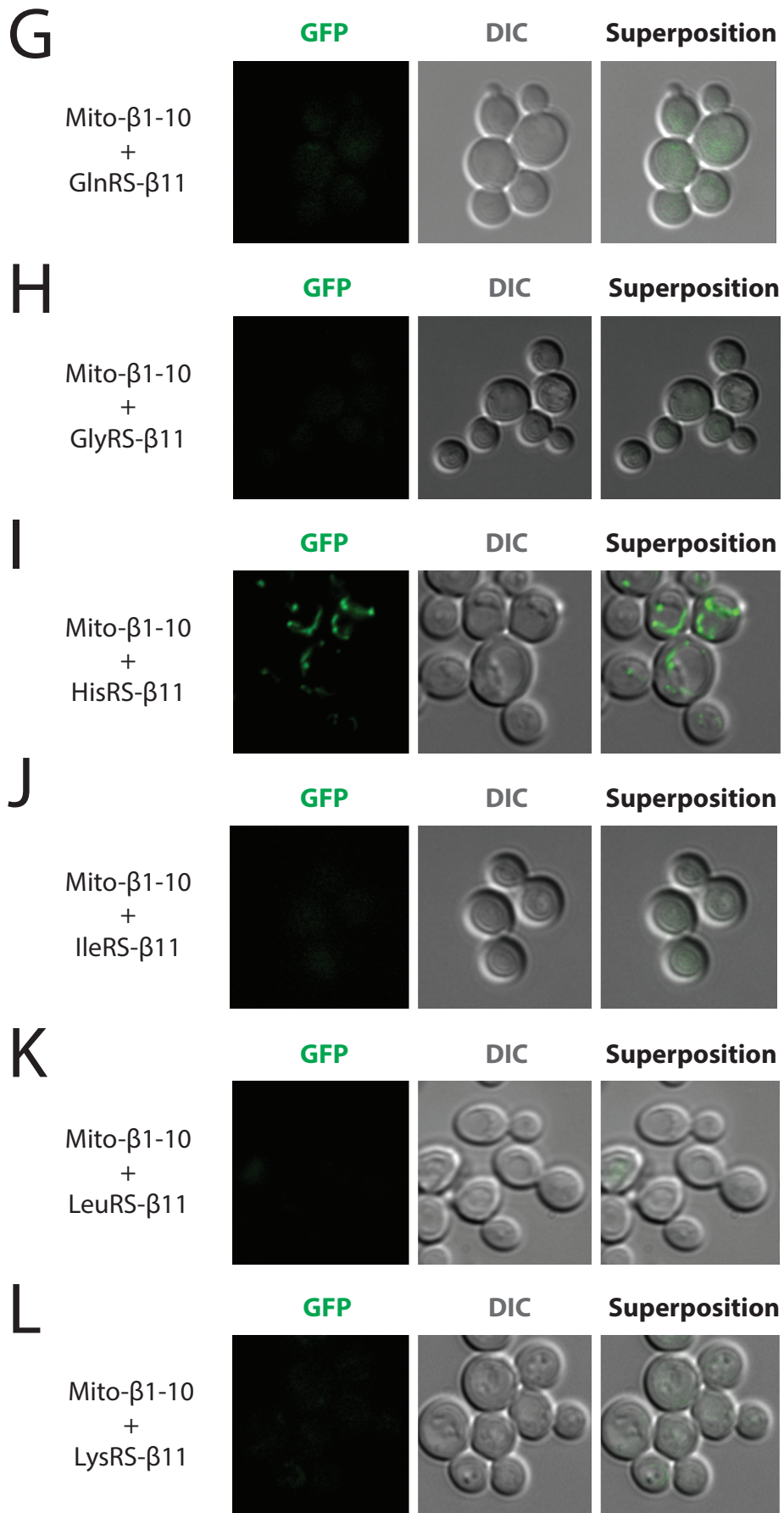
La souche Mito- β 1-10 a été transformée avec des vecteurs permettant l'expression de protéines contrôles : (A) Pam16- β 11, (B) Atp4- β 11, (C) Pgc1- β 11 ou (D) Om45- β 11 . La fluorescence émise est visualisée par microscopie à épifluorescence.

$mARNt^{Gln}$ fait intervenir la GluRS cytosolique, importée dans la mitochondrie, permettant l'aminocyclation de l' $mARNt^{Gln}$ (Frechin et al., 2009b). Cependant, la voie de formation de Cys- $mARNt^{Cys}$ n'est pas encore caractérisée. Finalement, une dernière activité d'aminocyclation n'a pas encore été décryptée, celle de la ProRS mitochondriale. Cependant, il existe un gène putatif qui code probablement pour cette enzyme : *YER087W*. Pour confirmer que ce gène code bien pour une protéine mitochondriale, j'ai également décidé de l'amplifier et de l'exprimer en fusion avec le fragment $\beta 11$. La validation de la localisation mitochondriale de la protéine exprimée à partir de ce gène nous donnera donc des informations sur sa potentielle relocalisation mitochondriale.

Dans un premier temps, j'ai réalisé la transformation des plasmides témoins (exprimant Pam16- $\beta 11$, Atp4- $\beta 11$, Pkg1- $\beta 11$ et Om45- $\beta 11$) et des plasmides d'expression des aaRS à partir de la même culture de la souche Mito- $\beta 1-10$; pour rappel, cette souche exprime le fragment $\beta 1-10$ à partir de son génome mitochondrial. J'ai tout d'abord validé les témoins et un signal fluorescent est détectable pour les souches exprimant Pam16- $\beta 11$ et Atp4- $\beta 11$ (**Figure 40A & B**), alors qu'aucun signal ne l'est pour celles exprimant Pkg1- $\beta 11$ et Om45- $\beta 11$ (**Figure 40C & D**). Le signal fluorescent est bien spécifique de la localisation intra-mitochondriale de la protéine d'intérêt.

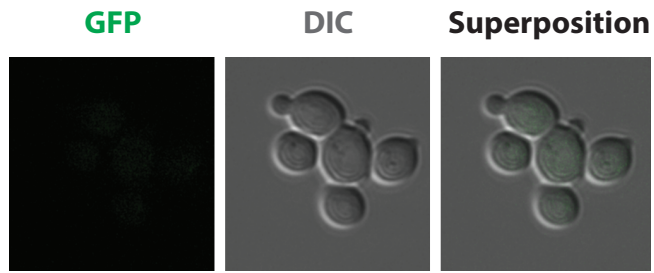
J'ai ensuite analysé la localisation mitochondriale des quatre aaRS où les versions organellaire et cytosolique sont exprimées à partir du même gène. Les souches exprimant l'AlaRS- $\beta 11$, l'HisRS- $\beta 11$, ainsi que la ValRS- $\beta 11$ présentent un signal fluorescent détectable par microscopie (**Figure 41A, I & T**), validant la localisation mitochondriale de ces aaRS et l'utilisation de l'outil Split-GFP mitochondrial pour ce type d'étude. Ces résultats confirment que la version mitochondriale de l'aaRS est codée par ce gène. Il serait maintenant intéressant de muter le premier codon d'initiation afin d'exprimer uniquement les aaRS cytosoliques. Il faudrait alors confirmer l'absence de relocalisation mitochondriale de celles-ci. La souche exprimant la GlyRS- $\beta 11$ n'émet, quant à elle, pas de fluorescence (**Figure 41H**). Ce résultat est plus surprenant car ce gène code également pour les versions cytosolique et mitochondriale de cette enzyme. Il est cependant possible que son expression soit insuffisante pour permettre la détection de sa relocalisation mitochondriale. Pour valider ces résultats, il faudrait vérifier l'expression de la protéine par immuno-détection sur un extrait total. De plus, nous voudrions augmenter le niveau d'expression de cette protéine par utilisation d'un vecteur à haut nombre





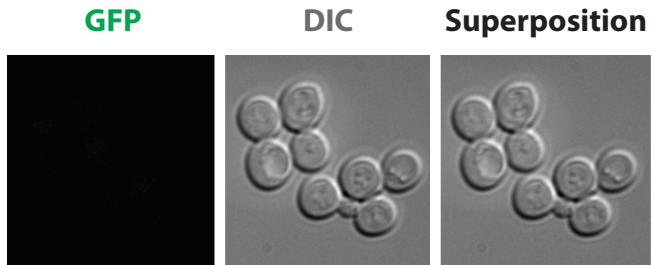
M

Mito- β 1-10
+
MetRS- β 11



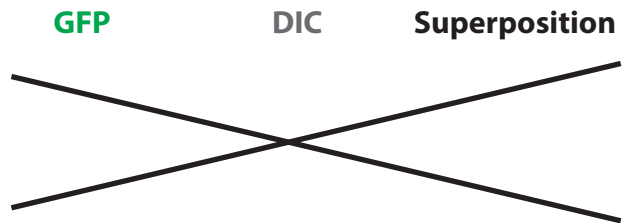
N

Mito- β 1-10
+
PheRS- β 11



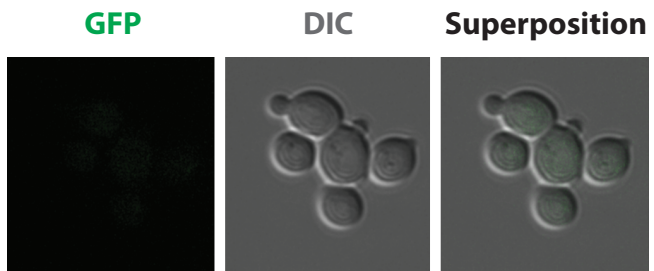
O

Mito- β 1-10
+
ProRS- β 11



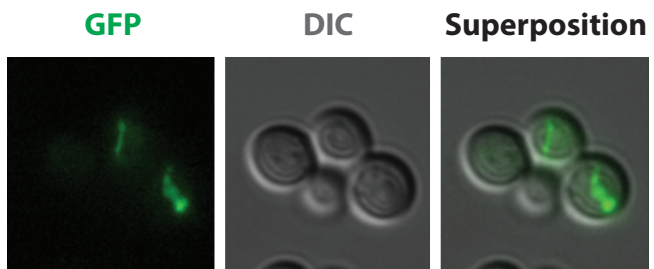
P

Mito- β 1-10
+
SerRS- β 11



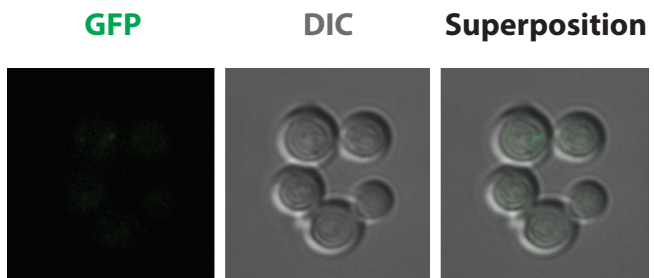
Q

Mito- β 1-10
+
ThrRS- β 11



R

Mito- β 1-10
+
TrpRS- β 11



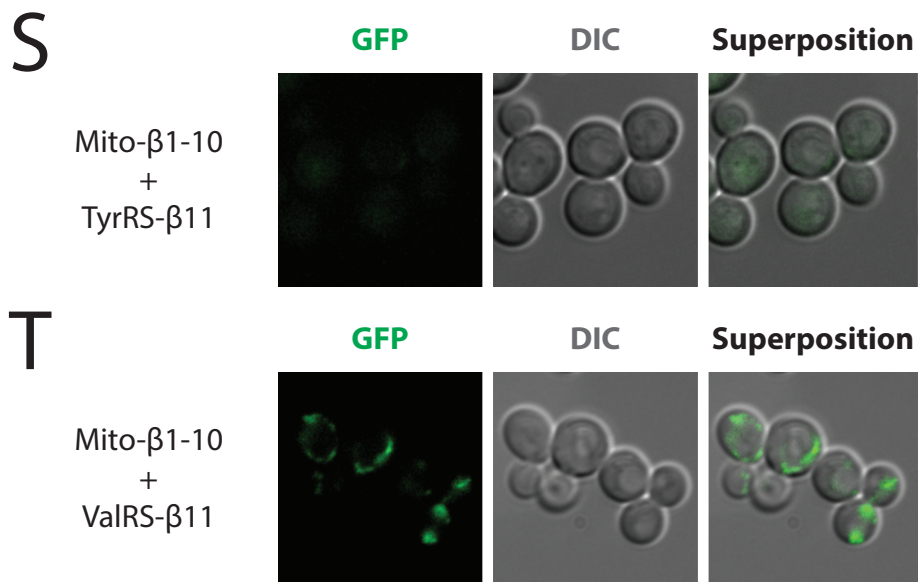


Figure 41 : Relocalisation mitochondriale des aaRS cytosoliques

La souche Mito- β 1-10 a été transformée avec les plasmides d'expression de 18 des 20 aaRS de levure étiquetées β 11 (A-T). L'apparition de fluorescence est visualisée par microscopie à épifluorescence et correspond à la relocalisation mitochondriale de l'aaRS- β 11. Les croix correspondent aux aaRS dont le sous-clonage n'a pas abouti.

de copies (p425-GPD-GlyRS- β 11 par exemple). Il est également envisagé de modifier directement le génome par recombinaison homologue et permettre la synthèse de cette protéine de fusion dans les conditions physiologiques car la régulation est potentiellement importante pour sa relocalisation.

Finalement, j'ai testé la relocalisation des quatorze autres aaRS disponibles. La majorité des enzymes ne sont pas importées dans la mitochondrie, dans les conditions testées (fermentation), du fait de l'absence de signal fluorescent (**Figure 41**). Un signal fluorescent est cependant détecté pour la GluRS, dont la localisation mitochondriale est avérée, mais également pour la ThrRS dont la localisation organellaire n'a jamais été suggérée auparavant (**Figure 41F & Q**). Ce résultat indique que cette dernière pourrait assurer une fonction au sein de ce compartiment. Sa relocalisation mitochondriale ne semble pas nécessaire pour son activité d'aminoacylation, puisqu'il existe une ThrRS mitochondriale fonctionnelle codée par un gène distinct et que la seule aaRS mitochondriale manquante est la CysRS. En effet, la ThrRS cytosolique ne reconnaît probablement pas l' m ARNt^{Cys}, suggérant qu'elle assure une fonction mitochondriale indépendante de sa fonction "primaire" d'aminoacylation.

Comme je l'ai décrit plus haut, l'activité de la ProRS mitochondriale n'a pas été validée et j'ai décidé d'exprimer cette protéine en fusion avec le fragment β 11 (AIM10- β 11). Après transformation dans la souche Mito- β 1-10, j'ai pu confirmer, par microscopie à épifluorescence, que la protéine de fusion est bien capable de se relocaliser dans les mitochondries (**Figure 42A**). La relocalisation de cette protéine, ainsi que la présence d'un domaine catalytique du type ProRS suggèrent qu'elle correspond à la ProRS mitochondriale, capable de synthétiser le Pro- m ARNt^{Pro}. Ce résultat doit néanmoins être confirmé par des tests d'aminoacylation sur des ARNt mitochondriaux de levure.

Finalement, j'ai également étudié la localisation d'Arc1, la protéine d'assemblage du complexe AME. Les localisations nucléaire et vacuolaire de cette enzyme ont été validées, mais aucune donnée concernant sa potentielle relocalisation mitochondriale n'a été publiée jusqu'à présent. Lors de l'expression de la protéine de fusion Arc1- β 11 dans la souche Mito- β 1-10, un signal fluorescent est détectable au niveau des mitochondries (**Figure 42B**). L'apparition d'un signal suggère que cette protéine peut se relocaliser dans ce compartiment. La détermination de la fonction mitochondriale d'Arc1 nécessitera de plus amples travaux et

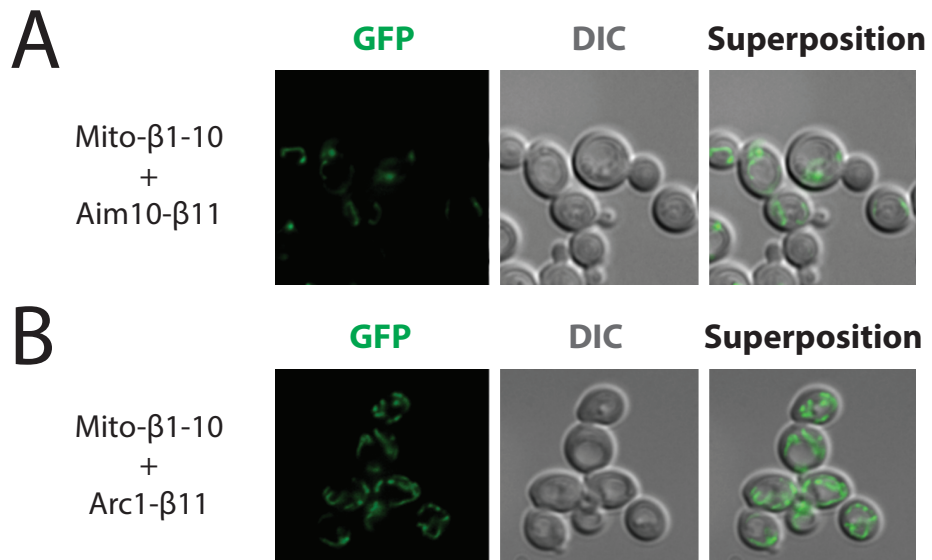


Figure 42 : Relocalisation mitochondriale d'Aim10 et d'Arc1

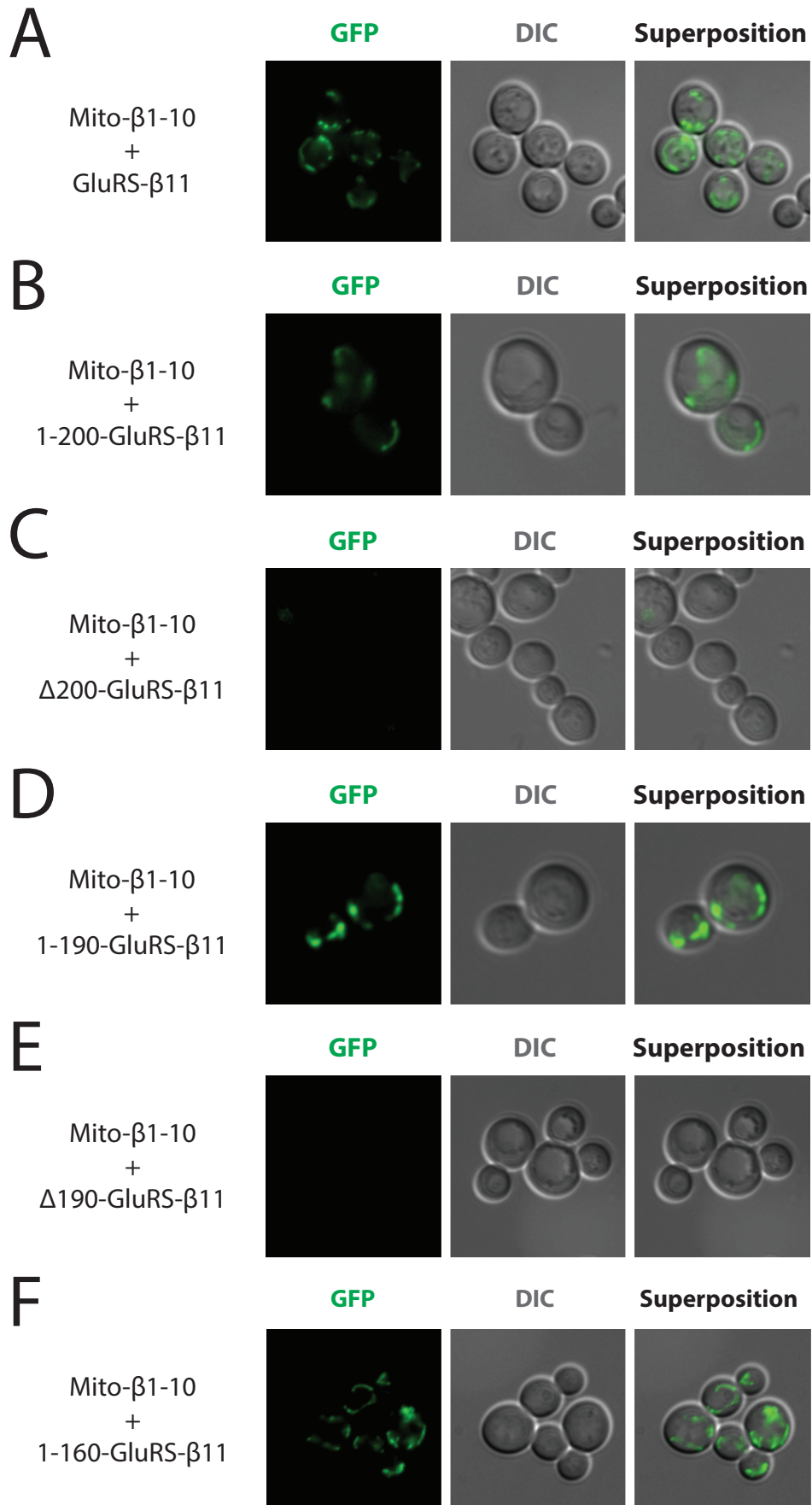
La souche Mito- β 1-10 a été transformée avec les plasmides d'expression d'Aim10- β 11, la ProRS mitochondriale putative (A) et Arc1- β 11, la protéine d'assemblage du MSC de levure (B). La fluorescence est visualisée par microscopie à épifluorescence.

passera notamment par la caractérisation de la région de la protéine contenant le MTS (décrit dans la partie suivante).

En perspectives de ces résultats et dans l'optique de déterminer si la relocalisation des aaRS est essentielle au sein de la mitochondrie et dans d'autres compartiments subcellulaires, nous pourrions utiliser une technique de rétention cytosolique de cette enzyme. En effet, au laboratoire, nous avons développé une stratégie qui vise à fusionner une protéine d'intérêt au niveau de la membrane vacuolaire par fusion avec Vph1. Nous pourrions alors créer 20 souches ayant chacune, une aaRS différente ancrée à la vacuole et *de facto*, incapable de rejoindre un autre compartiment subcellulaire. Ainsi, si la protéine assure une fonction essentielle dans un autre compartiment, la souche correspondante ne sera pas capable de se développer. Cependant, il est également possible que la fonction ne soit pas essentielle, mais uniquement importante. L'ancrage ne serait donc pas létal, mais pourrait potentiellement induire un phénotype observable. Il faudrait alors tester différentes conditions de croissance afin de trouver les conditions spécifiques où l'ancrage de l'aaRS est létal ou induit un phénotype observable. Les données obtenues permettraient de dégager des pistes concernant la relocalisation des aaRS cytosoliques dans la mitochondrie (en cas de létalité respiratoire), mais également dans d'autres compartiments cellulaires et surtout d'étoffer les connaissances de la communauté scientifique sur la capacité de relocalisation des enzymes appartenant à cette famille.

4. Etude des signaux d'import

Une facette très intéressante de l'outil de Split-GFP que j'ai développé est son utilisation pour la recherche des signaux d'adressage. Je me suis tout particulièrement focalisé sur la recherche des MTS cryptiques des aaRS cytosoliques, et autres protéines impliquées dans la formation des aa-ARNT, qui ont fait l'objet d'études précédentes au sein du laboratoire. Les signaux d'adressage mitochondriaux peuvent être de plusieurs types et leur positionnement dans la protéine donne beaucoup d'informations sur la localisation sub-mitochondriale de la protéine et la voie d'import mitochondriale utilisée pour la translocation dans l'organelle (**Figures 18 & 19**). Les protéines matricielles sont en effet caractérisées par une hélice amphiphile, retrouvée dans la région N-ter de la protéine, qui est reconnue par les complexes TOM et TIM permettant la translocation de la protéine dans la matrice de ce



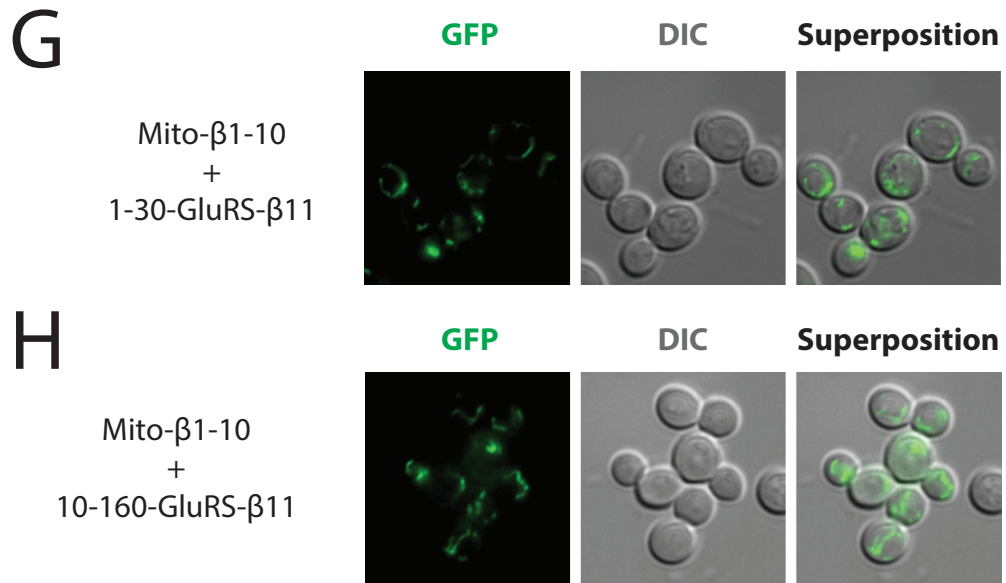


Figure 43 : Délimitation du signal d'import mitochondrial de la GluRS cytosolique

Pour délimiter le signal d'import de la GluRS, la souche Mito- β 1-10 a été transformée avec les plasmides d'expression de la GluRS (A) ou de variants de sa région N-ter (B-H). L'apparition de fluorescence est visualisée par microscopie à épifluorescence.

compartiment. Comme pour les autres protéines matricielles, les signaux d'import des aaRS sont invariablement retrouvés au niveau de ces extrémités N-ter et plus particulièrement dans les 50 premiers résidus. J'ai donc décidé de délimiter ces signaux pour différentes protéines dont l'import a été confirmé par l'utilisation de l'outil Split-GFP.

a. Caractérisation du signal d'import de la GluRS

Au laboratoire, nous avons montré que la GluRS se relocalise dans la mitochondrie pour participer à la formation du Gln-_mARNt^{Gln} *via* une voie alternative faisant intervenir l'AdT GatFAB (Frechin et al., 2009b). Cette protéine se retrouve dans la matrice de ce compartiment et doit, en théorie, posséder un MTS au niveau de son extrémité N-ter. Lors de l'étude bio-informatique des signaux putatifs d'adressage mitochondriaux, deux logiciels sur les trois utilisés prédisent la présence d'un MTS au sein de cette protéine. Selon TargetP1.1, ce signal est retrouvé au niveau des 45 premiers aa (Bader et al., 2017). Les données disponibles dans la littérature suggèrent que le MTS est retrouvé dans les 200 premiers aa et plus précisément entre les résidus 190 et 199 (Frechin et al., 2009b). Tout d'abord, j'ai décidé d'exprimer les 200 premiers aa de la GluRS associés au fragment β 11 (1-200-GluRS- β 11). La souche Mito- β 1-10 exprimant la protéine 1-200-GluRS- β 11 présente un signal fluorescent semblable à celui détecté pour la protéine entière (**Figure 43A et B**). De plus, lorsque les 200 premiers résidus de la GluRS sont absents (Δ 200-GluRS- β 11), aucun signal n'est visualisable (**Figure 43C**). Conformément aux données de la littérature, le signal d'import est bien présent au niveau des 200 aa N-ter de la GluRS. J'ai ensuite tenté de confirmer que le signal d'import est situé entre les résidus 190 et 200. Pour ce faire, j'ai exprimé la protéine 1-190-GluRS- β 11, où uniquement les 190 premiers aa de la GluRS sont fusionnés au fragment β 11. De manière contradictoire avec les données publiées précédemment, un signal fluorescent est détecté (**Figure 43D**). Par ailleurs, cette observation est confirmée par l'absence de reconstitution mitochondriale de la GFP lors de l'expression du mutant de délétion des 190 premiers aa : Δ 190-GluRS- β 11 (**Figure 43E**). Ces résultats suggèrent que le MTS ne se trouvent pas dans la région 190-200, mais qu'il est situé plus en amont dans la séquence. Par conséquent, j'ai décidé de délimiter grossièrement la région contenant le MTS. Pour cela, j'ai fusionné les 160 ou 30 premiers résidus de la GluRS au fragment β 11 (1-160-GluRS- β 11 et 1-30-GluRS- β 11). Quelle que soit la construction utilisée, un signal de fluorescence est observable (**Figure 43F & G**). Ainsi, le signal d'import est présent au niveau des 30 premiers aa de la GluRS.

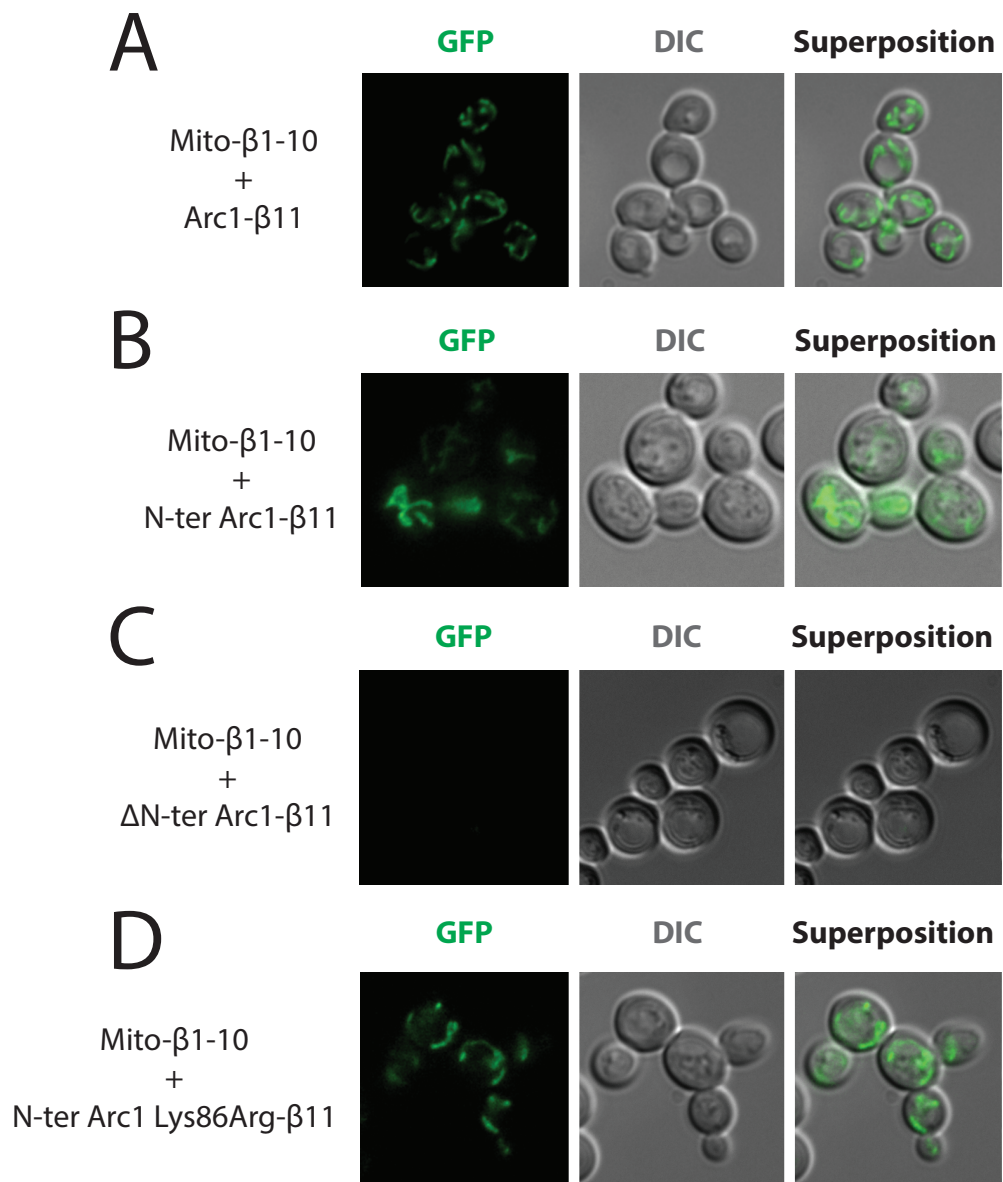


Figure 44 : Détermination de la région impliquée dans l'import mitochondrial d'Arc1

La souche Mito- β 1-10 a été transformée avec les plasmides d'expression d'Arc- β 11 (A) et de différents variants de la protéine (B-C). La région N-ter est biotinylée au niveau de son résidu Lys86 et la fonction de cette modification est inconnue. Le mutant N-ter Arc1 Lys86Arg- β 11 a également été transformé dans la souche Mito- β 1-10 pour déterminer si la modification de cet aa permet de moduler la localisation mitochondriale d'Arc1 (D). La fluorescence est visualisée par microscopie à épifluorescence.

Lors de l'expression de variants de la GluRS par l'équipe de Georges Simos, les neuf premiers résidus sont toujours conservés et la viabilité est assurée en fermentation. La viabilité respiratoire n'a cependant jamais été testée par cette équipe de recherche. J'ai alors voulu savoir si les dix premiers résidus sont importants pour la relocalisation mitochondriale de la GluRS. Pour cela, j'ai utilisé le vecteur contenant 1-160-GluRS- β 11 pour effectuer la délétion des résidus 2 à 10 et créer le variant 10-160-GluRS- β 11. Après transformation de la souche Mito- β 1-10 avec le plasmide d'expression de cette construction, j'ai pu confirmer l'import de cette protéine dans la mitochondrie par microscopie (**Figure 43H**). L'ensemble de ces résultats suggèrent la présence du MTS de la GluRS entre les résidus 10 et 30. En étudiant la structure cristallographique des régions N-ter d'interaction entre la GluRS et Arc1 (Simader et al., 2006), je me suis aperçu que la région composée des résidus 1 à 10 forme un feuillet, ne coïncidant pas avec la structure hélicoïdale des MTS N-ter. De manière plus intéressante, les résidus 20 à 30 adoptent une structure en forme d'hélice et du fait de la présence potentielle du MTS dans cette région, cette hélice pourrait correspondre au signal d'import de la GluRS.

A court terme, nous voudrions valider la présence du MTS de la GluRS au niveau des résidus 10-30. Pour cela, il faudrait effectuer une mutagenèse de ce dernier pour déterminer les résidus importants pour l'import mitochondrial de cette aaRS. Finalement, par mutagenèse de ces résidus, nous pourrions valider l'incapacité de relocalisation de la GluRS par l'utilisation de la souche Mito- β 1-10, ainsi qu'une létalité en métabolisme respiratoire par des techniques de chasse au 5FOA ou dissection de tétrades. En effet, une ségrégation 4/0 devrait être obtenue sur milieu fermentaire car l'élimination du MTS de la GluRS n'impacte probablement pas l'activité cytosolique de l'enzyme, tandis qu'une ségrégation 2/2 devrait être observée sur milieu respiratoire, validant l'incapacité de cette enzyme de se relocaliser dans la mitochondrie pour assurer la formation du Glu-_mARNt^{Gln}.

b. Caractérisation de la région contenant le signal d'import d'Arc1

La protéine Arc1, permettant l'assemblage du complexe AME, est capable de se relocaliser dans la mitochondrie (**Figure 44A**). L'import organellaire de cette protéine est un résultat étonnant, car elle est déjà retrouvée dans plusieurs compartiments subcellulaires (noyau, cytosol et à la surface de la vacuole). Pour déterminer la région d'Arc1 impliquée dans

sa relocalisation mitochondriale, j'ai utilisé des plasmides construits par Marine Hemmerlé, que j'ai encadrée lors de son stage de Master 2 (2016-17), et qui avait élaboré différents variants d'Arc1, fusionnés au fragment $\beta 11$, pour étudier leur localisation vacuolaire à l'aide de la Split-CFP. Supposant une localisation matricielle d'Arc1, j'ai décidé de débiter l'étude avec la région N-ter de cette protéine (Nter-Arc1- $\beta 11$), ainsi que le mutant de délétion de cette région (Δ Nter-Arc1- $\beta 11$). Ces protéines ont été exprimées dans la souche Mito- $\beta 1-10$ et les images de microscopie montrent que la région N-ter d'Arc1 est importée dans la mitochondrie, tandis que la protéine dépourvue de cette région ne l'est pas (**Figure 44B & C**). Ces résultats confirment qu'Arc1 peut se relocaliser dans la mitochondrie et que la région impliquée dans son import se situe au niveau des 131 premiers résidus de la protéine (organisation modulaire d'Arc1, **Figure 13**).

La région N-ter d'Arc1 peut être modifiée par la biotine:protéine ligase Bpl1 au niveau de la Lysine 86 (Kim et al., 2004a). La fonction biologique de cette modification n'ayant pas encore pu être établie, je me suis demandé si elle pouvait être impliquée dans l'import de cette protéine dans la mitochondrie. Pour ce faire, j'ai utilisé un plasmide d'expression de la région N-ter d'Arc1 possédant la mutation Lys86Arg (empêchant sa biotinylation). La présence d'un signal fluorescent dans la souche Mito- $\beta 1-10$ exprimant cette construction indique que la biotinylation de la Lys86 n'est pas requise pour la localisation mitochondriale de la région N-ter d'Arc1 (**Figure 44D**), mais également de la protéine Arc1 entière et mutée (Lys86Arg) (résultats non montrés).

Comme décrit je l'ai décrit dans l'introduction, la région N-ter d'Arc1 contient un domaine *GST-like* impliqué dans l'interaction avec la GluRS et la MetRS. Ainsi, il serait intéressant de déterminer précisément la région contenant le MTS d'Arc1 par création de mutants de délétion de cette protéine et visualisation de leurs capacités de relocalisation mitochondriale. En effet, la GluRS étant également importée dans la mitochondrie, l'import de ces deux protéines pourrait être simultanée et l'une d'entre elles pourrait induire leur import mitochondrial ou tout du moins, leur relocalisation à la surface des mitochondries. Les résidus impliqués dans l'interaction entre ces deux partenaires étant connus (Karanasios et al., 2007), il serait facile de tester cette hypothèse. En effet, les aa impliqués dans l'interaction avec Arc1 sont compris entre les résidus 125 et 164 de la GluRS. Or, le signal d'import de cette protéine est retrouvé entre les résidus 10 et 30 et l'interaction avec Arc1 ne devrait, en

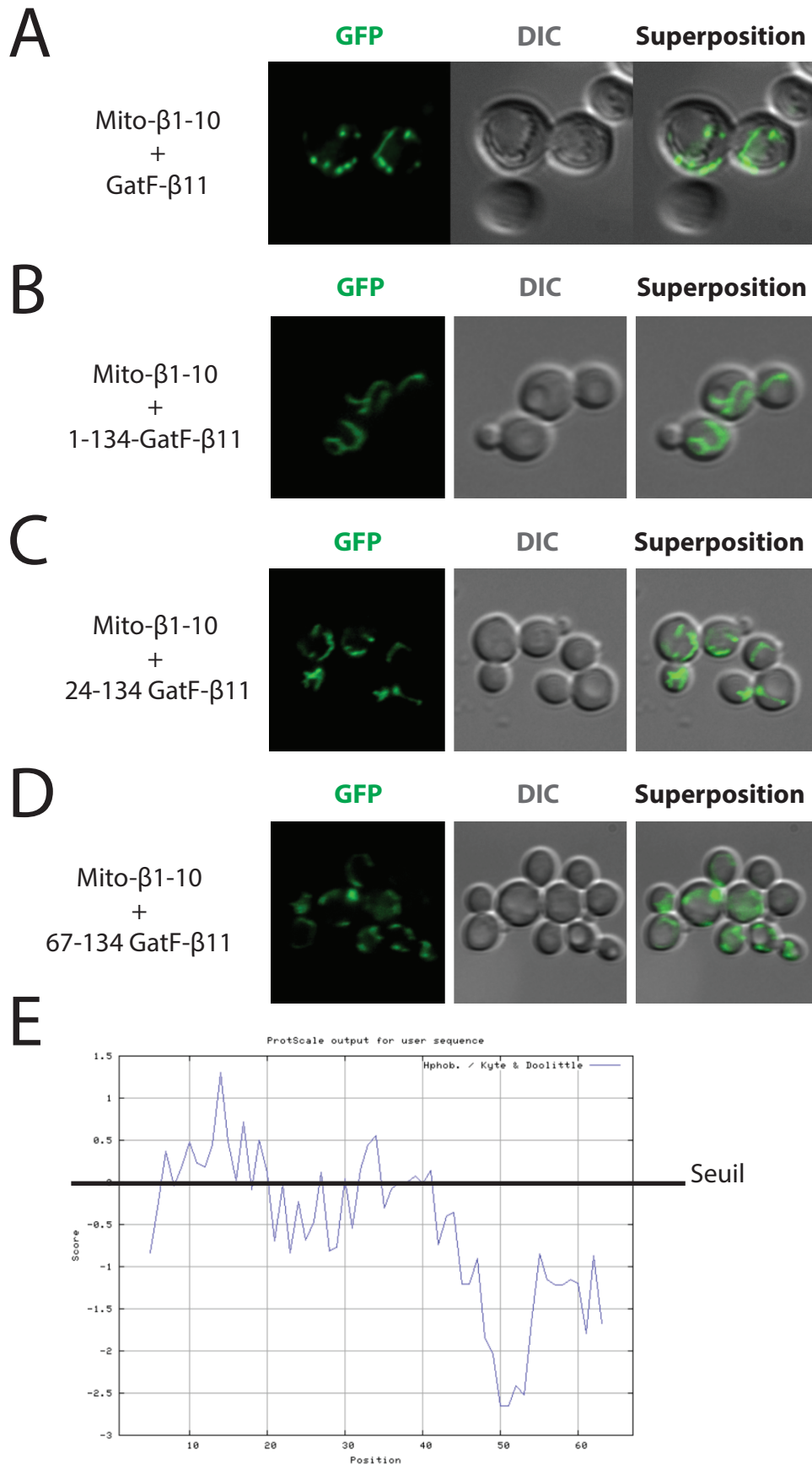


Figure 45 : Délimitation de la région d'import mitochondrial de GatF

La souche Mito- β 1-10 a été transformée avec les plasmides d'expression de GatF- β 11 (A) et de différents variants de délétion étiquetés β 11 (B-D). La reconstitution de la GFP active est visualisée par microscopie à épifluorescence. (E) L'hydrophobicité des résidus 67-134 a été testée bio-informatiquement par utilisation du logiciel <http://web.expasy.org/protscale/>. Les pics au dessus du seuil correspondent à des résidus hydrophobes.

théorie, pas intervenir dans l'import de la GluRS dans la mitochondrie ; elle peut néanmoins médier sa relocalisation au niveau de la MME. Au niveau d'Arc1, les résidus impliqués dans l'interaction avec la GluRS sont retrouvés au niveau de la région N-ter (entre les résidus 55 et 104) ; la région putative contenant le MTS contient ces résidus. L'association d'Arc1 avec la GluRS pourrait ainsi médier (i) sa relocalisation à la MME et/ou (ii) son import dans la matrice mitochondriale. Pour tester ces différentes hypothèses, plusieurs solutions sont envisagées. Elles nécessiteront la construction de souches Mito- β 1-10 possédant (i) une GluRS incapable de se relocaliser dans la mitochondrie (Δ 30-GluRS par exemple), (ii) une GluRS ne pouvant plus interagir avec Arc1 ou (iii) une protéine Arc1, incapable de se lier à la GluRS. L'expression d'Arc1- β 11 nous renseignerait donc sur la capacité de la GluRS à favoriser l'import mitochondrial d'Arc1, tandis que l'expression de la GluRS- β 11 affirmerait ou réfuterait l'hypothèse inverse. Quelle que soit les résultats obtenus avec ces souches, il est nécessaire de délimiter plus précisément la région contenant le signal d'import d'Arc1 et de les valider dans une souche où Arc1- β 11 est dépourvue de ces résidus.

c. Caractérisation de la région impliquée dans l'import de GatF

La voie de formation du Gln- m ARNt^{Gln} fait intervenir la GluRS cytosolique se relocalisant dans la mitochondrie et l'AdT mitochondriale GatFAB (Frechin et al., 2009b). Cette dernière contient, entre autres, la s-u GatF, dont le mécanisme d'import est encore inconnu. Au laboratoire, la structure de l'AdT mitochondriale a été élucidé au début de ma thèse (Araiso et al., 2014). Cependant la région impliquée dans l'import de GatF n'est pas délimitée. Cette protéine fait une taille de 183 aa et se relocalise probablement dans la matrice mitochondriale, mais pourrait également, après translocation et assemblage avec les deux autres s-u de l'AdT, être associée à la MMI. Afin de caractériser le MTS de GatF, le Dr. Araiso avait construit plusieurs fragments de cette protéine fusionnés au fragment β 11, que j'ai utilisé pour transformer la souche Mito- β 1-10. Les MTS des protéines matricielles étant retrouvés dans la région N-ter, j'ai étudié la relocalisation mitochondriale des variants 1-134-GatF- β 11, 24-134-GatF- β 11 et 67-134-GatF- β 11. L'expression de GatF- β 11 induit l'émission de fluorescence confirmant bien sa relocalisation mitochondriale (**Figure 45A**). De manière inattendue, les souches exprimant les trois variants de délétion présentent une fluorescence mitochondriale, validant leur relocalisation au sein de ce compartiment (**Figure 45B, C & D**). Contrairement à ce que l'on pensait, la région N-ter n'est donc pas impliquée dans la

relocalisation de GatF et le MTS est plutôt retrouvé dans la région centrale ou C-ter de la protéine. La présence du MTS entre les résidus 67 et 134 est atypique et pourrait suggérer que GatF n'est pas importée dans la matrice mitochondriale. Les protéines incorporées dans la MMI possèdent un MTS dans leur région centrale, caractérisé par la présence d'une région hydrophobe suivie d'une hélice amphiphile (**Figure 18**). J'ai tenté de confirmer la présence d'une région hydrophobe au sein de la région 67-134. Pour cela, j'ai effectué une recherche avec un logiciel de prédiction (<http://web.expasy.org/protscale/>). Les résultats sont présentés dans la **Figure 45E** et suggèrent la présence d'une région hydrophobe entre les résidus 77 et 87. De plus, au sein de la structure 3D, cette région est suivie d'une hélice, confortant l'hypothèse d'un signal interne permettant l'intégration de GatF au sein de la MMI.

Néanmoins, cette protéine s'associe avec les s-u GatA et GatB pour former l'AdT trimérique GatFAB. Par conséquent, il est crucial de déterminer plus en détail la localisation sub-mitochondriale des deux autres partenaires de GatF. En effet, la région 67-134 est également impliquée dans l'interaction avec GatA et le MTS de cette dernière favorise potentiellement l'import de GatF dans la mitochondrie. Cette hypothèse est pour l'instant très spéculative et une étude approfondie devra être réalisée pour confirmer la localisation membranaire de GatF et potentiellement du complexe en entier. D'ailleurs, lors de l'étude qui a conduit à l'établissement de la structure 3D de l'AdT, nous avons remarqué qu'en comparant toutes les AdT dont la structure avait été établie, la GatFAB de mitochondrie de levure est la seule à avoir une charge de surface électropositive. Cette caractéristique structurale pourrait conforter l'hypothèse de l'ancrage de ce complexe trimérique à une membrane, dont le potentiel de surface est électronégatif (Araiso et al., 2014).

5. Etude de la relocalisation mitochondriale de protéines hétérologues

L'étude de la physiologie de la levure a révélé une forte conservation de divers processus biologiques entre cet unicellulaire et les autres eucaryotes. En effet, environ 30 % des gènes de levure ont un homologue humain. La levure a également été un hôte de choix pour l'expression hétérologue de protéines humaines afin d'étudier divers processus biologiques essentielles, mais également afin de comprendre certaines pathologies. En outre, 50 % de la totalité des ADN complémentaires (gènes humains sans les introns) ont pu être exprimés chez la levure et uniquement 6 % sont toxiques (Sekigawa et al., 2010). L'expression

hétérologue a permis notamment l'étude de la famille des *poly ADP ribose polymerase* (PARP), utilisée comme cible pour le cancer du sein ou des ovaires (Perkins et al., 2001), la découverte de gènes potentiellement impliqués dans des maladies neurodégénératives (Tardiff et al., 2013) ou l'étude des facteurs de risques de la maladie d'Alzheimer (Treusch et al., 2011). Cet unicellulaire a prouvé, au fil du temps, qu'il est un modèle fiable pour l'étude des processus se déroulant chez l'Homme. Les aaRS appartiennent à ces familles d'enzymes fortement conservées entre la levure et l'Homme. Cette conservation ne se limite pas seulement à l'homologie de séquence et d'activité, puisque ces deux organismes contiennent des MSC, favorisant l'activité d'aminocyclation des aaRS et servant de plateformes de rétention pour des aaRS se relocalisant dans d'autres compartiments subcellulaires. La levure possède un MSC plus petit que celui de l'Homme, mais les enzymes retrouvées dans celui-ci le sont également au sein du MSC humain (GluRS et MetRS). En parallèle, la protéine d'ancrage Arc1 a des homologues fonctionnel et structural, retrouvés au sein du MSC, chez l'Homme (p18 et p43 respectivement). Les deux aaRS communes aux deux MSC, mais également d'autres, présentes uniquement dans le MSC de l'Homme, ont la capacité d'assurer des fonctions additionnelles dans le cytosol, ou dans d'autres compartiments cellulaires. Ces fonctions peuvent également être conservées entre l'Homme et la levure. Ainsi, la MetRS humaine, comme son homologue de levure, a la capacité de se relocaliser dans le noyau dans des conditions spécifiques pour permettre la régulation de la transcription (Frechin et al., 2014; Ko et al., 2000). De plus, la LeuRS peut se relocaliser au niveau de la vacuole et du lysosome, chez la levure et l'Homme respectivement, afin de jouer un rôle de senseur de leucine (Bonfils et al., 2012; Han et al., 2012). Etudier la relocalisation des aaRS de levures, humaines, mais aussi d'autres organismes comme les plantes pourrait faciliter la compréhension des fonctions additionnelles de celles-ci. Dans cette optique, j'ai tenté d'exprimer des protéines venant de différentes origines pour vérifier si elles sont exprimées chez la levure et surtout si elles peuvent se relocaliser dans la mitochondrie. J'ai procédé comme pour les aaRS et les séquences codantes ont été sous-clonées dans la série de plasmides du système Gateway®. J'ai ensuite exprimé ces enzymes en fusion avec le fragment β 11 dans la souche Mito- β 1-10.

Dans un premier temps, j'ai voulu déterminer si des protéines humaines peuvent être exprimées chez la levure et si elles sont capables de se relocaliser dans la mitochondrie. Pour cela, j'ai tout d'abord utilisé deux protéines : la GluRS mitochondriale humaine (*Hsa-mGluRS-*

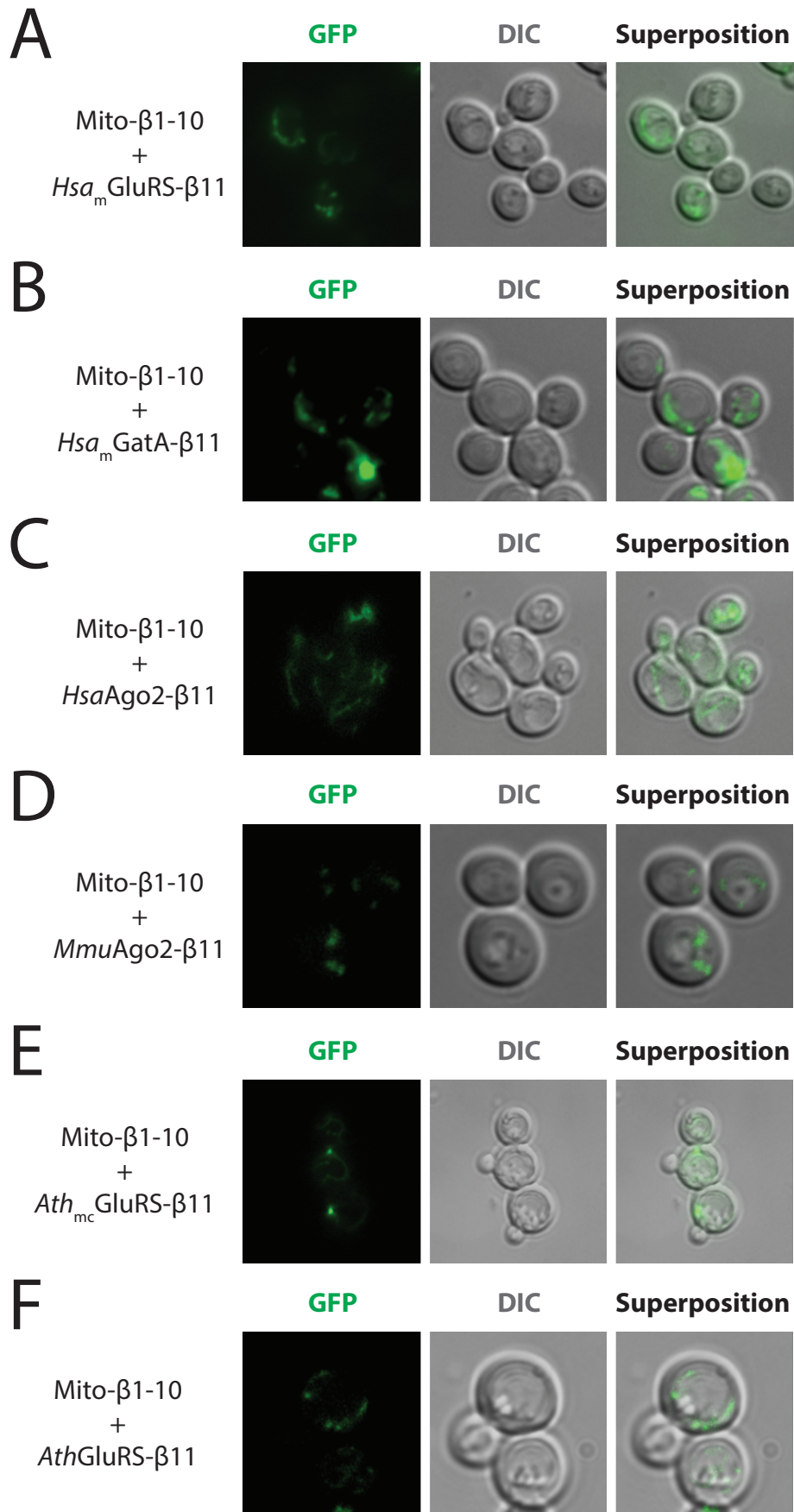


Figure 46 : Relocalisation mitochondriale de protéines hétérologues

La souche Mito- β 1-10 est transformée avec des plasmides d'expression de protéines humaines (*Hsa_mGluRS- β 11*, *Hsa_mGatA- β 11* et *HsaAgo2- β 11*), de souris (*MmuAgo2- β 11*) et de plantes (*Ath_{mc}GluRS- β 11* et *AthGluRS- β 11*) (A-F). La capacité d'import mitochondrial de ces protéines est visualisée par microscopie à épifluorescence.

β 11) et la s-u GatA de l'AdT mitochondriale humaine (*Hsa-mGatA- β 11*). Lors de l'expression de ces deux témoins dans la souche Mito- β 1-10, un signal fluorescent est visualisable (**Figure 46A & B**). La séquence des gènes correspondants n'ayant pas été optimisée pour l'usage des codons de levure, l'apparition d'un signal fluorescent confirme qu'elles sont exprimées, qu'elles se replient correctement, et surtout que la machinerie d'import mitochondriale de levure est capable de les transloquer dans la matrice. L'étude de la relocalisation mitochondriale de protéines hétérologues humaines est possible grâce à l'utilisation de l'outil de Split-GFP mitochondrial.

En parallèle, j'ai également exprimé les versions étiquetées β 11 des protéines Argonaute 2 (Ago) de souris (*MmuAgo2*) et humaine (*HsaAgo2*). L'import mitochondrial des protéines Ago a été suggéré par certains de nos collaborateurs, mais aucun mécanisme d'ARN interférence n'a été mis en évidence dans ce compartiment. Lorsque ces protéines sont exprimées dans la souche Mito- β 1-10, il y a apparition d'un signal fluorescent (**Figure 46C & D**). Ces résultats suggèrent que les protéines Ago2 humaine et de souris possèdent des signaux d'adressage favorisant leur relocalisation mitochondriale. Les perspectives qui en découlent sont très intéressantes et l'étude de la relocalisation mitochondriale de ces protéines pourrait permettre de découvrir et comprendre le mécanisme d'interférence au niveau de ce compartiment. Des études d'ancrage dans le cytosol devront ainsi être réalisées sur des cellules humaines pour caractériser plus en détail la fonction de ces protéines dans la mitochondrie et déterminer s'il y a accumulation de certains ARN mitochondriaux qui sont des cibles potentielles pour le système Argonaute. De plus, avec la souche Mito- β 1-10, nous pourrions délimiter précisément le MTS de ces enzymes. Il serait alors possible de créer des lignées cellulaires possédant des protéines Ago2 non importables, ce qui faciliterait l'étude de ce mécanisme dans la mitochondrie.

Finalement, j'ai voulu savoir si l'outil était également utilisable pour l'étude de l'import mitochondrial de protéines issues de plantes. J'ai amplifié et exprimé la GluRS mitochondriale et chloroplastique (*Ath_{MC}GluRS*) et la GluRS cytosolique (*AthGluRS*) d'*A. thaliana* en fusion avec le fragment β 11. Lors de l'expression de la *Ath_{MC}GluRS- β 11*, un signal de fluorescence est observable ; cette enzyme peut donc être importée dans les mitochondries de levure. De manière intéressante, l'expression de la GluRS cytosolique de cet organisme (*AthGluRS- β 11*) induit également la reconstitution de la GFP, malgré une absence de données dans la

littérature décrivant cette possibilité (**Figure 46E & F**). Chez les plantes, les MTS sont très différents de ceux retrouvés chez la levure et il existe des signaux d'import ambigus induisant l'import de protéines dans la mitochondrie et le chloroplaste. De manière intéressante, les complexes d'import mitochondriaux de levure sont capables de reconnaître ces signaux ambigus et permettent de transloquer la *Ath_{MC}GluRS* dans ce compartiment.

L'apparition d'un signal fluorescent pour l'ensemble de ces protéines hétérologues est très intéressante. En effet, malgré l'absence d'optimisation dans l'usage des codons, ces résultats suggèrent que ces protéines sont (i) bien exprimées et (ii) capables de se relocaliser dans la mitochondrie. Finalement, tous ces résultats réunis montrent la puissance de l'outil pour étudier l'import des protéines mitochondriales ou à double localisation cytosolique et mitochondriale d'autres espèces.

Partie IV : Etude de la GluProRS^{N1}
humaine et de sa fonction
mitochondriale

1. Contexte de la recherche

Comme je l'ai décrit dans l'introduction, la formation du Gln-ARNt^{Gln} organellaire est catalysée par une voie indirecte faisant intervenir une GluRS ND et une AdT organellaire (**Figure 10**). Lorsque l'on compare les mécanismes de formation de ces aa-ARNt au travers de l'arbre phylogénétique, on constate que même si la plupart des eucaryotes utilisent cette voie, il existe des différences marquantes concernant les enzymes y participant. En effet, chez la levure, cette voie fait intervenir la GluRS cytosolique, importée dans la mitochondrie (Frechin et al., 2009b), tandis que chez l'Homme, c'est la GluRS mitochondriale qui génère le Glu-mARNt^{Gln} (Nagao et al., 2009). Il est à noter que la GluRS mitochondriale de levure n'est pas capable d'aminocycler l'mARNt^{Gln} (Rinehart et al., 2005). La différence majeure entre les deux mécanismes est donc l'origine de la GluRS ND qui génère l'intermédiaire mésacylé. Ainsi, chez la levure, la GluRS cytosolique est capable de reconnaître l'ARNt^{Glu} cytosolique (cARNt^{Glu}) et l'mARNt^{Gln}, alors que la GluRS mitochondriale ND humaine reconnaît deux espèces d'mARNt (mARNt^{Glu} et mARNt^{Gln}) et assure deux activités d'aminocyclation dans le même compartiment.

A l'heure actuelle, les déterminants d'identité indispensables à la reconnaissance de l'ARNt^{Glu} par la GluRS, ont uniquement été étudiés chez les procaryotes. Cependant, ces données permettent de spéculer quant à la conservation de ces déterminants chez les eucaryotes. De la sorte, la GluRS mitochondriale ND humaine est semblable à l'aaRS de *B. subtilis*. Au sein des ARNt reconnus par cette enzyme bactérienne, la position 34 est primordiale, car les ARNt possédant une U34 sont aminocyclés, tandis qu'une C34 empêche la fixation de l'aa (Sekine et al., 2001). De plus, la GluRS mitochondriale discriminante de levure ressemble fortement à son homologue d'*E. coli*. Les données de la littérature indiquent que cette dernière est capable de discriminer les ARNt^{Glu} et ^{Gln} du fait de la présence, dans l'ARNt^{Glu}, d'un bras D plus long et d'une boucle variable plus courte. Cependant, lorsque le bras accepteur de l'ARNt^{Glu} est transféré au reste de l'ARNt^{Gln}, l'ARNt chimère résultant peut être aminocyclé par la GluRS, suggérant qu'en plus de ces déterminants, un déterminant crucial se trouve au niveau du bras accepteur de l'ARNt^{Glu} (Chang and Hendrickson, 2009). Pour finir, la GluRS cytosolique de levure ne semble pas avoir d'équivalent chez les procaryotes. La spécificité de reconnaissance des cARNt^{Glu} et ^{Gln} ou des mARNt^{Glu} et ^{Gln} n'est pas permise par la longueur du bras D et de la boucle variable ou la première position de l'anticodon (position 34). Cependant, les ARNt aminocyclés par la GluRS cytosolique

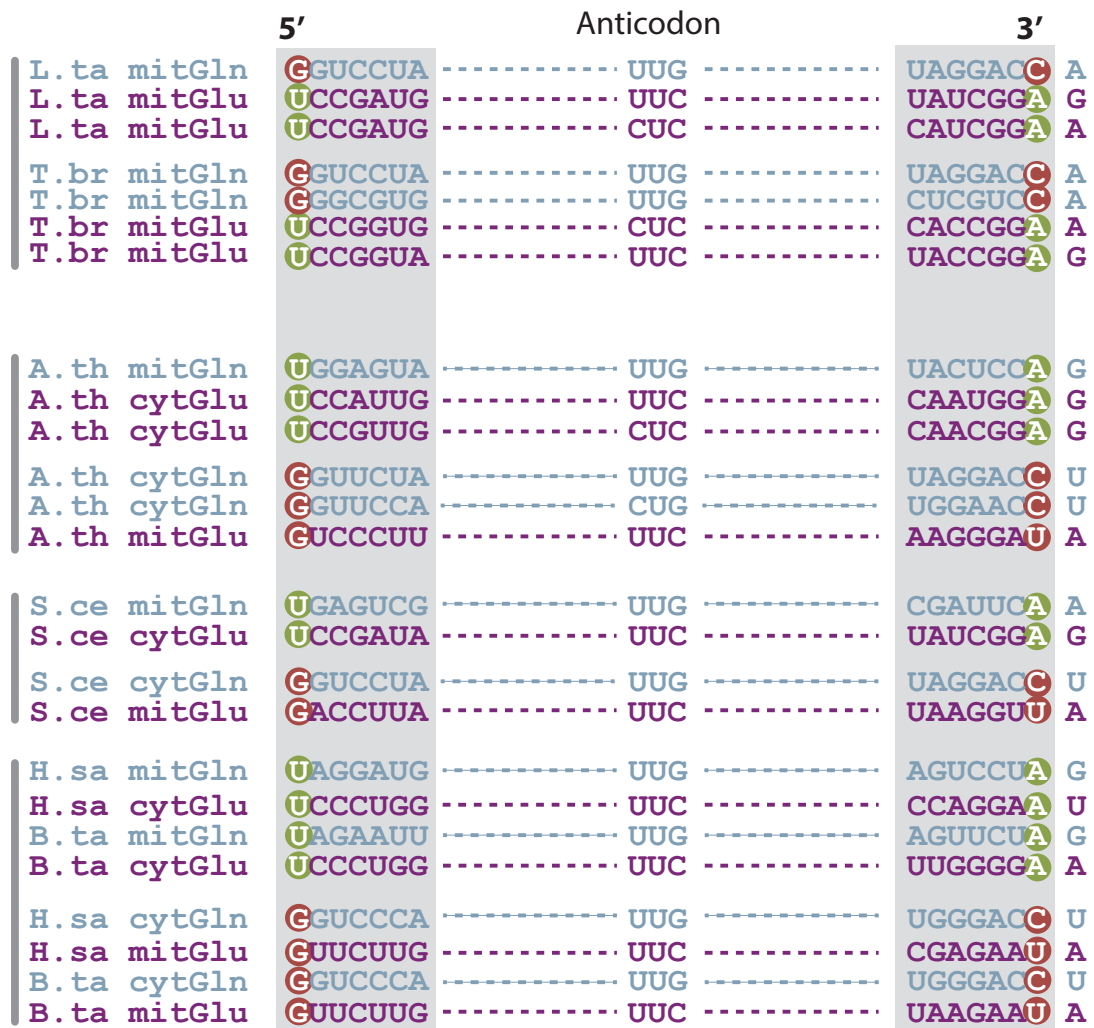


Figure 47 : Alignement multiple des ARNt^{Glu} et ^{Gln} cytosolique et organellaire de diverses espèces

Alignement de plusieurs séquences d'ARNt^{Glu} (violet) et ^{Gln} (bleu clair) cytosoliques et mitochondriaux. Les bases présentes dans le bras accpeteur sont grisées. La première paire de bases du bras accepteur (1-72) est mise en évidence par des boules vertes (pour les paires de bases U-A) et par des boules rouges (pour les paires de bases G-U/C). Elle contribue probablement à la reconnaissance spécifique de l'ARNt par l'aRS correspondant.

B. ta : *Bos taurus* ; *S. ce* : *Saccharomyces cerevisiae* ; *A. Th* : *Arabidopsis thaliana* ; *T. br* : *Trypanosoma brucei* ; *L. Ta* : *Leishmania tarentolae*.

présentent tous une paire de bases U_1-A_{72} ($cARnt^{Glu}$ et $mARnt^{Gln}$), tandis que cette paire de bases est remplacée dans $mARnt^{Glu}$, reconnu par la GluRS mitochondriale de levure, par une paire de bases G_1-U_{72} .

Au laboratoire, nous avons alors entrepris l'alignement des séquences des $ARnt^{Glu}$ et $ARnt^{Gln}$ cytosoliques et mitochondriaux de différentes espèces (**Figure 47**). Le constat est flagrant : les $cARnt^{Glu}$ présentent tous une paire de bases U_1-A_{72} au niveau du bras accepteur, qui est remplacée par G_1-C_{72} dans tous les $cARnt^{Gln}$. A l'inverse, lorsque l'on analyse les séquences des $mARnt$, on constate que la paire de bases U_1-A_{72} est retrouvée dans le bras accepteur des $mARnt^{Gln}$, alors que les $mARnt^{Glu}$ possèdent la paire de bases G_1-U_{72} (**Figure 47**). En résumé, chez tous les eucaryotes dont la formation du $Gln-mARnt^{Gln}$ est catalysée par voie indirecte, les $cARnt^{Glu}$ et $mARnt^{Gln}$ possèdent les mêmes déterminants. Il semble donc peu probable qu'une enzyme capable d'aminoacyler l' $ARnt^{Glu}$ dans un compartiment, puisse également charger l' $ARnt^{Gln}$ dans ce même compartiment. Et pourtant, les données de la littérature suggèrent que la GluRS organellaire en est capable chez l'Homme, mais également chez les plantes, où une GluRS mitochondriale et chloroplastique assurent la charge des deux $mARnt$ (Pujol et al., 2008). Pour cette dernière activité, la capacité d'aminoacylation a été mesurée sur des extraits mitochondriaux totaux, et non sur des enzymes purifiées. Comme nous l'avons vu dans la partie précédente, la GluRS cytosolique d'*A. thaliana* est retrouvée dans les mitochondries et il n'est pas à exclure que ce soit cette enzyme qui catalyse la formation du $Glu-mARnt^{Gln}$ dans ce compartiment.

Nous avons décidé de tester l'hypothèse selon laquelle la voie indirecte de formation du $Gln-ARnt^{Gln}$ organellaire implique l'import de la GluRS cytosolique chez tous les organismes utilisant cette voie, et notamment chez l'Homme. En effet, si l'on compare les GluRS humaine et de levure, on constate qu'elles ont des points communs, comme leur capacité d'incorporation au sein de MSC et de libération, afin d'assurer des fonctions additionnelles. Cependant, il existe une différence majeure : l'aaRS humaine est bifonctionnelle et possède, en plus de son activité GluRS, une activité de ProRS (GluProRS). L'activité ProRS dans la mitochondrie étant déjà existante, l'import d'une seconde enzyme capable d'assurer cette fonction serait donc étonnant. Cependant, s'il existait une version tronquée de cette protéine présentant uniquement l'activité GluRS, celle-ci serait un candidat idéal pour corroborer notre hypothèse. De manière intéressante, l'étude des fonctions additionnelles de la GluProRS, et

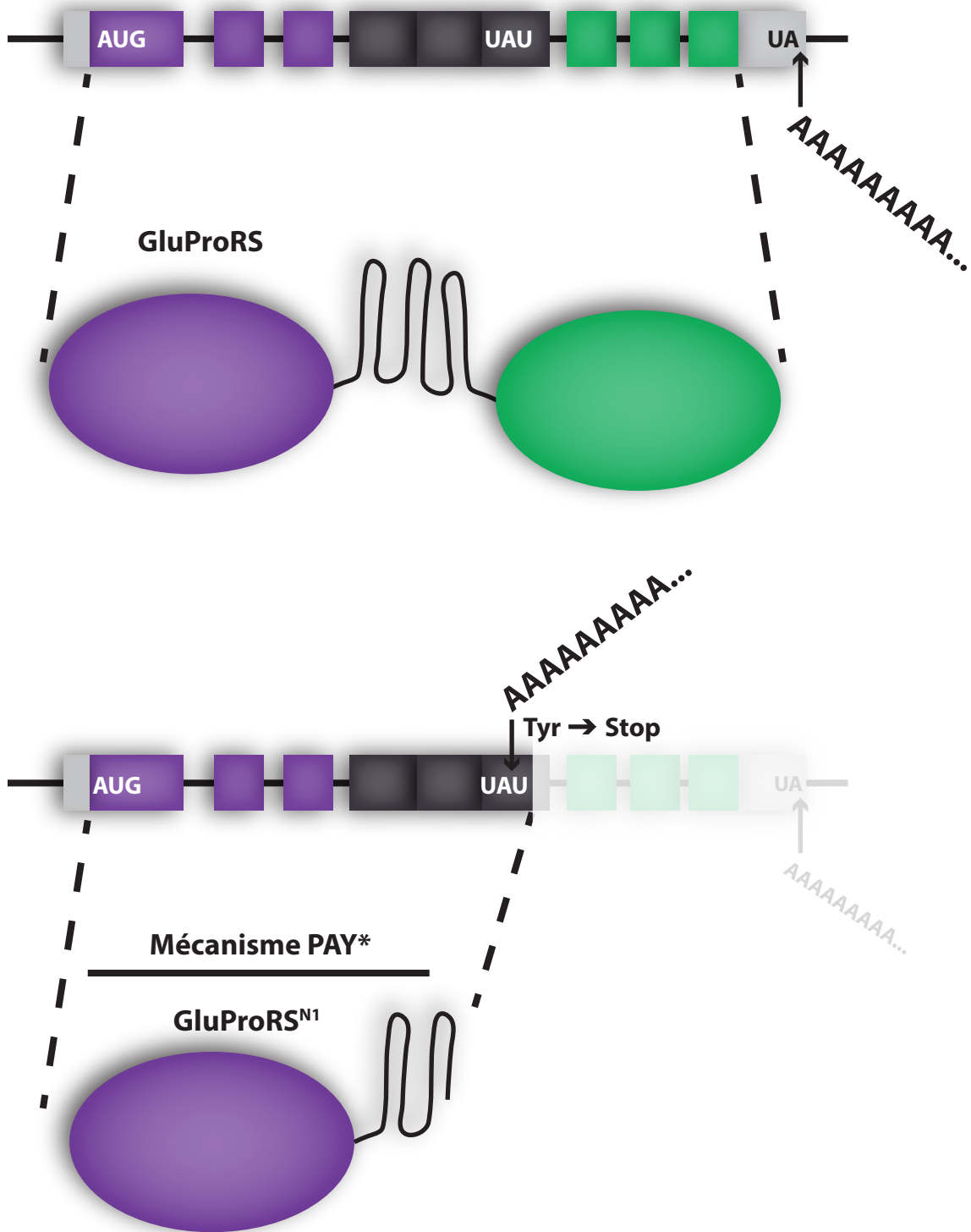


Figure 48 : Mécanisme d'expression de la GluProRS^{N1}

Expression de la GluProRS et de son variant traductionnel, la GluProRS^{N1}. L'expression de ce dernier est médiée par un mécanisme de polyadénylation alternative (PAY*), qui induit l'apparition d'un codon stop (UAA), en remplacement d'un codon Tyr (UAU). La GluProRS^{N1} comprend la région catalytique de la GluRS et quasiment deux répétitions de WHEP ; uniquement les douze derniers aa sont éliminés.

notamment son implication dans la régulation de la traduction *via* le complexe GAIT (voir introduction § VI.4.b), a dévoilé l'existence d'un variant transcriptionnel de la GluProRS. Celui-ci, nommé GluProRS^{N1}, est exprimé par un mécanisme de polyadénylation alternative (PAY* : *Polyadenylation-mediated Tyr-to-stop codon conversion*), permettant la conversion du codon Tyr (UAU) en codon stop (UAA) (**Figure 48**). La protéine résultante a une taille d'environ 95 KDa et comprend la région GluRS associée au premier et à la majorité du second domaine WHEP (uniquement les douze derniers aa de ce domaine sont éliminés). La GluProRS^{N1} est capable de s'associer avec une grande affinité aux régions en 3'UTR des ARNm, reconnues par le complexe GAIT (Yao and Fox, 2012; Yao et al., 2012). Elle inhibe alors l'activité de ce dernier et favorise une traduction basale des ARNm pro-inflammatoires. Cependant, l'absence du dernier domaine WHEP empêche son incorporation au sein du MSC. De plus, ce variant n'est pas restreint aux cellules myéloïdes et il est exprimé constitutivement dans divers types cellulaires ; notamment les cellules hépatiques (HepG2), les cellules embryonnaires de rein (HEK293T) et les cellules Hela. Il est également exprimé chez les primates et chez d'autres mammifères, tels que le chien, le chat ou la souris (Yao and Fox, 2012; Yao et al., 2012).

La GluProRS^{N1} est un variant très intéressant, car il ressemble fortement à la GluRS de levure. En effet, la GluProRS, dépourvue du domaine catalytique de la ProRS, arbore, comme son homologue de levure, un domaine *GST-like* au sein de son extrémité N-ter. Ces deux protéines étant très proches, il ne serait pas impossible que ce variant se relocalise également dans les mitochondries, d'autant qu'il n'est pas incorporé dans le MSC, pour assurer une fonction mitochondriale essentielle, et peut être la formation du Glu_{-m}ARNt^{Gln}.

2. Résultats préliminaires

a. Etude de la GluRS humaine

Au laboratoire, les premières études de la GluRS cytosolique humaine ont débuté avant la découverte de la GluProRS^{N1} et remontent au courant de l'année 2010. Matthieu Fréchin, alors doctorant au laboratoire, a étudié la capacité de la GluRS humaine, associée à trois domaines WHEP (GluRS^{3WHEP}) ou dépourvue de ceux-ci (GluRS^{0WHEP}), à remplacer son homologue de levure (constructions reçues du Dr. Mirande, Gif-sur-Yvette). Pour cela, il a utilisé une technique de chasse de plasmide (décrite dans les résultats § II.3). Pour ce faire, la

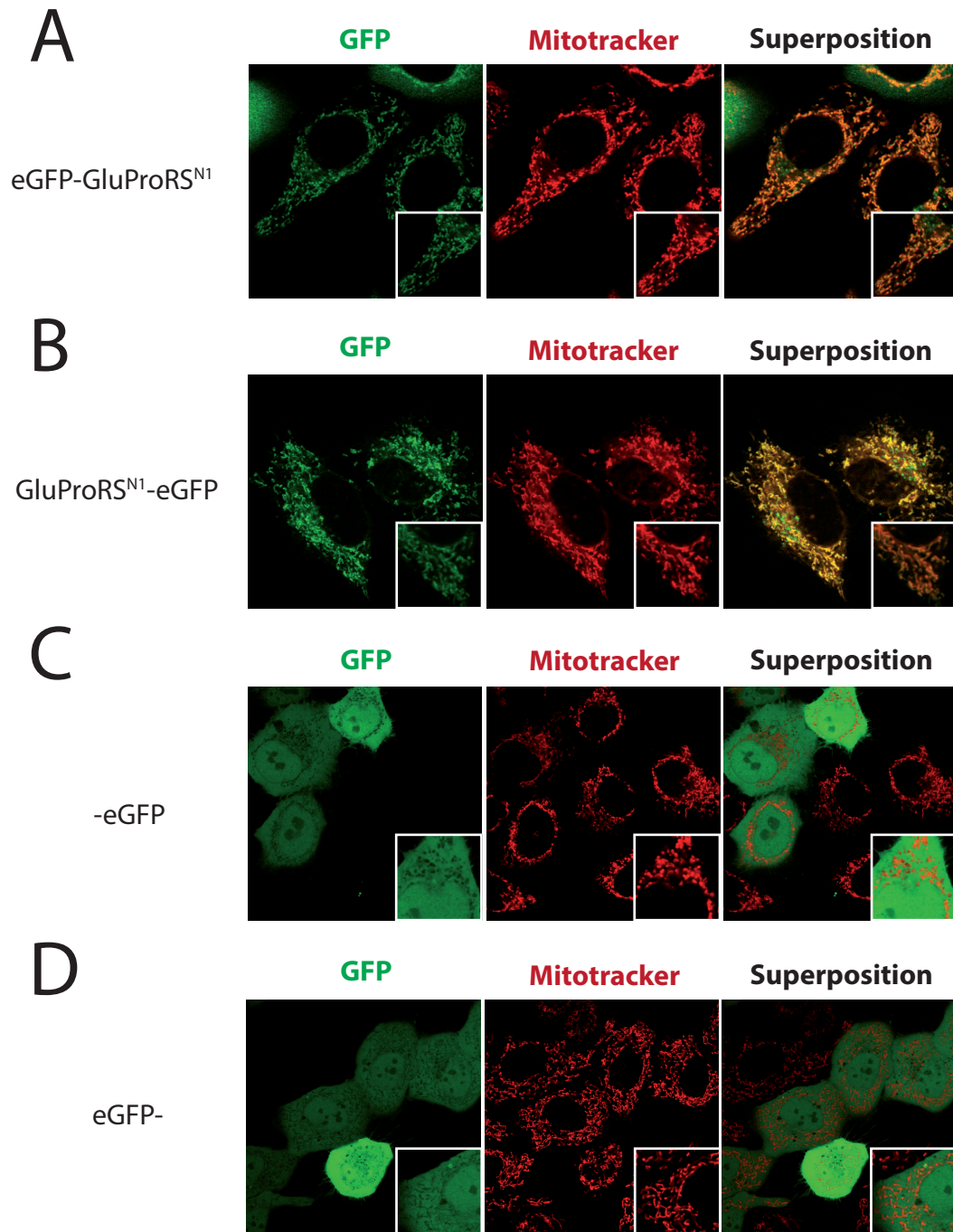


Figure 49 : Localisation subcellulaire de la GluProRS^{N1} étiquetée GFP dans des cellules HeLa

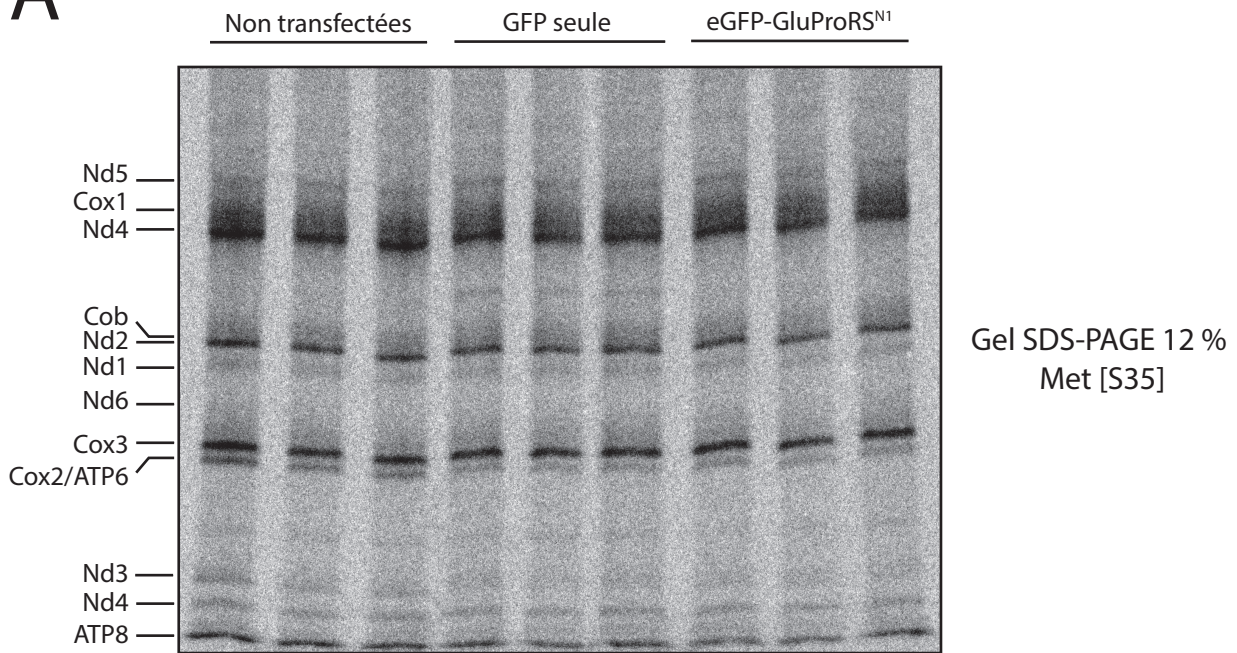
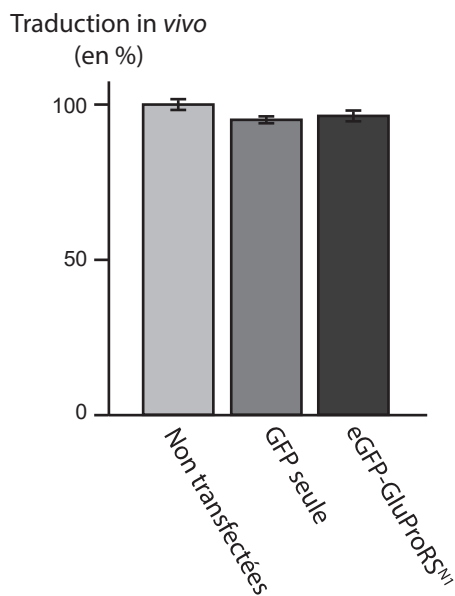
Les vecteurs d'expression de la GluProRS^{N1} étiquetée en N-ter (eGFP-GluProRS^{N1}) (A), C-ter (GluProRS^{N1}-eGFP) (B) ou les vecteurs vides (-eGFP ou eGFP-) (C-D) ont été transfectés dans des cellules HeLa (A-D). La fluorescence est observée deux jours après transfection par microscopie confocale et les mitochondries sont marquées au Mitotracker.

souche *gus1Δ Shuffle* a été transformée par les plasmides d'expression de GluRS^{3WHEP} et GluRS^{0WHEP}. Après sélection au 5FOA, la capacité de croissance fermentaire et respiratoire de ces souches a été évaluée. De manière intéressante, les deux souches après chasse sont capables de se développer en milieu fermentaire, confirmant la capacité de la GluRS humaine à compenser l'absence de son homologue de levure et, de ce fait, à aminoacyler l' ϵ -ARNt^{Glu} de levure. De plus, la souche exprimant la GluRS^{0WHEP} se développe en respiration, contrairement à celle contenant la GluRS^{3WHEP}. Ces résultats suggèrent que la GluRS^{0WHEP} est capable de se relocaliser dans la mitochondrie et de catalyser la première étape de la voie alternative de formation du Gln- ϵ -ARNt^{Gln}. De plus, il semblerait que les domaines WHEP, présents dans l'autre version de la GluRS interfèrent avec sa relocalisation mitochondriale et/ou sa capacité d'aminoacylation de l' ϵ -ARNt^{Gln}.

b. Etude de la GluProRS^{N1} dans des cellules Hela

Suite à la découverte d'un variant de la GluProRS, Guillaume Octobre (post doctorant), Ludovic Enkler (doctorant) et Aline Huber (technicienne), ont débuté le clonage de son gène dans des vecteurs d'expression humains dans l'optique d'étudier la localisation subcellulaire de la version étiquetée GFP dans des cellules humaines en culture. Pour ce faire, ils ont transfecté des cellules Hela avec les vecteurs d'expression de la GluProRS^{N1} étiquetée en N- ou C-ter avec la GFP ou les vecteurs vides correspondants, exprimant la GFP seule. De manière intéressante, que ce soit la version étiquetée en N- ou C-ter, le signal est clairement mitochondrial car il colocalise parfaitement avec le marquage des mitochondries au Mitotracker (**Figure 49A & B**). De plus, la GFP seule n'est restreinte qu'au cytosol et au noyau, confirmant bien la capacité de la GluProRS^{N1} à se relocaliser dans la mitochondrie (**Figure 49C & D**). La localisation mitochondriale de ce variant a également été confirmée par immuno-détection sur des extraits totaux de cet organelle purifié (résultats non montrés).

L'import forcé de la GluRS de levure dans la mitochondrie (souches nE et nE/nM) permet d'augmenter la traduction organellaire, alors que sa rétention cytosolique la diminue. Sachant que la GluProRS^{N1} se relocalise dans les mitochondries de cellules Hela, il était donc intéressant de vérifier si la traduction organellaire est modifiée lors d'une surexpression de cette protéine. Des tests de traductions *in vivo* ont alors été réalisés en triplicata sur des cellules non transfectées (NT), transfectées avec le plasmide vide (GFP seule) ou transfectées

A**B****Figure 50 : Capacité de stimulation de la traduction mitochondriale par la GluProRS^{N1}**

(A) Visualisation de l'activité traductionnelle mitochondriale *in vivo* sur des cellules Hela (i) non transfectées, (ii) exprimant la GFP seule ou (iii) exprimant la GluProRS^{N1} étiquetée GFP en N-ter (eGFP-GluProRS^{N1}). La traduction cytosolique est stoppée par traitement à l'émétine suivi d'addition de Met-[³⁵S] qui sera incorporée dans les protéines naissantes d'origine mitochondriale. Les extraits totaux correspondant sont séparés par SDS-PAGE et transférés sur membrane de PVDF avant exposition de celle-ci durant 24 h. Un triplicata biologique est réalisé pour chaque expérience. (B) Mesure de la radioactivité relative de chaque piste (relatif à la quantité de protéines traduites). Les barres d'erreurs correspondent à la moyenne des triplicatas.

avec le plasmide d'expression de la GluProRS^{N1} étiquetée GFP en N-ter (eGFP-GluProRS^{N1}). La traduction mitochondriale est quasiment identique dans les trois souches testées, suggérant qu'elle n'est pas stimulée lors de l'expression accrue de la GluProRS^{N1} (**Figure 50A & B**). Ces résultats corroborent les données de la littérature indiquant que la voie de formation du Gln-mARNt^{Gln} humain fait intervenir la GluRS mitochondriale ND. Cela suggère également que la fonction de la fraction de la GluProRS^{N1} importée dans la mitochondrie est indépendante de la synthèse protéique ou que son implication dans ce processus est dépendante de conditions physiologiques que nous n'avons pas encore identifiées.

Quel que soit son rôle dans ce compartiment, la GluProRS^{N1} est indéniablement capable de se relocaliser dans les mitochondries de cellules humaines dans des conditions normales de croissance. Cependant, il n'est pas à exclure que l'activité mitochondriale de cette protéine soit spécifique de conditions de croissance particulières. Il serait maintenant intéressant de déterminer la capacité de relocalisation mitochondriale de la GluProRS^{N1} dans différentes conditions physiologiques, et notamment dans des conditions de stress. N'étant pas des spécialistes en culture cellulaire, nous avons débuté une collaboration avec le laboratoire du Dr. Karin Musier-Forsyth afin de déterminer si l'import mitochondrial de la GluProRS^{N1} est constitutif dans un grand nombre de types cellulaires. Cependant, nous n'avons pas encore de retour de nos collaborateurs. En revanche, nous avons entrepris l'étude de la fonction mitochondriale de la GluProRS^{N1} chez la levure *S. cerevisiae*. C'est lors de cette période que je suis arrivé au laboratoire et que j'ai pris part au projet concernant la GluProRS^{N1}.

3. Caractérisation de la fonction de la GluProRS^{N1} chez la levure

L'utilisation de la levure pour caractériser des protéines venant d'autres organismes a permis la compréhension d'une multitude de processus biologiques essentiels. Pour élucider la fonction mitochondriale de la GluProRS^{N1}, différents plasmides d'expression (marqueur de sélection *LEU2*) ont été construits. Ceux-ci permettent l'expression de la version sauvage de cette protéine et de son variant dépourvu de WHEP, fusionnés ou non à la GFP (GluProRS^{N1}-GFP ; GluProRS^{N1} ; GluProRS^{N1}^{0WHEP}-GFP et GluProRS^{N1}^{0WHEP}).

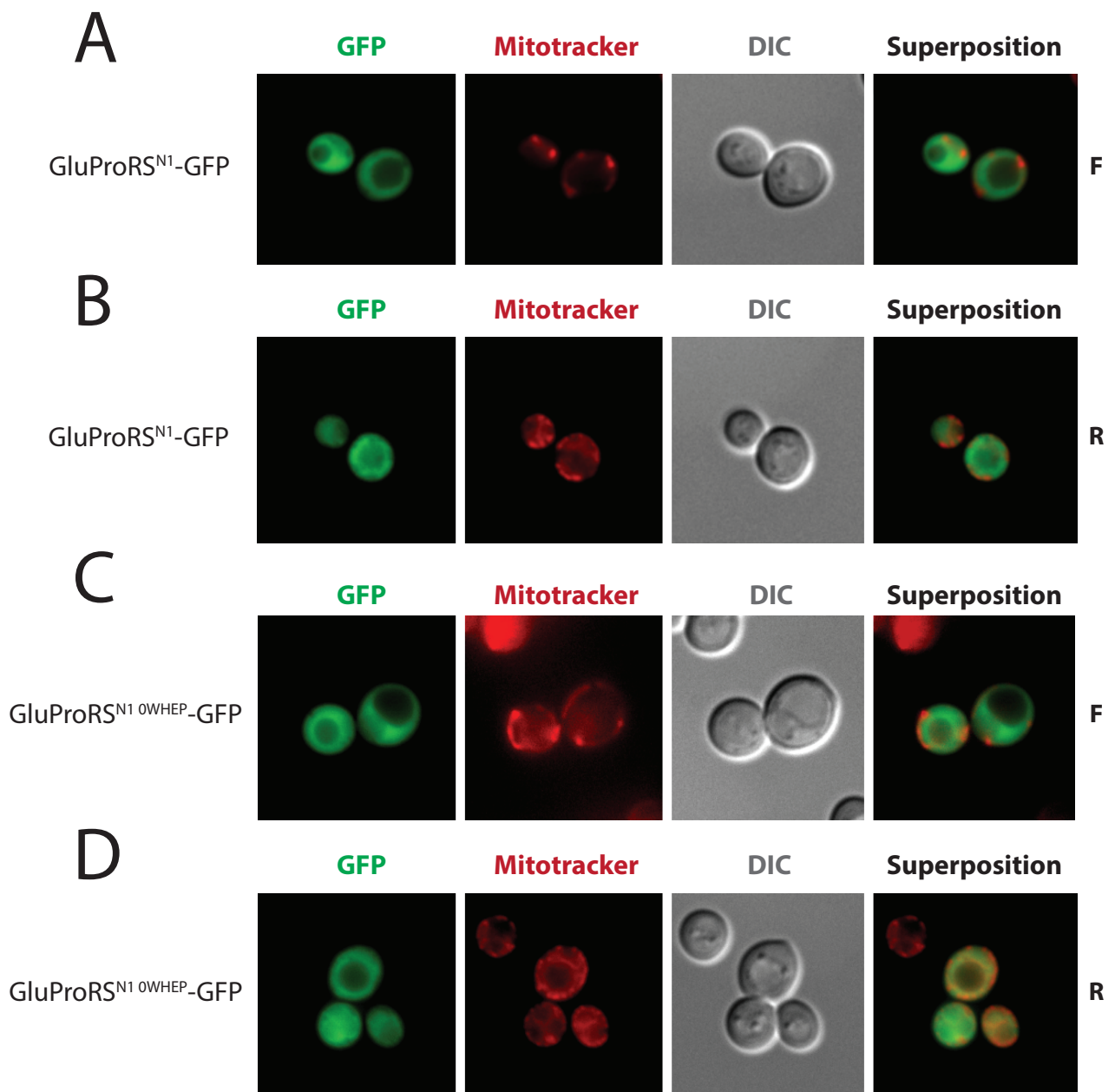


Figure 51 : Localisation subcellulaire de la GluProRS^{N1} chez la levure *S. cerevisiae*

Les plasmides d'expression de la GluProRS^{N1}-GFP et GluProRS^{N1}^{OWHEP}-GFP sont transformés dans une souche sauvage BY4742. La localisation des variants est alors visualisée par microscopie à épifluorescence et les mitochondries sont marquées au Mitotracker. Les cultures sont réalisées en fermentation (F) et respiration (R).

Nous avons, dans un premier temps, étudié la localisation subcellulaire des versions étiquetées GFP. Pour cela, les souches ont été cultivées en fermentation ou en respiration et les protéines de fusion ont été visualisées par microscopie à épifluorescence. Que ce soit en fermentation ou en respiration, la GluProRS^{N1}-GFP est retrouvée dans le cytosol et dans les mitochondries ; il y a colocalisation entre la fluorescence verte et le marquage des mitochondries au Mitotracker (**Figure 51A & B**). Cependant, la localisation mitochondriale n'est que partielle et il est possible que l'import de ce variant ne soit pas optimal. De plus, la même répartition de la fluorescence est observée dans les souches exprimant la GluProRS^{N1}_{OWHEP}-GFP (**Figure 51C & D**). Ces résultats valident (i) l'expression hétérologue de la GluProRS^{N1} humaine, ainsi que (ii) la localisation cytosolique et mitochondriale de cette enzyme. De plus, les données suggèrent que la présence des domaines WHEP n'influent pas sur sa localisation mitochondriale.

Suite à ces résultats, nous avons voulu déterminer si les différentes protéines de fusion sont capables de remplacer la GluRS de levure. Pour cela, nous avons utilisé une technique de chasse de plasmide. Tout d'abord, une souche *gus1Δ Shuffle* a été transformée avec les quatre plasmides d'expression de la GluProRS^{N1} et GluProRS^{N1}_{OWHEP}, étiquetées GFP ou non. Les souches exprimant ces protéines ont ensuite été étalées sur le milieu 5FOA, permettant de forcer la chasse du plasmide d'expression du gène *GUS1*, codant pour la GluRS (marqueur *URA3*). Suite à l'apparition de colonies, nous avons réalisé des tests en gouttes sur divers milieux de sélection pour étudier la capacité fermentaire et respiratoire des souches *gus1* (dépourvues de GluRS) exprimant la GluProRS^{N1}. Toutes les souches sont capables de se développer sur un milieu minimum (SC) (**Figure 52**). De manière intéressante, contrairement à la souche *gus1Δ Shuffle*, toutes les souches exprimant les versions étiquetées ou non de la GluProRS^{N1} poussent sur milieu de sélection 5FOA. Ces résultats confirment la capacité de cette enzyme et de ses variants sans domaines WHEP à compenser l'absence de la GluRS endogène et à reconnaître et aminoacyler l'^cARNt^{Glu} de levure (**Figure 52**). Pour confirmer la perte du plasmide permettant l'expression de la GluRS de levure, les souches ont également été déposées sur milieu SC-Ura. La souche *gus1Δ Shuffle* est capable de croître sur ce milieu, tandis que les souches après chasse ne le sont pas (**Figure 52**). Ces souches ne contiennent donc plus le plasmide d'expression de *GUS1*. Les souches dépourvues de ce plasmide sont également capables de croître sur milieu SC-Leu, confirmant bien la présence des plasmides

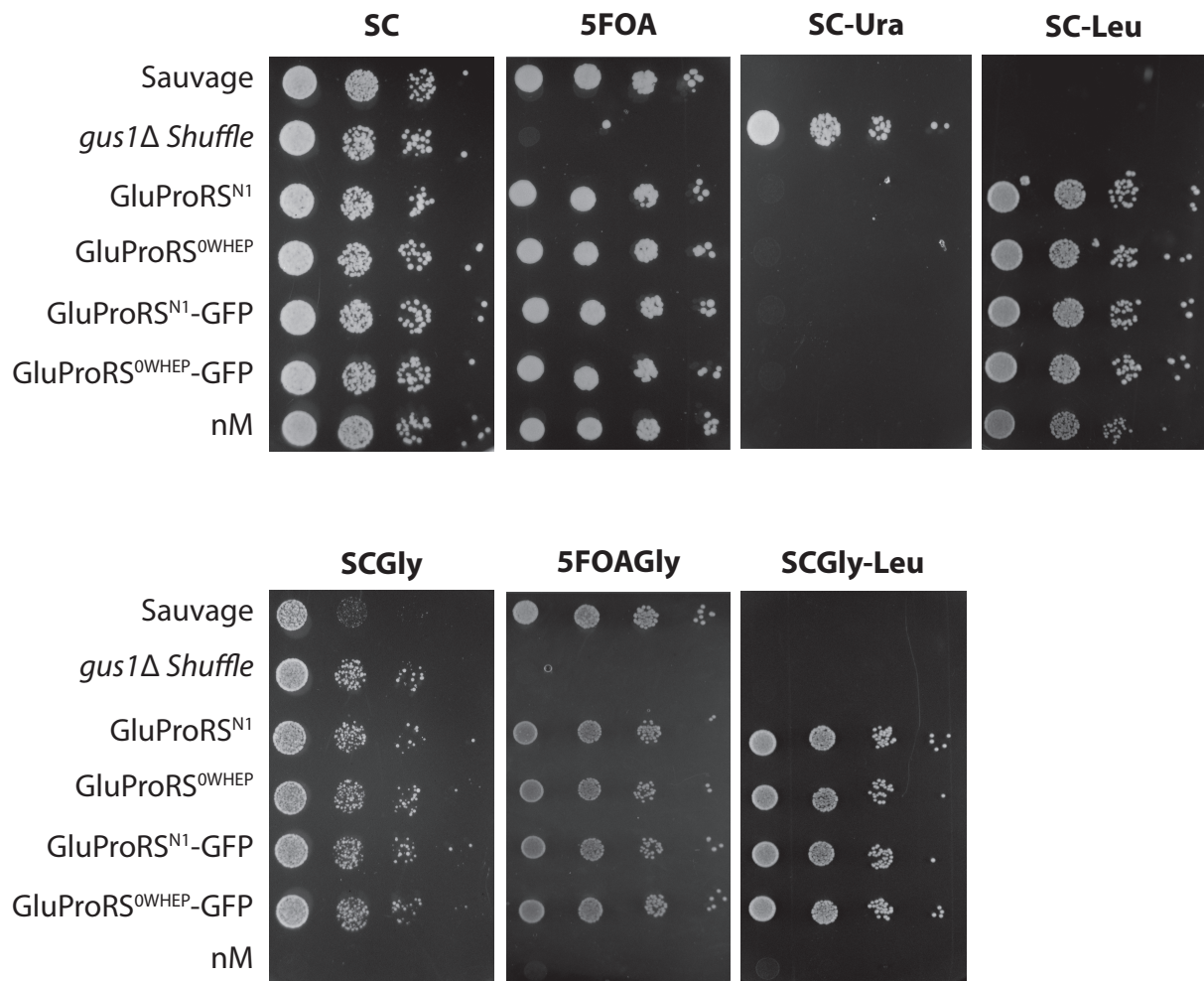


Figure 52 : Capacité de croissance des souches exprimant la GluProRS^{N1} après chasse au 5FOA

Après sélection au 5FOA, les souches exprimant la GluProRS^{N1} étiquetées GFP ou non sont déposées, selon des dilutions en cascade de raison 10, sur différents milieux de culture afin de tester leur capacité de croissance. En guise de témoin, les souches *gus1Δ Shuffle*, et nM (incapable de croître en respiration) sont également déposées. Après deux à quatre jours de croissance, les boîtes sont prises en photo.

d'expression de la GluProRS^{N1} et de ses variants, et son absence dans la souche *gus1Δ Shuffle* (qui ne pousse pas) (**Figure 52**). Pour étudier la capacité respiratoire de ces souches, elles ont été déposées sur des milieux respiratoires (SCGly, 5FOAGly et SCGly-Leu). Comme attendu, la souche nM (contrôle négatif) ne pousse pas, alors qu'une croissance est visualisable pour toutes les autres souches (**Figure 52**). Ces résultats suggèrent que l'enzyme humaine est importée dans les mitochondries de levure et participe, en coopération avec l'AdT GatFAB, à la formation du Gln-_mARnt^{Gln}. De plus, la présence des domaines WHEP ne perturbe, ni l'import mitochondrial, ni l'activité de cette enzyme dans ce compartiment. En parallèle, Ludovic Enkler a également validé l'expression des versions étiquetées de la GluProRS^{N1} et de la GluProRS^{N1} ^{OWHEP} par immuno-détection sur des extraits totaux (résultats non montrés), confirmant bien l'expression et la présence des variants après chasse du plasmide. Finalement, la croissance respiratoire est assurée, malgré la relocalisation mitochondriale partielle des versions étiquetées GFP, observée par microscopie.

4. Validation de la fonction mitochondriale de la GluProRS^{N1} chez la levure

Tous nos résultats vont donc dans le même sens et confirment que la GluProRS^{N1} est capable de se relocaliser dans les mitochondries (humaines et de levure). Cependant, chez l'Homme, elle n'augmente pas la capacité traductionnelle de ce compartiment dans des conditions standard. L'expression hétérologue chez la levure a permis de suggérer que cette enzyme et de son variant, dépourvu de domaines WHEP, à remplacer la GluRS de levure en fermentation *via* l'aminoacylation de l'_cARnt^{Glu}, mais également en respiration, suite à son import mitochondrial, *via* l'aminoacylation de l'_mARnt^{Gln}. La capacité de la GluProRS^{N1} à assurer cette fonction mitochondriale chez l'Homme est cependant plus controversée. En effet, cette voie de formation fait intervenir la GluRS mitochondriale ND, capable d'aminoacyler le _mARnt^{Gln}. En outre, pourquoi la cellule synthétiserait-elle un variant de la GluProRS, capable d'entrer dans les mitochondries pour aminoacyler l'_mARnt^{Gln}, si elle possède déjà une enzyme capable d'assurer cette fonction ? Nous avons cherché à valider les résultats précédents par une approche de génétique de levure alternative. Pour cela, j'ai d'abord construit un plasmide intégratif permettant l'expression de la GluProRS^{N1} étiquetée HA (pAG306-GPD-GluProRS^{N1}-HA). Nous avons alors décidé d'étudier la ségrégation des spores d'une souche *gus1Δ* diploïde exprimant la GluProRS^{N1}. Pour cela, avec Laurence Huck, technicienne au laboratoire, nous avons intégré cette construction au niveau du locus *trp1*.

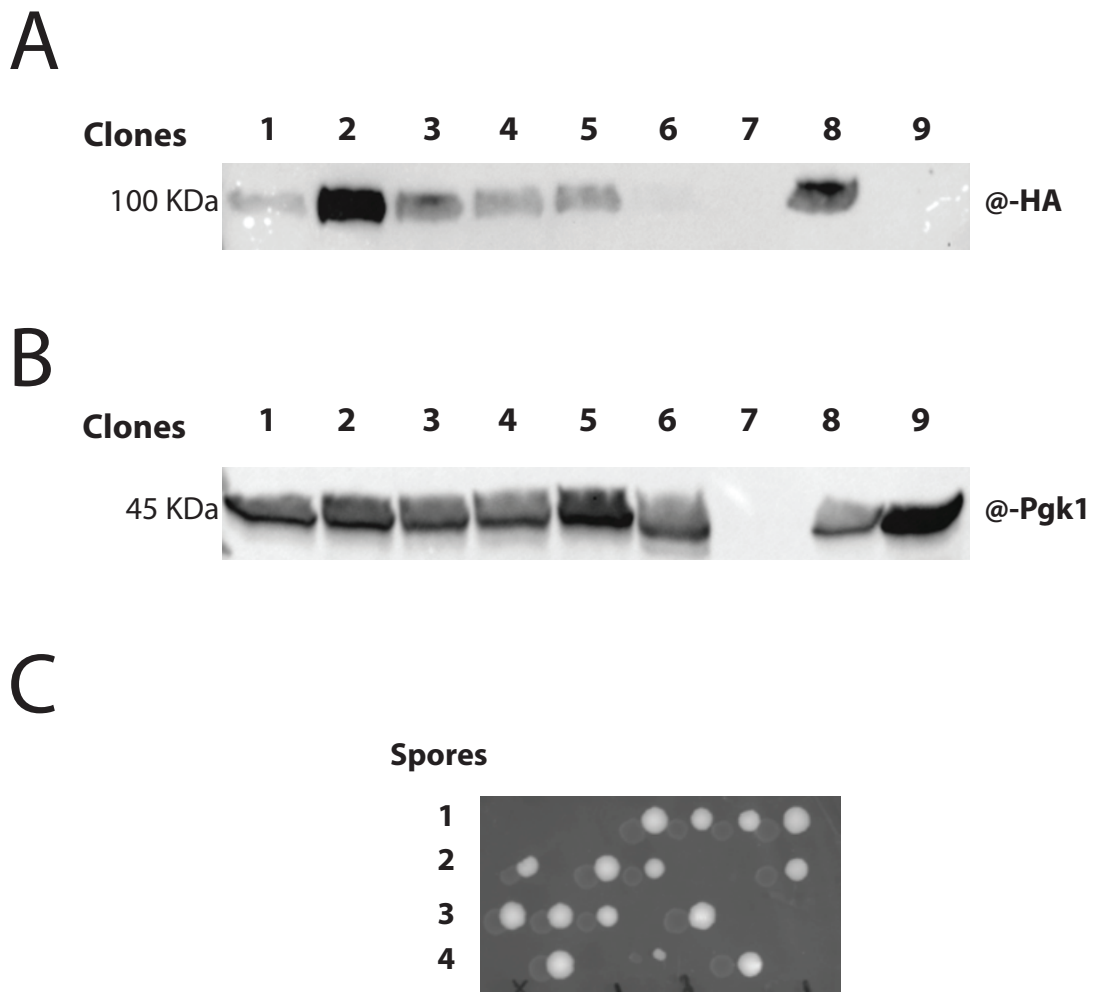


Figure 53 : Expression de la GluProRS^{N1}-HA et test de complémentation

La souche *gus1Δ* diploïde a été transformée avec le vecteur intégratif pAG306-GPD-GluProRS^{N1}-HA. Ensuite, neuf colonies isolées ont été récupérées et la présence de la protéine d'intérêt a été validée par immuno-détections, avec un anticorps @-HA, sur des extraits totaux (A). Le contrôle de charge (@-Pgk1) confirme les résultats d'expression (B). Le clone 2 a alors été mis à sporuler et les spores isolées sont déposées sur milieu riche (C). La ségrégation 2/2 indique une absence de complémentation.

Pour rappel, cette souche, décrite dans les résultats § II.3, contient une copie sauvage (*GUS1*) et une copie mutante (*gus1*) du gène codant la GluRS. Ainsi, après dissection, deux spores posséderont *GUS1* et exprimeront donc cette enzyme, tandis que les deux autres contiendront l'allèle *gus1* (muté) et ne se développeront pas. En parallèle, deux spores sur les quatre exprimeront la GluProRS^{N1}. Par conséquent, quatre types de spores seront potentiellement retrouvés : (i) les spores sauvages (*GUS1*) n'exprimant pas la GluProRS^{N1}, (ii) les spores sauvages (*GUS1*) exprimant la GluProRS^{N1}, (iii) les spores "mutées" (*gus1*) n'exprimant pas la GluProRS^{N1} et (iv) les spores "mutées" (*gus1*) exprimant la GluProRS^{N1}. La présence de ce dernier type de spores nous permettra de valider la compensation de l'activité de la GluRS de levure par l'enzyme humaine. Nous avons, tout d'abord, récolté neuf colonies isolées pour déterminer si la version étiquetée de cette protéine était exprimée. Les immuno-détections avec l'anticorps @-HA, sur des extraits totaux, montrent bien qu'il y a expression de la protéine d'intérêt dans la majorité des clones avec une intensité plus forte pour les clones 2 et 8 (**Figure 53A**). De plus, cette intensité n'est pas due à un dépôt de protéines plus important comme en témoigne le contrôle de charge (@-PGK) (**Figure 53B**). Les clones 2 et 8 ont alors été sélectionnés et mis à sporuler, car l'expression de la GluProRS^{N1} est bien supérieure à celle des autres clones. Après dissection des tétrades, la ségrégation est étudiée par étalement sur un milieu riche (YPD). A notre grand étonnement, la dissection de ces deux clones indique le même résultat, à savoir une ségrégation 2/2 (**Figure 53C**). De plus, les spores qui se développent contiennent uniquement l'allèle sauvage de *GUS1* et, de ce fait, aucune spore *gus1* n'a pu être observée. Pour confirmer ces résultats, j'ai effectué, en collaboration avec le Dr. De Craene, la dissection de plusieurs clones indépendants. En parallèle, j'ai transformé la souche *gus1Δ* diploïde avec un vecteur centromérique (pAG416-GPD-GluProRS^{N1}-HA) afin d'étudier la ségrégation des spores. En effet, l'insertion de la cassette contenant le gène codant la GluProRS^{N1} au niveau du génome, restreint la présence de la protéine à deux spores. Pour augmenter la probabilité d'observer des spores *gus1* mutées exprimant la GluProRS^{N1}, l'expression sur un vecteur était nécessaire. Cependant, malgré une expression de cette enzyme dans les quatre spores, uniquement les spores exprimant la GluRS de levure se développent (*GUS1* sauvage). Ces résultats suggèrent que la délétion du gène codant la GluRS de levure ne peut pas être compensée par l'expression de la GluProRS^{N1} humaine. Contrairement à ce qui avait été montré avec les sélections au 5FOA, ces résultats soulignent l'incapacité de l'enzyme humaine à remplacer son homologue de levure.

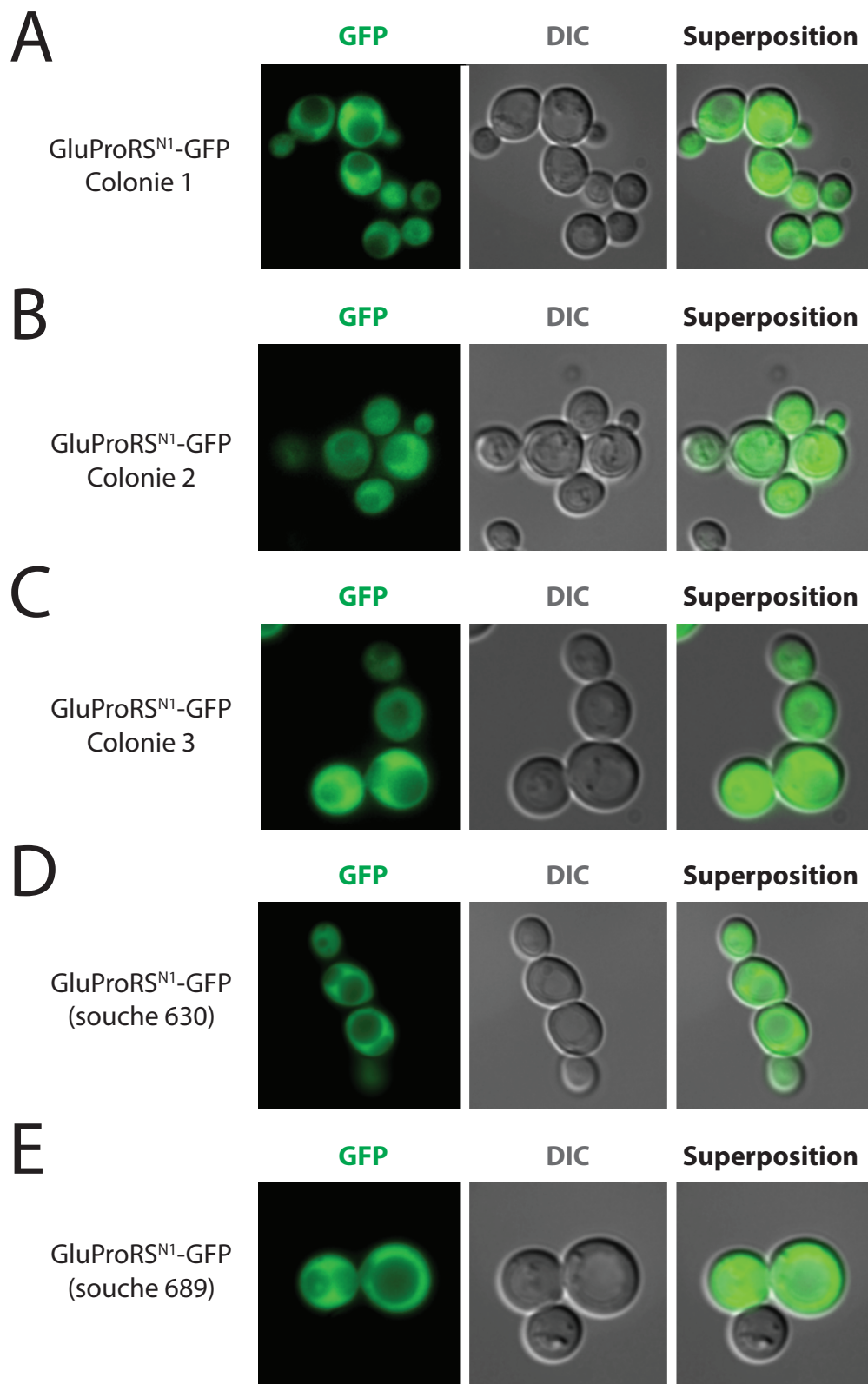


Figure 54 : Expression de la GluProRS^{N1}-GFP chez la levure

Trois colonies isolées provenant d'un souche sauvage BY4742, transformée avec le pAG304-GPD-GluProRS^{N1}-GFP, sont visualisées par microscopie à épifluorescence (A-C). Les souches de la collection avant (D) et après chasse (E) au 5FOA, exprimant la version plasmidique (p415-ADH-GluProRS^{N1}-GFP), sont également observées.

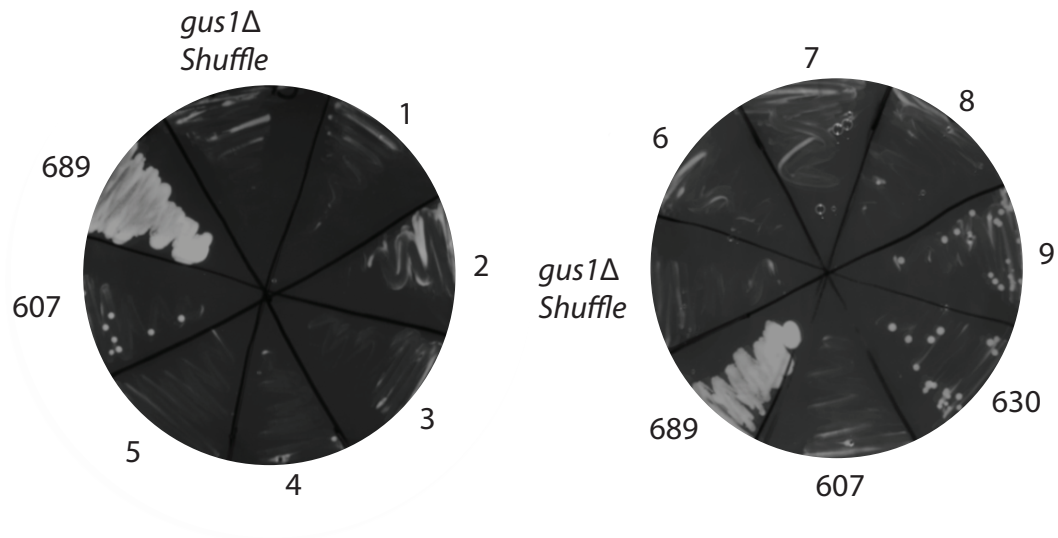


Figure 55 : Tests de complémentation de la GluProRS^{N1} par sélection au 5FOA

La capacité de complémentation de la GluProRS^{N1} a été testée par sélection au 5FOA. Les souches 1 à 3 correspondent aux des souches *gus1Δ Shuffle* transformées avec le vecteur intégratif pAG304-GPD-GluProRS^{N1}-GFP, les souches 4 à 6 le sont avec le vecteur intégratif pAG304-GPD-GluProRS^{N1}-HA et finalement, les souches 7 à 9 sont transformées avec le vecteur centromérique pAG414-GPD-GluProRS^{N1}-HA. En guise de témoin, la souche *gus1Δ Shuffle* transformée avec le plasmide vide, les souches 607 et 630, correspondant aux souches *gus1Δ Shuffle* exprimant la GluProRS^{N1} et GluProRS^{N1}-GFP avant chasse et la souche 689, correspondant *gus1Δ Shuffle* exprimant la GluProRS^{N1}-GFP après chasse sont également déposées. Les boîtes sont prises en photo trois à quatre jours après incubation à 30 °C.

Pour comprendre les différences de phénotypes observées entre les techniques de dissections et de chasse de plasmides, j'ai décidé de transformer la souche *gus1Δ Shuffle*, soit avec un plasmide d'expression de la GluProRS^{N1} étiquetée GFP (pAG304-GPD-GluProRS^{N1}-GFP), soit avec des plasmides permettant son expression étiquetée HA (pAG304-GPD-GluProRS^{N1}-HA et pAG414-GPD-GluProRS^{N1}-HA). Par microscopie à épifluorescence, j'ai alors vérifié l'expression et la localisation de la protéine d'intérêt à partir de trois colonies isolées. Toutes les souches présentent une fluorescence cytosolique intense (**Figure 54A, B & C**), validant l'expression de la GluProRS^{N1}-GFP. J'ai également observé l'expression de cette protéine dans deux souches de la collection (construites par Ludovic Enkler) : la souche *gus1Δ Shuffle* exprimant la GluProRS^{N1}-GFP avant chasse au 5FOA (souche 630) et (iii) cette même souche après chasse au 5FOA (souche 689) (**Figure 54D & E**). J'ai alors étalé les souches exprimant les différentes versions de la GluProRS^{N1} (étiquetées GFP ou HA) sur milieu de sélection 5FOA. En guise de témoin, j'ai déposé les souches 630, 689 et la souche *gus1Δ Shuffle* exprimant la GluProRS^{N1} (sans étiquette) avant chasse au 5FOA (souche 607). La souche 689 se développe normalement sur milieu 5FOA. De plus, des colonies isolées sont présentes pour les souches 607 et 630, validant leur capacité à perdre le plasmide *GUS1* (**Figure 55**). Sur ces mêmes boîtes, j'ai déposé neuf clones exprimant la protéine fusionnée à la GFP ou à l'étiquette HA. Ainsi, les souches 1, 2 et 3 proviennent d'une souche *gus1Δ Shuffle* transformée avec le vecteur intégratif pAG304-GPD-GluProRS^{N1}-GFP et les souches 4, 5 et 6 le sont avec le vecteur intégratif pAG304-GPD-GluProRS^{N1}-HA. Finalement, les souches 7, 8 et 9 sont transformées avec le vecteur centromérique pAG414-GPD-GluProRS^{N1}-HA. La plupart des clones ne permettent pas la chasse du plasmide *URA3* (**Figure 55**). En effet, des colonies isolées apparaissent uniquement pour les clones 8 et 9, provenant d'une souche *gus1Δ Shuffle* transformée avec le pAG414-GPD-GluProRS^{N1}-HA. Ceci peut être dû à l'étalement d'une trop faible quantité de cellules, mais également à l'incapacité de ces souches à se passer de la GluRS endogène. De plus, la quantité de GluProRS^{N1} est probablement plus faible lors de son expression à partir du génome par rapport à celle à partir du plasmide ; ceci pourrait expliquer l'absence de complémentation. J'ai néanmoins utilisé les clones obtenus pour les souches 8 et 9 afin de vérifier la perte du plasmide *URA3*. Pour cela, j'ai à nouveau réalisé des tests en gouttes en déposant : (i) la souche *gus1Δ Shuffle*, (ii) les clones 8 et 9 avant chasse et (iii) des colonies apparues après chasse (8A, 9A et 9B). J'ai alors cultivé ces souches quatre jours sur différents milieux et les résultats sont présentés dans la **Figure 56**. Comme attendu, la souche

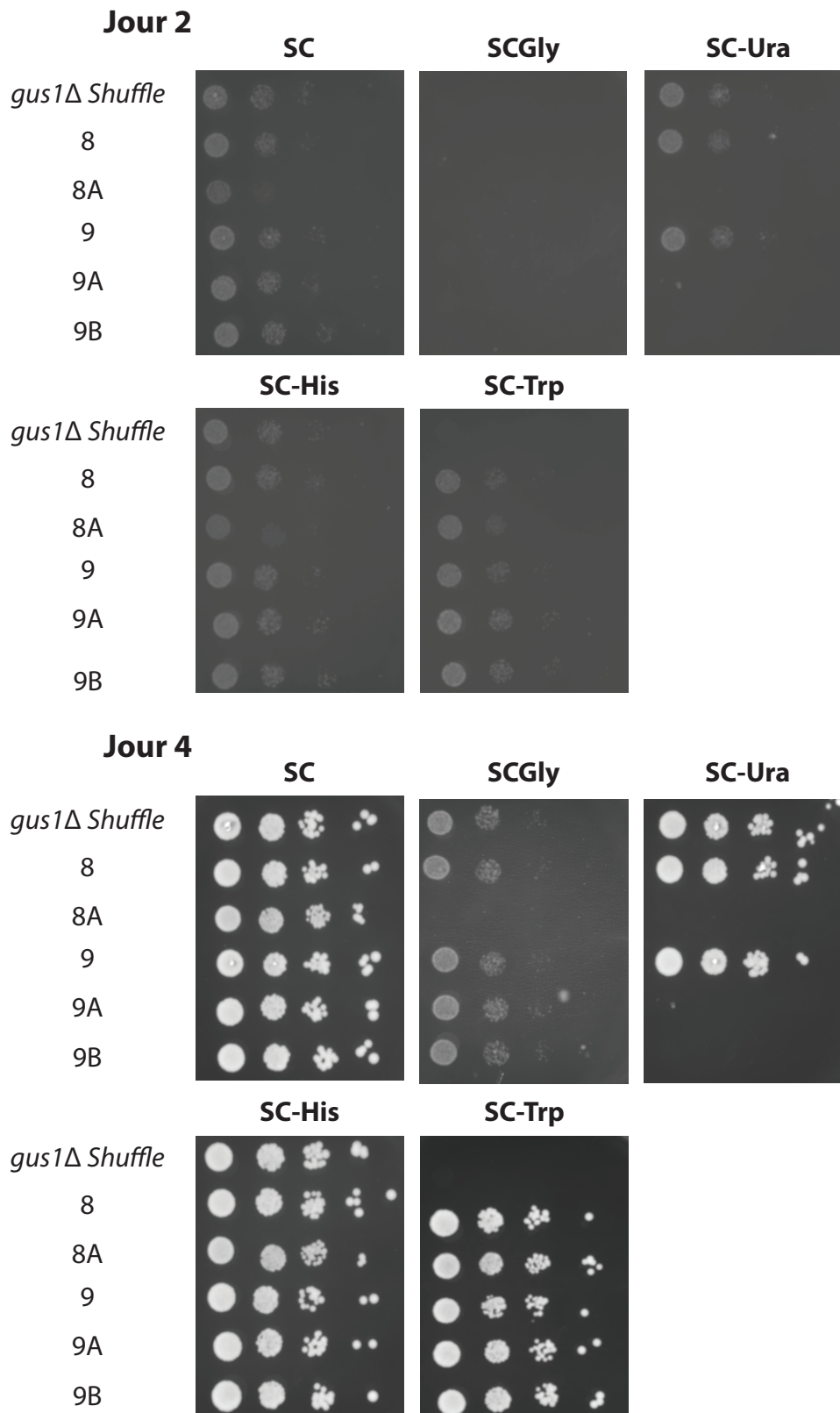
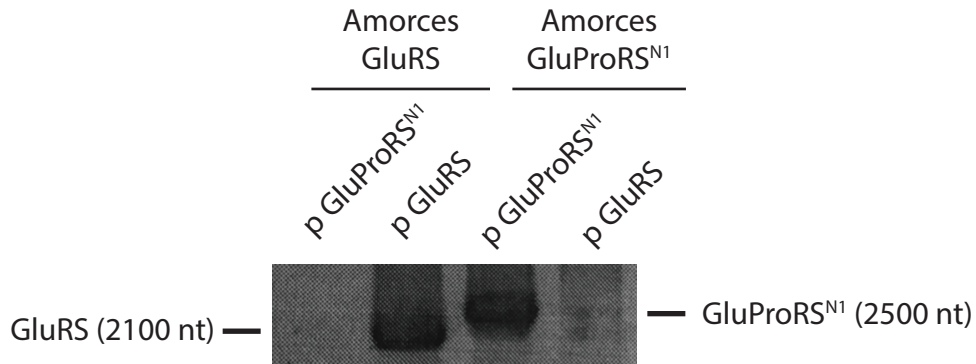


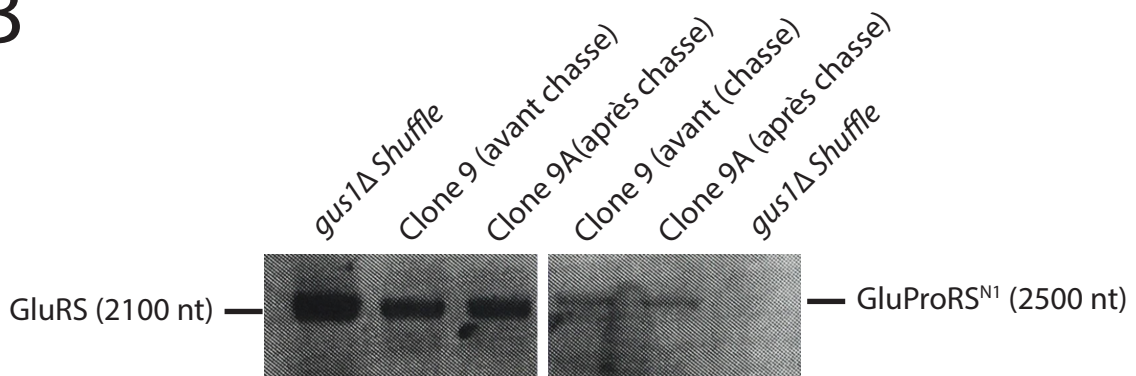
Figure 56 : Capacité proliférative des souches avant et après sélection FOA

Les souches exprimant la GluProRS^{N1} étiquetées HA sont déposées, selon des dilutions en cascade de raison 10, sur différents milieux de croissance afin de tester leur capacité proliférative. Ainsi, les souches 8 et 9 correspondent aux clones avant chasse au 5FOA et les souches 8A, 9A et 9B correspondent à des clones isolés de ces mêmes souches après chasse. La souche *gus1Δ Shuffle* est déposée en tant que témoin. Les photos de boîtes sont prises à deux et quatre jours.

A



B



C

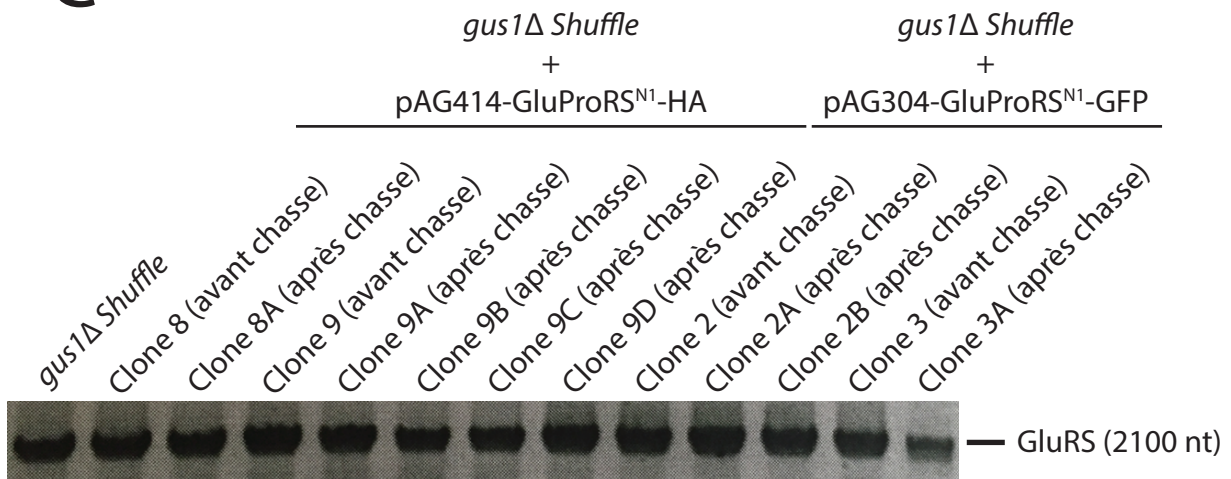


Figure 57 : Vérification de l'absence de *GUS1* par PCR

L'élimination du vecteur d'expression de *GUS1* est vérifiée par PCR avec des amorces spécifiques de *GUS1* et du gène codant la GluProRS^{N1}. (A) Les amorces sont tout d'abord validées par PCR sur des plasmides d'expression contenant *GUS1* et le gène codant la GluProRS^{N1}. (B) Les souches *gus1Δ Shuffle*, *gus1Δ Shuffle* transformée avec le vecteur d'expression de la GluProRS^{N1} avant (9) et après (9A) chasse sont alors testées avec les deux paires d'amorces. (C) Finalement, la présence de *GUS1* est testée chez tous les clones utilisés pour la chasse (2, 8 et 9), ainsi que les clones obtenus après chasse (2A, 2B, 3A, 8A, 9A, 9B, 9C et 9D).

témoin *gus1Δ Shuffle* se développe en fermentation (SC), en respiration (SCGly) et sur milieu dépourvu d'uracile (SC-Ura), confirmant bien la présence du plasmide d'expression de la GluRS de levure. En parallèle, les souches 8 et 9, correspondant aux clones avant chasse, se développent sur ces mêmes milieux et sur milieu dépourvu de tryptophane (SC-Trp), car elles possèdent le plasmide d'expression de la GluProRS^{N1}-HA. Les colonies apparues après chasse du plasmide (9A et 9B), sont capables de se développer sur milieu SC, SC-Trp et SCGly, mais ne sont plus capables de croître sur SC-Ura, confirmant la perte du plasmide *URA3* (**Figure 56**). De manière surprenante, la souche 8A se développe sur les mêmes milieux que les autres souches après chasse, mais pas en respiration (SCGly). Le gène *GUS1* est remplacé par la cassette *HIS3*. Toutes les souches testées poussent sur SC-His ; nous sommes en présence des souches délétées. Ces résultats confirment la capacité des souches ayant perdu le plasmide *GUS1* à se développer en fermentation et en respiration et sont donc concordants avec les résultats de chasse préliminaires. La capacité proliférative de ces différentes souches valide finalement la complémentarité de la perte de la *GUS1* plasmidique par la GluProRS^{N1}.

Pour confirmer ces résultats, j'ai décidé de réaliser des PCR sur colonies avec des amorces spécifiques du gène *GUS1* ou de celui codant la GluProRS^{N1}. J'ai alors utilisé le kit Phire *plant direct PCR master Mix* pour tenter d'amplifier ces deux gènes à partir de l'ADN cellulaire (génomique et plasmidique) des souches avant ou après chasse sur milieu 5FOA. J'ai, dans un premier temps, vérifié la spécificité des amorces en réalisant des PCR sur deux plasmides exprimant ces gènes. Les amorces de *GUS1* sont bien spécifiques et ne permettent pas l'amplification de la GluProRS^{N1}, et inversement, les amorces du gène codant la GluProRS^{N1} ne permettent pas l'amplification de *GUS1* (**Figure 57A**). Afin de vérifier l'absence du gène *GUS1* dans la souche après chasse au 5FOA, j'ai réalisé des PCR sur les souches *gus1Δ Shuffle*, 9 et 9A. J'ai, par la même occasion, vérifié la présence du gène codant la GluProRS^{N1}. Les résultats indiquent la présence de *GUS1* dans les trois souches testées. De plus, le gène codant la GluProRS^{N1} est présent uniquement dans les souches 9 et 9A, confirmant la présence du plasmide exprimant l'enzyme humaine (**Figure 57B**). Malgré la chasse du plasmide lors de la sélection au 5FOA, les résultats suggèrent la présence du gène d'expression de la GluRS. Ces résultats sont déroutants, car les souches ne se développent plus sur milieu SC-Ura. La présence du gène *GUS1* après chasse du plasmide suggère que la protéine est potentiellement

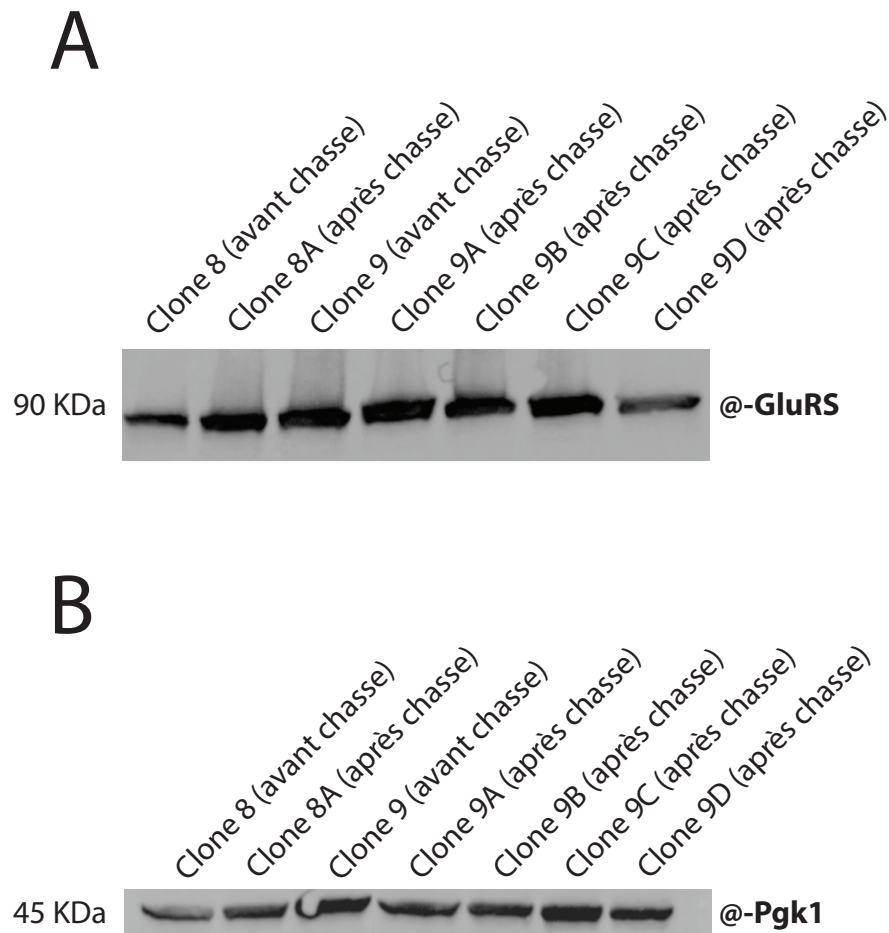


Figure 58 : Vérification de l'expression de la GluRS de levure

L'expression de la GluRS de levure est vérifiée sur des extraits totaux de clones avant (8 et 9) et après (8A, 9A, 9B, 9C et 9D) chasse au 5FOA par immuno-détection avec un anticorps @-GluRS (1/3000) (A). Le témoin de charge est réalisé avec un @-Pgk1 (1/4000) (B).

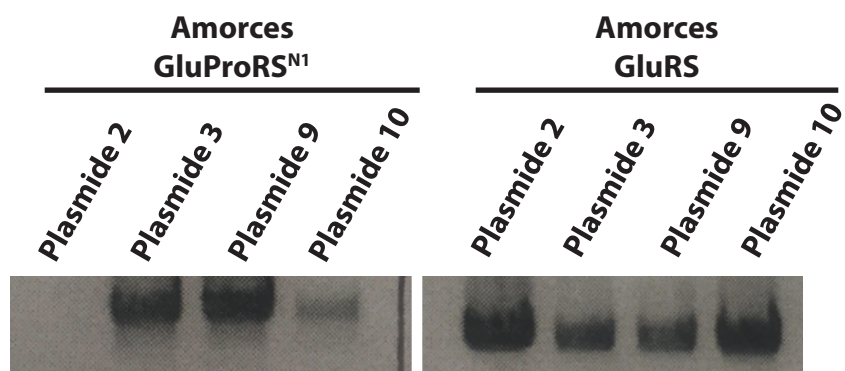


Figure 59 : Vérification de la recombinaison de *GUS1*

Après extraction de l'ADN plasmidique à partir de la souche 9A, l'ADN est utilisé pour réaliser une transformation de bactéries. Quatre plasmides issus de colonies isolées sont utilisés pour vérifier la présence du gène *GUS1* par PCR. La présence du gène codant la GluProRS^{N1} est également vérifiée par PCR.

exprimée dans cette souche et que le phénotype observé ne dépend donc pas de la présence de la GluProRS^{N1}.

En parallèle, j'ai à nouveau réalisé une sélection au 5FOA avec les souches exprimant la GluProRS^{N1}-HA ou la GluProRS^{N1}-GFP au génome. J'ai obtenu quelques colonies isolées pour cette seconde souche. J'ai décidé de tester, par PCR avec les amorces spécifiques de *GUS1*, la présence de ce gène dans tous les clones avant et après chasse du plasmide. Comme c'est le cas pour les premiers clones testés, une bande apparaît à la taille de *GUS1*, confirmant la présence de ce gène dans toutes les souches étudiées (**Figure 57C**). Ceci est très surprenant, car ce gène ne devrait plus être présent après chasse du plasmide et aucune amplification ne devrait être observée. Si *GUS1* est présent, il permet probablement l'expression de la GluRS de levure et le phénotype observé n'est pas spécifique de la complémentation par la GluProRS^{N1}, mais dépend de l'activité de la GluRS endogène. J'ai néanmoins voulu confirmer l'expression de l'aaRS de levure dans les souches avant et après chasse du plasmide. Pour cela, j'ai préparé des extraits totaux et effectué une immuno-détection avec l'anticorps @-GluRS. Conformément aux résultats obtenus par PCR, la GluRS est bien exprimée dans les souches *gus1Δ Shuffle*, 8 et 9, mais également dans les souches après chasse du plasmide (8A, 9A, 9B, 9C et 9D) (**Figure 58A & B**). Tous ces résultats indiquent que les souches après chasse du plasmide *URA3* possèdent encore le gène *GUS1* et que la GluRS de levure est exprimée. Ces données confirment également que les phénotypes observés ne sont pas dépendants de l'activité de la GluProRS^{N1}, mais sont probablement dus à la présence de la GluRS de levure.

La présence du gène *GUS1* suggère une potentielle recombinaison de ce gène directement au niveau du génome de levure ou au sein du plasmide d'expression de la GluProRS^{N1}. J'ai donc cherché à valider ou réfuter l'une de ces hypothèses. Pour cela, j'ai réalisé une préparation de plasmide directement à partir de la souche 9A, puis j'ai transformé une souche bactérienne avec l'ADN obtenu. J'ai ensuite utilisé quinze colonies isolées pour déterminer, par restriction enzymatique, si ce plasmide était modifié par rapport à celui d'origine. Les profils obtenus ne correspondent pas à ceux attendus. J'ai alors sélectionné quatre clones pour déterminer, par PCR, la présence du gène *GUS1*, mais également celui codant la GluProRS^{N1}. Comme attendu, le plasmide permet l'amplification de *GUS1* et, la plupart du temps, du gène de la GluProRS^{N1} (uniquement le plasmide 2 ne le contient pas), confirmant qu'il y a eu recombinaison au sein de ce vecteur (**Figure 59**). Les souches exprimant

l'aaRS humaine directement au niveau du génome contiennent également la copie endogène de *GUS1*. Cependant, la chasse est beaucoup moins efficace que lors de l'expression plasmidique de la GluProRS^{N1}. Il est possible que la recombinaison de *GUS1* dans ces souches intervienne, soit au niveau de la région d'intégration (ici *TRP1*), soit au niveau du locus de *GUS1*. Cependant, je n'ai pas testé le locus précis de la recombinaison génomique de *GUS1* dans ces souches.

Finalement, j'ai tenté de forcer la perte du plasmide en absence du stress engendré par la sélection au 5FOA. Pour cela, j'ai fait croître les différentes souches dans du milieu riche (YPD) pendant sept jours, favorisant la perte d'un des deux, voire des deux, plasmides. J'ai alors étalé environ 300 colonies sur milieu riche. J'ai finalement réalisé des répliques velours sur milieu SC-Ura pour visualiser les souches ayant perdu le plasmide d'expression de *GUS1*, et pouvant donc se développer avec la GluProRS^{N1} comme seule enzyme assurant l'aminocyclation de l'^cARNt^{Glu}. Cependant, toutes les colonies ont maintenu le plasmide d'expression *GUS1*, confirmant que la perte de la GluRS de levure n'est pas favorable à la croissance de ces souches.

L'ensemble des résultats obtenus suggère que le variant humain n'est pas capable de remplacer la GluRS de levure et, de ce fait, qu'il n'est pas capable de charger l'^cARNt^{Glu}. En parallèle, ils ne nous permettent pas de conclure quant à la capacité de charge de l'^mARNt^{Gln} par la GluProRS^{N1}. En effet, que ce soient les spores après dissection ou les colonies obtenues après les expériences de chasse de plasmide, aucune souche dépourvue de GluRS n'a pu être isolée et la capacité de charge de l'^mARNt^{Gln} par la GluProRS^{N1} n'a donc pas pu être testée. Pour surmonter l'absence de complémentarité de l'activité cytosolique de la GluRS par la GluProRS^{N1} et pouvoir tester l'implication de cette dernière dans la glutamylation de l'^mARNt^{Gln}, nous envisageons de construire une souche exprimant la GluProRS^{N1}, où la GluRS de levure serait ancrée au niveau du cytosol (par fusion Vph1 par exemple). En effet, la viabilité fermentaire devrait être assurée par la GluRS de levure et la capacité de charge de l'^mARNt^{Gln} par la GluProRS^{N1} pourrait alors être testée. Cette enzyme étant présente dans les mitochondries humaines et de levure, il est tout à fait possible d'envisager que son rôle mitochondrial puisse être crucial pour la cellule et il serait donc intéressant de pouvoir l'identifier.

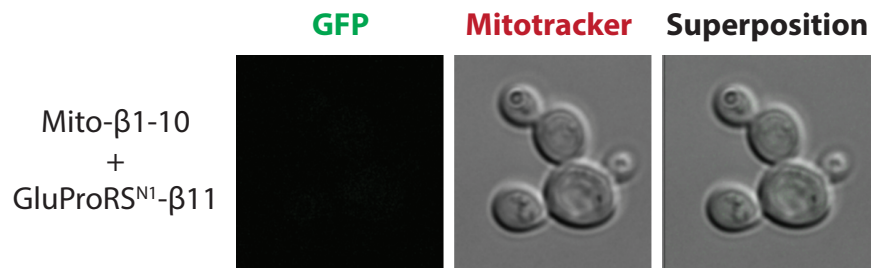


Figure 60 : Relocalisation mitochondriale de la GluProRS^{N1}

La souche Mito- β 1-10, exprimant le fragment β 1-10 à partir du génome mitochondrial, a été transformée avec le plasmide d'expression de la GluProRS^{N1}- β 11. La fluorescence est visualisée par microscopie à épifluorescence.

5. Relocalisation mitochondriale de la GluProRS^{N1} avec l'outil Split-GFP

La localisation mitochondriale de la GluProRS^{N1} n'étant pas claire car le signal du Mitotracker n'est colocalisé que partiellement avec la fluorescence GFP (**Figure 51**), j'ai décidé d'utiliser l'outil de Split-GFP que j'ai élaboré pour étudier sa relocalisation mitochondriale. Pour ce faire, j'ai construit un plasmide d'expression de la GluProRS^{N1}- β 11 que j'ai transformé dans la souche Mito- β 1-10, exprimant la fragment β 1-10 à partir du génome mitochondrial. J'ai alors étudié la présence mitochondriale de cette enzyme par microscopie à épifluorescence. Cette souche ne présente pas de signal de fluorescence visualisable (**Figure 60**). Cette enzyme ne semble donc pas se relocaliser dans la mitochondrie. Cependant, il est possible que la relocalisation de cette enzyme ne soit pas suffisante pour permettre l'apparition d'un signal fluorescent visualisable. Il faudrait donc augmenter l'expression de cette protéine de fusion à l'aide d'un vecteur multicopie. De plus, nous ne savons pas si elle est exprimée. Nous allons tout d'abord vérifier l'expression de cette protéine par immuno-détection sur des extraits totaux de souches exprimant une version étiquetée de la GluProRS^{N1}. Ensuite, nous allons purifier des mitochondries à partir de ces mêmes souches pour étudier la présence mitochondriale de cette enzyme. Dans le cas où aucun signal ne serait observable dans les fractions mitochondriales, l'incapacité de relocalisation mitochondriale de la GluProRS^{N1} chez la levure serait confirmée. Il serait possible que cette enzyme interagisse avec une protéine cytosolique, empêchant ainsi sa relocalisation mitochondriale (sur le principe du complexe AME). Néanmoins, il est également envisageable que l'import mitochondrial de cette enzyme fasse intervenir un cofacteur qui n'est pas présent chez la levure et qui induit une localisation cytosolique stricte de la GluProRS^{N1} chez cet organisme.

Conclusions et perspectives générales

J'ai consacré ma thèse à l'étude des aaRS et plus particulièrement à celle des aaRS cytosoliques des organismes eucaryotiques. Ces enzymes, impliquées dans la synthèse protéique, peuvent s'associer entre elles, mais également avec des protéines non synthétasiques afin de former des MSC. Chez la levure *S. cerevisiae*, le complexe AME lie la GluRS et la MetRS à la protéine d'assemblage Arc1 (Simos et al., 1996). Ce complexe est retrouvé au niveau du cytosol et la protéine Arc1 permet l'ancrage des deux aaRS dans ce compartiment. Néanmoins, la GluRS et la MetRS peuvent toutes les deux rejoindre des compartiments subcellulaires différents pour assurer des fonctions additionnelles pouvant être reliées ou non à la synthèse protéique. Ainsi, la relocalisation mitochondriale de la GluRS et la relocalisation nucléaire de la MetRS est simultanée, et cette synchronisation est rendue possible du fait de la régulation transcriptionnelle d'*ARC1* qui induit la diminution de la concentration de l'ancre cytosolique de ces deux aaRS. Cette relocalisation synchrone est cruciale puisque toute tentative visant à la désynchroniser, conduit à un phénotype de défaut respiratoire. La fraction nucléaire de la MetRS va moduler la transcription des gènes du complexe III de la CR et d'une s-u du domaine F₁ du complexe V, tandis que la fraction mitochondriale de la GluRS est cruciale pour la formation du Gln-mARN^{Gln}, utilisé pour la synthèse de Var1 et des s-u des complexes III, IV, V de la CR (Frechin et al., 2014; Frechin et al., 2009b). La relocalisation organellaire synchrone des deux aaRS peut ainsi être comparée à un dialogue entre le noyau et la mitochondrie permettant de coordonner l'expression de leur génome respectif, et ce, directement en réponse aux conditions physiologiques de croissance.

Lors de ma thèse, j'ai également été amené à participer à l'étude d'un variant d'une aaRS humaine, la GluProRS^{N1}. Cette dernière ressemble fortement à la GluRS de levure et nous nous demandions donc si elle pouvait avoir la même fonction mitochondriale que son homologue de levure, d'autant plus que les études microscopiques du Dr. Enkler avaient montré la capacité de la GluProRS^{N1} étiquetée GFP de se relocaliser dans les mitochondries de cellules Hela. Nous avons pu montrer, chez la levure, que la version étiquetée GFP de cette enzyme était présente dans le cytosol, mais également plus partiellement dans les mitochondries. Nous avons également montré que la GluProRS^{N1} était capable de compenser l'absence de la GluRS de levure. Les différents contrôles étaient également concordants avec cette hypothèse et nous étions donc confiant quant à l'activité mitochondriale de cette enzyme. Nous avons finalement voulu confirmer les résultats obtenus par sélection au 5FOA,

via l'utilisation d'une souche diploïde et l'étude de la ségrégation des tétrades. Cependant, cette expérience a conduit à l'obtention de résultats contradictoires, réfutant la capacité de l'enzyme humaine à remplacer son homologue de levure. Suite à diverses expériences contrôles, nous avons observé que lors de la chasse de plasmides, une recombinaison homologue permettait de conserver le gène codant la GluRS de levure et *de facto*, que les phénotypes observés ne dépendaient plus de la seule activité de la GluProRS^{N1}. Ces résultats sont tout à fait surprenants et montrent l'importance de la GluRS pour la croissance de levure. En effet, malgré une forte conservation du domaine catalytique, la GluRS ne peut pas être remplacée par l'aaRS humaine. Il est donc évident que la GluRS possède des fonctions essentielles, non seulement en lien avec la traduction cytosolique, mais vraisemblablement également avec d'autres processus cellulaires dans d'autres compartiments subcellulaires. La détermination de toutes les localisations additionnelles de la GluRS de levure est donc primordiale pour la compréhension de ses autres fonctions essentielles.

L'étude du complexe AME et de la GluProRS^{N1}, montrent l'importance des aaRS dans divers processus biologiques. Cependant, ces deux études nous a également permis de comprendre les difficultés techniques et conceptuelles de l'étude de protéines multi-localisées, ayant des fonctions essentielles dans plusieurs compartiments subcellulaires. Ainsi, le projet principal de ma thèse a été d'élaborer un outil de localisation, basé sur une Split-GFP, permettant l'étude de ces protéines. Cette technique consiste en la séparation de la GFP en deux fragments (β 1-10 et β 11), capables de s'autoassembler lorsqu'ils sont présents dans le même compartiment. En collaboration avec le Dr. Arais, j'ai donc développé cet outil pour son utilisation chez la levure. Un des fragments (β 11) a été fusionné aux différentes protéines d'intérêts, tandis que l'autre (β 1-10) a été associé à des signaux d'adressage permettant de le restreindre constitutivement à un compartiment donné. Finalement, la production de fluorescence est spécifique de la présence de la protéine testée dans ledit compartiment. Au travers de l'ingénierie de l'outil, j'ai pu comprendre l'importance de sa validation avec de multiples témoins ayant des propriétés différentes. Ainsi, pour les différents compartiments dans lesquels j'ai voulu tester l'import des protéines, l'utilisation d'une protéine présente dans un autre compartiment me permet de n'obtenir aucun signal aspécifique. Cependant, lorsque j'utilise une protéine libre de diffuser n'importe où dans la cellule, un signal fluorescent aspécifique est détectable. Il reste en effet quelques petits ajustements à faire pour que l'outil

soit totalement fonctionnel pour l'étude de protéines dans divers compartiments subcellulaires. Ces ajustements passeront par l'étude de l'affinité des deux fragments, mais également par des mutagenèses aléatoires ou dirigées permettant de réduire leur affinité tout en conservant leur capacité d'autoassemblage. Pour cela, nous envisageons une collaboration avec l'équipe du Dr. Cabantous qui a développé la méthodologie Split-GFP.

De plus, nous nous sommes lancés dans la création d'une palette de variants Split-GFP ayant des propriétés spectrales différentes, permettant la visualisation simultanée de plusieurs signaux de fluorescence, et favorisant l'étude de la relocalisation d'une protéine d'intérêt dans plusieurs compartiments au même moment. Cependant, nous envisageons de modifier les mutations utilisées pour créer des variants présentant une plus forte intensité de fluorescence. Ainsi, l'utilisation des variants Cerulean et Venus pourrait nous permettre d'augmenter drastiquement l'émission des fluorescences jaune et cyan, diminuant le bruit de fond et augmentant l'intensité de fluorescence lors des études de relocalisation.

Finalement, le résultat le plus intéressant de ma thèse a été la capacité d'expression du fragment $\beta 1-10$ à partir du génome mitochondrial. En effet, l'expression de fragments d'origine génomique mixte nous permet d'éliminer toute forme de bruit de fond dû à un autoassemblage pré-import, une diffusion de la protéine d'intérêt, mais également un signal issu de l'affinité de deux fragments et non de la localisation de la protéine étudiée. Ainsi, l'apparition d'un signal fluorescent dépend uniquement de la présence de la protéine d'intérêt dans la mitochondrie. Les perspectives qui en découlent sont également des plus intéressantes et importantes car cette souche (Mito- $\beta 1-10$) pourrait être utilisée pour des études systématiques de relocalisation des protéines de levure et pour la délimitation de l'ensemble des signaux d'import mitochondriaux des aaRS ou d'autres familles de protéines. L'ensemble des résultats obtenus pourrait être un moteur pour : (i) la compréhension des règles gouvernant la relocalisation mitochondriale des protéines, (ii) l'étude des conditions induisant la relocalisation mitochondriale d'une protéine donnée (stress, UV, source de carbone...) et plus globalement, (iii) l'étude de la localisation des protéines de fonctions inconnues, permettant ainsi de spéculer quant à leur fonction dans ce compartiment. De plus, la capacité d'expression du fragment $\beta 1-10$ à partir du génome mitochondrial nous permettra de créer une souche exprimant deux fragments $\beta 1-10$ de couleurs différentes, exposés dans l'IMS et dans la matrice. Cette souche ouvrira alors une nouvelle voie d'étude du protéome

de ces sous-compartiments mitochondriaux. De plus, elle pourra également être utilisée pour l'étude de la maturation des protéines de l'IMS et permettra de déterminer si ce processus se déroule en partie dans la matrice ou totalement dans l'IMS. Finalement, il sera également possible de réaliser des mitoprotéomes sub-mitochondriaux dans diverses conditions de croissance permettant de comprendre l'ensemble des régulations ciblant ce compartiment.

L'ensemble des données récoltées par l'étude systématique de la relocalisation des protéines dans la matrice mitochondriale et l'IMS sera sans aucun doute essentiel pour la compréhension des mécanismes impliqués dans la mitophagie, l'apoptose ou encore dans ceux impliqués dans le vieillissement ou les maladies mitochondriales humaines.

Annexe 1



Review

Exploring the evolutionary diversity and assembly modes of multi-aminoacyl-tRNA synthetase complexes: Lessons from unicellular organisms



Daphné Laporte, Jonathan L. Huot, Gaétan Bader, Ludovic Enkler, Bruno Senger, Hubert Dominique Becker*

UMR 'Génétique Moléculaire, Génomique, Microbiologie', CNRS, Université de Strasbourg, 21 rue René Descartes, 67084 Strasbourg Cedex, France

ARTICLE INFO

Article history:

Received 27 August 2014

Revised 3 October 2014

Accepted 6 October 2014

Available online 12 October 2014

Edited by Wilhelm Just

Keywords:

Multi-aminoacyl-tRNA synthetase complex

Unicellular

AIMP

GST-like domain

WHEP domain

EMAPII domain

Assembly mode

Evolution

ABSTRACT

Aminoacyl-tRNA synthetases (aaRSs) are ubiquitous and ancient enzymes, mostly known for their essential role in generating aminoacylated tRNAs. During the last two decades, many aaRSs have been found to perform additional and equally crucial tasks outside translation. In metazoans, aaRSs have been shown to assemble, together with non-enzymatic assembly proteins called aaRSs-interacting multifunctional proteins (AIMPs), into so-called multi-synthetase complexes (MSCs). Metazoan MSCs are dynamic particles able to specifically release some of their constituents in response to a given stimulus. Upon their release from MSCs, aaRSs can reach other subcellular compartments, where they often participate to cellular processes that do not exploit their primary function of synthesizing aminoacyl-tRNAs. The dynamics of MSCs and the expansion of the aaRS functional repertoire are features that are so far thought to be restricted to higher and multicellular eukaryotes. However, much can be learnt about how MSCs are assembled and function from apparently 'simple' organisms. Here we provide an overview on the diversity of these MSCs, their composition, mode of assembly and the functions that their constituents, namely aaRSs and AIMPs, exert in unicellular organisms.

© 2014 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Introduction

During ribosomal protein synthesis, mRNA codons are decoded through base pairing with the anticodon of aminoacylated tRNAs. Therefore, maintenance of the genetic code and of the information stored as nucleotide triplets in genes relies on the rules that govern the recognition of a given tRNA by its corresponding aminoacyl-tRNA synthetase (aaRS) [1]. The term 'second genetic code' has been introduced to describe this concept [2]. AaRSs belong to one of the most ancient family of proteins and are mainly known for their primary function, which is to charge each tRNA species with its cognate amino acid (aa) [3]. Originally characterized and isolated in the late 50s and early 60s [4–6], aaRSs have been extensively studied since then. Interestingly, the propensity of several aaRSs to form complexes called multi-synthetase complexes (MSCs) was already discovered in the early 1970s [7,8]. Their potential to assemble into

supramolecular complexes was initially discovered in mammals, in which the MSC is composed of eight aaRSs and three auxiliary proteins, named 'aminoacyl-tRNA synthetase-interacting multifunctional proteins' (AIMP) [7,8]. MSC formation was believed to be restricted to higher eukaryotes, as the 20 aaRSs isolated either from *Saccharomyces cerevisiae* or from *Escherichia coli* were found to be only in a free form [9,10]. However, although all aaRSs from *E. coli* were initially purified in their free forms using a methodology combining ion-exchange and adsorption chromatography, the existence of an MSC in *E. coli* was reported a decade later [11]. So far, this is the only study reporting the presence of MSCs in the bacterial branch of the tree of life. This study was not pursued any further, as, during the next decades, research mostly focused on deciphering the molecular determinants that determine the tRNA-aaRS recognition [12].

From an evolutionary point of view, the accepted scheme is that primitive aaRSs contained solely a contemporary catalytic core that was capable of activating aa. New domains were later recruited to these amino acid-activating polypeptides, either by

* Corresponding author. Fax: +33 (0)3 88 65 13 65.

E-mail address: h.becker@unistra.fr (H.D. Becker).

fusion or insertion, allowing the ancestral aaRS to bind RNA helices that corresponded to the contemporary tRNA acceptor arm (accepting minihelix) [13]. These primitive tRNA–aaRS pairs then evolved, mainly through accretion of new domains (e.g. of the anticodon minihelix and anticodon-binding domain) to give rise to the modern tRNAs and aaRSs. Although the details of their evolutionary history are still a matter of debate, the modular organization of aaRSs strongly suggests that they result from a piece-wise sequential addition of domains to the aa-activating core [3,14]. Besides the anticodon-binding domain, other functional domains, including the editing domains [15] were later added to the aaRS sequence, as well as modules that have not been assigned any function so far [16–18]. Contemporary aaRSs should thus be regarded as catalytic folds that attracted, throughout evolution, several other independent modules, which are not necessarily associated with tRNA aminoacylation.

Some of these additional modules are made of folds that are typically involved in protein–protein interactions, as, for example, the glutathione transferase (GST) domain. Such modules might have allowed aaRSs to assemble into complexes, like the MSCs [19,20]. Others domains, like the EMAP II domains, might have allowed aaRSs to accomplish tasks that are completely unrelated to protein synthesis [21]. It has even been proposed that the capacity of these multitasking enzymes to assemble in MSCs allows for both enhanced efficiency of tRNA aminoacylation of the participating aaRSs [22–24] and annexation of this essential and central process of aminoacyl-tRNA (aa-tRNA) synthesis to other cellular pathways [25–27].

The evolutionary origin of MSCs is still unclear, and the advantage brought in by the accretion of aaRSs into MSCs was for a long time an open question. Recently, it was proposed that metabolic events might have facilitated the incorporation of as many as eight aaRSs into the MSCs of higher eukaryotes [25]. The consensual concept that has emerged during the last decade is that MSCs are reservoirs for releasable and multifunctional aaRSs [28–30]. Sequestration inside MSCs confines aaRSs activities to their primary tRNA charging activity, while they can carry out alternative regulatory roles upon release. As most of these non-conventional roles include regulation of gene expression [27,31] or participation in a signaling pathway [32], the released aaRSs often relocate to a new subcellular compartment (e.g. the nucleus or the mitochondrion) [26,27,31]. However, it is not known if the nuclear relocalization of MSC-released aaRSs is independent or not of regular nuclear import mechanisms that have been associated with aaRS-mediated tRNA channeling or the control of tRNA integrity and functionality [33,34]. All eukaryotic MSCs characterized so far were primarily isolated from cytosolic fractions [24,35], thus, these complexes are considered to be exclusively restricted to the cytosol. Consequently, they can act as a dynamic cytosolic anchors for relocating aaRSs [27]. However, the presence of MSCs in other subcellular compartments was so far not exhaustively investigated. There is even one study reporting the presence of a complex containing 13 aaRSs in nuclear extracts from Chinese hamster ovarian cells and rabbit kidney cells [36]. This complex was proportionally less abundant than its cytosolic orthologue, but was shown to be more stable, suggesting that it might not be as dynamic as the cytosolic MSC and, thus, more likely to be involved in tRNA export from the nucleus.

In this review, we focus on the various features of the MSCs that have been identified in low-complexity organisms such as bacteria, archaea, fungi and protozoans. We describe their components (aaRSs and AIMPs), their domain architectures and assembly modes and by comparing them to those found in MSCs from multicellular organisms, we propose that MSCs segregate in two different evolutionary types.

2. Main text

2.1. MSCs composition and functions in low-complexity organisms and in metazoan

Overall, most of the knowledge on MSCs almost exclusively originates from studies made on a few metazoan species and especially on human cells. As a consequence of this narrow sampling of the species in which MSCs have been looked for and characterized, we have a biased view of the function of these complexes. Based on their expected roles of dynamical platforms for relocating aaRSs, the presence of MSCs in prokaryotes, in which the quasi-totality of the cellular processes takes place in a single compartment, would be difficult to explain. In addition, there are contradicting reports concerning the presence of MSC in some prokaryotes and their compositions have not always been fully deciphered. Finally, when tested, the only role that could be attributed to MSCs from low-complexity organisms was an enhancement of the tRNA aminoacylation efficiency of their participating aaRSs [24], making these complexes apparently less interesting to study than their metazoan orthologues.

2.1.1. Bacterial MSC

In 1987, almost a decade after the purification of 20 individual aaRSs from *E. coli*, a study reported the isolation of a MSC from the same bacterial species [11]. Strangely, depending on the method used to break cells, the molecular weight of the complex varied from 1 MDa (freeze press) to 400 kDa (sonication). The complex could only be isolated on 6% agarose columns but not by regular size-exclusion chromatography on Sephadex-based columns. This complex does not seem to contain all *E. coli* aaRSs, since, given the pool of aaRSs present in this species, the expected size of a MSC containing all of them would be of 2.4 MDa. This discrepancy in the size of the complex, suggests that the strength of the protein–protein interactions mediating the assembly of the complex is extremely weak. Moreover, such a MSC has never been isolated from any of the numerous other bacterial species from which aaRSs have been extracted and purified; and this study still is the only one reporting the presence of a MSC in *E. coli*. For these reasons, the existence of bacterial MSCs remains, to our opinion, to be clearly demonstrated, despite the fact that bacterial aaRSs have indeed been found to form complexes [37,38]. However, these complexes cannot be considered as MSCs *per se*, since they only contain a single aaRS species.

2.1.2. Archaeal MSCs

2.1.2.1. *Haloarcula marismortui*. The first archaeal MSC that has been described was isolated from *H. marismortui* [39]. This halophilic and marine organism contains one or two big complexes that include almost all twenty aaRSs (Fig. 1). This complex is resistant to dissociation when subjected to various chromatographic separations (size-exclusion, ion exchange, hydroxyapatite and reversed phase). However, no additional information concerning the presence or absence of non-enzymatic anchoring proteins (AIMPs) or concerning the potential role of this MSC is available.

2.1.2.2. *Methanothermobacter thermautotrophicus*. The methanogenic archaea *M. thermautotrophicus* contains a well-characterized MSC. This complex represents the best-studied archaeal MSC, so far [30,40–42]. This macromolecular edifice is composed of 3 aaRSs: leucyl- (LRS); prolyl- (PRS); and lysyl-tRNA synthetase (KRS), and the elongation factor 1A (EF-1A) (Fig. 1). The presence of LRS seems to be essential, as it probably acts as an assembly factor for PRS, KRS and EF-1A. When subjected to size-exclusion chromatography, the three aaRSs coelute at high molecular weight (600 kDa) suggesting that they assemble with a 2-2-2

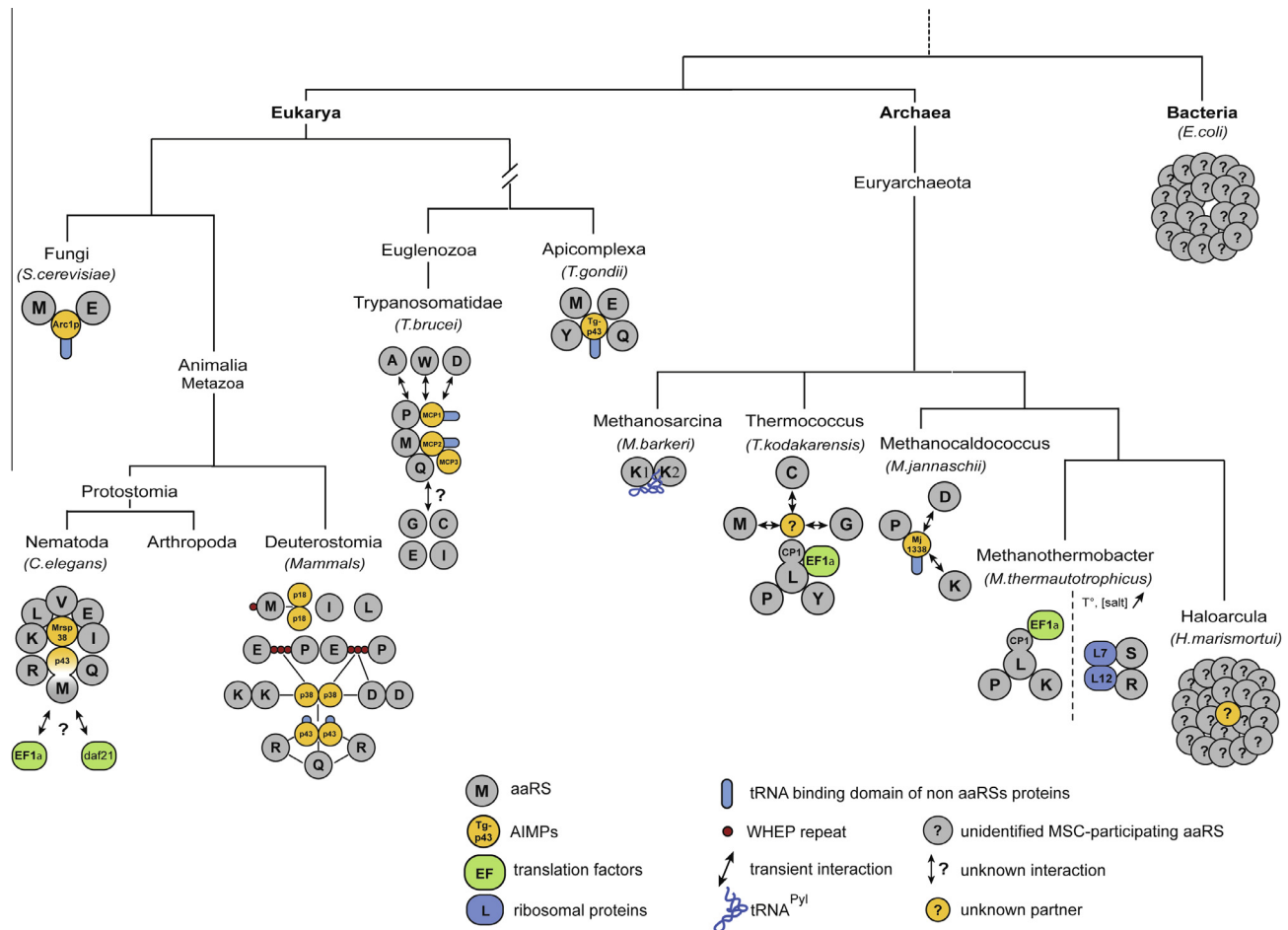


Fig. 1. Composition and architecture of all characterized MSCs. Schematized phylogenetic tree presenting the various organisms in which a MSC has been identified and the corresponding illustration of the MSC's architecture and/or composition. The one-letter code is used for each aaRS (gray). K1 and K2 refer to Class 1 and Class 2 KRSS, respectively.

stoichiometry. KRS and PRS bind to the N- and C-terminal part of LRS, respectively, whereas EF-1A binds the connective peptide (CP1) of LRS that is responsible for its amino acid proofreading activity [43]. This CP1 domain is a post-transfer editing domain, and its activity is not affected by EF-1A binding. Within the complex, the catalytic efficiency of PRS and KRS is improved. When bound to EF-1A, LRS displays an increased tRNA aminoacylation efficiency owing to an 8-fold increased k_{cat} for Leu-tRNA^{Leu} synthesis compared to free LRS. A modest enhancement of EF-1A GTP hydrolysis is also observed when this factor is bound to LRS. In the light of these observations, it was proposed that the *M. thermotrophicus* MSC intensifies aa-tRNA channeling (at least in the case of Leu-tRNA^{Leu}, Pro-tRNA^{Pro}, Lys-tRNA^{Lys}) from aminoacyl-tRNA synthesis to its delivery to translating ribosomes.

A second MSC has been identified in *M. thermotrophicus* via protein-protein interaction detection methods, in which arginyl-tRNA synthetase (RRS) was used as bait. This second complex contains two aaRSs: seryl-tRNA synthetase (SRS); and RRS. The latter was found to interact with its tRNA^{Arg} substrate within the complex [23] (Fig. 1). The ribosomal proteins L7 and L12, which are normally located at the ribosome stalk, have also been found to participate in this complex [44]. Although the activity of RRS was shown to remain unaffected within this complex, the activity of SRS shows a 4-fold increase as compared to the free enzyme. Similarly, the catalytic activity of SRS increases with temperature and salt concentration when associated with RRS. It was proposed that this MSC is involved in tRNAs recycling, as its association with

translating ribosomes may help to avoid tRNA diffusion, and promote fast tRNA re-aminoacylation to enable immediate re-used for the decoding of the next downstream mRNA codons.

2.1.2.3. *Methanocaldococcus jannaschii*. In *M. jannaschii*, PRS was co-purified with the protein encoded by the Mj1338 ORF. This protein has been annotated as a putative metabolic enzyme involved in methanogenic metabolism [45] (Fig. 1). This enzyme has been identified as an inactive paralogue of Hmd (H_2 -forming methylene-tetrahydromethanopterin dehydrogenase). It has been shown that in *Methanothermobacter marburgensis*, HmdII and HmdIII, two other paralogues of Hmd, do not exhibit Hmd activity [46]. Interestingly, transcription of the genes coding for HmdII and HmdIII is regulated by H_2 concentration: transcription of *hmdII* is up regulated and that of *hmdIII* is down regulated under H_2 -limited growth conditions. These findings suggest that HmdII and HmdIII are indirectly involved in methanogenesis. In *M. jannaschii*, HmdII/Mj1338 and HmdIII can bind to both PRS and tRNA^{Pro}, thereby forming ternary complexes. However the precise architectures of these particles have not yet been established [47]. HmdIII/Mj1338 also binds other archaeal aaRSs such as KRS and aspartyl-tRNA synthetase (DRS), suggesting the existence of another HmdII/Mj1338-containing MSC in this archaeal species. The systematic presence of an enzyme involved in methanogenesis in the *M. jannaschii* MSCs pinpoints toward a possible link between energy production and translation. Thus, it is possible that in methanogenic archaea MSCs can couple energy production to energy use for

protein synthesis in response to environmental conditions. However, an Mj1338-containing MSC remains questionable, since it was not ruled out whether these aaRSs co-bind together to Mj1338 or bind separately in the original paper. Mj1338 was shown to also bind, separately, human KRS and PRS by the same method, suggesting a potential non-specific protein interaction.

2.1.2.4. *Thermococcus kodakarensis*. In *T. kodakarensis*, affinity purification chromatography using LRS as bait resulted in the co-elution of 5 other aaRSs: PRS; tyrosyl- (YRS); glycyl- (GRS); methionyl- (MRS); and cysteinyl-tRNA synthetase (CRS), together with EF-1A, elongation factor 2 (EF-2), initiation factor 2 (IF-2) and -2B (IF-2B) and several ribosomal proteins [48] (Fig. 1). In agreement with the composition of this MSC, purified ribosome fractions of *T. kodakarensis* exhibited the tRNA aminoacylation activities of PRS, LRS and YRS, confirming the association between these aaRSs and the protein synthesis machinery. Additionally, the catalytic efficiency of LRS was found to be increased when it associates with EF-1A in its GTP-bound form but not in the GDP-bound state. This indicated that only the activated elongation factor could act as an enhancer for LRS. As a consequence, the *T. kodakarensis* MSC was proposed to increase the activities of its associated aaRSs in order to boost the rate of protein synthesis.

2.1.2.5. *Methanosarcina barkeri*. In *M. barkeri*, monomethylamine methyltransferase genes contain amber stop codons that are reassigned to pyrrolysine (Pyl) [49]. For their translation, the monomethylamine methyltransferase mRNAs require synthesis of a Pyl-charged amber suppressor tRNA pyrrolysine (tRNA^{Pyl}). *M. barkeri* was shown to use two different strategies to generate pyrrolysyl-tRNA^{Pyl} (Pyl-tRNA^{Pyl}). The first pathway is essential, and involves the direct aminoacylation of tRNA^{Pyl} with Pyl by a pyrrolysyl-tRNA synthetase [50]. A second and additional pathway involves both class 1 KRS (KRS1) and class 2 KRS (KRS2), which misacylate tRNA^{Pyl} with lysine, as well as a still unidentified tRNA-dependent amino acid-modifying enzyme that converts this mischarging lysine into Pyl. Interestingly, lysylation of tRNA^{Pyl} by KRS1 and KRS2 can only be achieved through the formation of a ternary complex, in which tRNA^{Pyl} is sandwiched by both KRSs, as each KRS is individually unable to mischarge tRNA^{Pyl} [51] (Fig. 1). In this complex, KRS1 acts as a stabilizer, whereas KRS2 catalyses misacylation of tRNA^{Pyl} with lysine. However, the rate of tRNA^{Pyl} charging is 100-fold slower, compared to the one of the cognate tRNA^{Lys} by each of these KRSs. It was thus surmised that this complex is more likely to be involved in modulating the level of suppressing Pyl-tRNA^{Pyl} available in the cell.

2.1.3. Fungal MSCs

2.1.3.1. *S. cerevisiae*. The AME (Arc1p-MRS-ERS) complex from *S. cerevisiae* is one of the most studied MSCs. This complex was discovered in the mid 1990's and is made of two aaRSs—MRS and glutamyl-tRNA synthetase (ERS)—and one AIMP, Arc1p, which is homologous to the AIMP1/p43 assembly factor of the mammalian MSC [24,52] (Fig. 1). Arc1p was initially identified as a cofactor for nuclear export of tRNA by Hurt and colleagues, who used the synthetic lethality approach to identify a genetic relationship between the *LOS1* (encoding an exportin) and *ARC1* (encoding Arc1p) genes [24].

Arc1p acts as an ark for the two aaRSs [53,54]. It specifically binds the N-terminal GST-like domains of MRS and ERS via the GST-like domains that are located in its N-terminus. Complete deletion of N-terminal extensions of the two aaRSs is not lethal, but abolishes their interaction with Arc1p. Formation of the AME complex causes nuclear exclusion of the three AME proteins, as Arc1p is actively excluded from the nucleus through the Xpo1p-dependent nuclear export pathway [55]. Indeed, even when

artificially fused to a strong nuclear localization signal, Arc1p is still very efficiently exported from the nucleus, suggesting that Arc1p contains a strong nuclear export signal (NES), although a classical leucine rich NES (target of Xpo1) has not been clearly identified. The accumulation of Arc1p-NLS fusion protein in the nucleus of yeast *xpo1* mutant strain confirmed the presence of a strong NES. Notably, deletion of the N-terminal sequence of Arc1p allows its entry in the nucleus, suggesting that the Arc1p NES is located within this protein region [56]. Along the same line, deletion of *ARC1* results in nuclear relocation of MRS and to a lesser extent, of ERS [56]. However, the mechanism and structural determinants governing the nuclear relocation of Arc1p, MRS and ERS remain to be identified. Moreover, although nuclear localization of Arc1p can be fully understood given its expected role in exporting tRNA from this organelle, the function of ERS and MRS in the nucleus remains puzzling. One possibility would be that they participate in a quality control step for newly transcribed tRNAs, as also suggested for other nucleus-localizing aaRSs in other organisms [33,34]. It is still unknown whether all or only a subset of the aaRSs are capable of entering the nucleus, and whether MSCs, like the AME complex, control nuclear localization of their aaRS components. As only a few such MSCs have been characterized so far, one could speculate that, if they exist, they must be highly dynamic.

For a long time after its discovery, the function of the *S. cerevisiae* AME complex had been associated with the improvement of the tRNA aminoacylation efficiency (k_{cat}/K_M) of MRS and ERS [24,57]. This efficiency enhancement is mediated by the tRNA binding domain in the C-terminus of Arc1p that probably compensates for the low affinity of MRS and ERS for their cognate tRNAs [58]. Therefore, the catalytic efficiency of Arc1p-bound MRS is increased by 500-fold as compared to the free enzyme. However, another study reported only a 3 to 4-fold increase of the aminoacylation efficiency of MRS upon binding to Arc1p, leaving unclear the degree by which Arc1p enhance the charging efficiency of MRS [59]. This gain in tRNA aminoacylation efficiency is mainly due to an Arc1p-mediated lowering of the K_M of MRS for its cognate tRNA [24], which allows AME-assembled MRS to be far more efficient at low tRNA^{Met} concentrations. Similarly to MRS, the catalytic efficiency of Arc1p-bound ERS is enhanced compared to the free enzyme, owing to a 100-fold increase in the affinity of Arc1p-bound ERS for its cognate tRNA^{Glu} [57].

Interestingly, both Arc1p-bound MRS and free MRS have been shown to mismethylate tRNAs in vivo and in vitro under oxidative stress [60]. These mischarged Met-tRNAs are subsequently used during translation generating proteins containing misincorporated methionines. Previous reports showed that these additional methionine residues can protect the proteins against reactive oxygen species (ROS) [61,62]. If these misincorporated methionines replacing non-methionine residues are, for example, located near the active site of an enzyme, their oxidation will shield this strategic position against ROS. The AME complex would not be necessary for this defense mechanism in vivo, since the extent of mismethylation does not vary between a wild type and an *arc1Δ* strain, suggesting that MRS alone might be able to mismethylate tRNAs [60]. However, it is still not known whether binding to Arc1p also enhances the charging efficiency of MRS for non-cognate tRNAs, like it does for its cognate tRNA^{Met}. If this would be the case, the AME complex would not then increase the extent of protection against ROS but could modify the speed of the response to oxidative stress.

The fact that *S. cerevisiae* MRS and ERS are far more efficient in tRNA charging when assembled in the AME complex than in their free forms suggested that the role of this MSC would strictly be confined to improve aa-tRNA synthesis. For that reason the AME MSC was not considered to be a dynamic complex and the

existence and localization of Arc1p-unbound MRS and ERS was not exhaustively checked. Recently it was shown that when yeast cells switch from fermentation to respiration expression of Arc1p is decreased [27] triggering the release of ERS. Free ERS is then able to relocate to the mitochondria and act as a mitochondrial non-discriminating ERS that misacylates mitochondrial-encoded tRNA^{Gln} with glutamate [27]. The glutamate acylating tRNA^{Gln} is then transamidated by the mitochondrial GatFAB amidotransferase into glutamine yielding a correctly paired Gln-tRNA^{Gln} species. These findings show that Arc1p, and thus the AME complex, plays the role of an anchoring platform that regulates the subcellular location of ERS in response to which metabolism—fermentation or respiration—is used by *S. cerevisiae*. Since the mitochondrial transamidation reaction involving the cytosolic ERS and the GatFAB AdT, is the only pathway through which yeast cells generate the mitochondrial Gln-tRNA^{Gln}, the dynamics of AME complex is essential for the expression of the mitochondrial genes and thus the ability of yeast cells to respire.

To summarize, the yeast AME complex would have two different roles. In fermenting yeast cells binding of ERS and MRS to the AIMP protein Arc1p improves their aminoacylation efficiency. Concomitantly, Arc1p (and thus the AME complex) also serves as an anchoring platform sequestering ERS in the cytoplasm in fermenting cells, and releasing ERS so that it can relocate in the mitochondria when cells need to respire. What happens to MRS in respiring yeast cells is still unknown. Obviously, the decrease in Arc1p levels upon switch to respiration should also trigger the release of MRS, and, since deletion of *ARC1* triggers nuclear relocation of MRS, one could imagine that MRS could relocate to this compartment in respiring cells to perform yet undetermined functions.

2.1.3.2. *Schizosaccharomyces pombe*. Analysis of the initiation factor 3 (eIF3p) interactome in the fission yeast *S. pombe* [63] revealed the presence of a translasome: a super-complex linking protein synthesis and degradation. Seven aaRSs were identified within this complex: threonyl-tRNA synthetase (TRS); SRS; MRS; DRS; phenylalanyl-tRNA synthetase (FRS); ERS; and GRS together with a protein that was identified as a cofactor for MRS and ERS. The 7 aaRSs and the MRS and ERS cofactor protein may constitute a MSC in *S. pombe*. Notably, MRS, ERS and their cofactor are the most highly abundant proteins in this putative MSC. Thus, this study suggests the existence in fission yeast of an MSC that would be similar to the *S. cerevisiae* AME complex, composed of MRS, ERS, and a cofactor protein that resembles Arc1p. However, it is still not known whether the other aaRSs found in the translasome are part of this MSC. The role of the putative *S. pombe* MSC is still unknown, but it has been shown that the function of the translasome, of which this MSC is part of, is to increase translation efficiency.

2.1.4. Protozoan MSCs

2.1.4.1. *Trypanosoma brucei*. A recent study revealed that the protozoan pathogen *T. brucei*, which causes human African trypanosomiasis, harbors a 1.2 MDa MSC [22].

Remarkably, this complex is organized like a mammalian MSC. Indeed, it contains 3 AIMPs called MCP1–2–3 (for MARS Complex-associated Proteins), which bind to at least 6 aaRSs: PRS, MRS, DRS, glutaminyl- (QRS), alanyl- (ARS) and tryptophanyl-tRNA synthetase (WRS) (Fig. 1). Four of these aaRSs (PRS, MRS, DRS and QRS) are also found in the mammalian MSC. Sequence homology searches have revealed that the three AIMPs of *T. brucei* have homologues in various organisms. For example, MCP1 has a tRNA-binding domain and a GST-like domain and therefore resembles the yeast protein Arc1p that binds ERS and MRS. Given the sequence similarity between MCP2 and the bacterial tRNA-binding

protein Trbp111 or eukaryotic tRNA binding proteins, it is possible that MCP2 contains a tRNA-binding domain and does not simply serve as a scaffold protein within the protozoan MSC. This hypothesis has been confirmed in *in vitro* studies, in which MCP2 was shown to bind various tRNAs. Interestingly, MCP3 shares some sequence similarity with the bacterial protein YbaK, which has been shown to be a free-standing editing domain that deacylates mischarged cysteinyl-tRNA^{Pro} or alanyl-tRNA^{Pro}, thereby compensating the lack of these editing activities of ProRS [64,65]. Thus, in the *T. brucei* MSC, MCP3 might be involved in misacylated tRNA editing and the proofreading of aa-tRNAs generated by the MSC aaRSs.

Overall, the role of the *T. brucei* MSC is probably to enhance the catalytic efficiency of the aaRSs building the complex. Indeed, it was shown that, *in vitro*, MCP2 increases the rate of tRNA^{Met} aminoacylation. Other, *in vivo* experiments showed that an MCP2-deficient *T. brucei* strain is less infectious than a wild-type strain. Taken together, these results suggest that the MSC of *T. brucei* is important for the parasite's fitness. Consequently, it represents an attractive drug target in the search for treatments against human African trypanosomiasis.

2.1.4.2. *Toxoplasma gondii*. Recently, an MSC was isolated from the apicomplexan, *T. gondii*, which is the pathogen causing toxoplasmosis [66]. This 1 MDa complex is localized in the cytoplasm of the parasite and is composed of 4 aaRSs: MRS; ERS; QRS; and tyrosyl-tRNA synthetase (YRS), and a single AIMP called Tg-p43 (*T. gondii* p43). The 4 aaRSs are bound to Tg-p43 through their N-terminal GST-like domains, which are similar to the GST-like domains of the *S. cerevisiae* Arc1p (Fig. 1). Tg-p43 also contains a tRNA binding domain that shares sequence similarity with both human AIMP1/p43 and yeast Arc1p. Notably, the *T. gondii* MSC resembles to the *S. cerevisiae* AME complex in several ways: the subcellular localization of the complex is the same; both MSCs contain MRS and ERS; and their AIMPs are highly similar in terms of sequence homology and structural domains.

Interestingly, the mutant parasite in which the gene encoding Tg-p43 has been deleted is viable and its pathogenicity is not affected by the loss of Tg-p43. The MSC is therefore not involved in the parasite virulence. However, it is still not known whether binding of the 4 aaRSs to Tg-p43 modulates their catalytic efficiencies, leaving the role of the *T. gondii* MSC still enigmatic.

2.1.5. Metazoan MSCs

2.1.5.1. Nematodes. The worm *Caenorhabditis elegans* also has a 1 MDa MSC. This MSC is composed of MRS, LRS, IRS, ERS, QRS, RRS, KRS and valyl-tRNA synthetase (VRS) and contains an AIMP called mrsp-38 (MARS Scaffold Protein 38 orthologue of the AIMP2/p38) [67]. The main differences with the mammalian MSC are that AIMP3/p18, DRS and PRS are not part of this complex whereas VRS is. Another discrepancy is that MRS is fused to a domain that resembles the mammalian AIMP1/p43 [68]. This appended domain replaces AIMP3/p18 for anchoring of MRS to the complex. In addition to the classical MSC components, the chaperone protein HSP90/Daf21 and the elongation factor EF-1A have been co-purified with the *C. elegans* MSC. However, the role of these additional proteins within the MSC is not known. Hsp90/Daf21 might be involved in the MSC assembly, whereas EF-1A might contribute to aa-tRNA channeling to the ribosome. Some experiments show that inactivation of MRS leads to sterility phenotype, but the exact role of the *C. elegans* MSC remains unknown.

2.1.5.2. The mammalian MSC. The mammalian MSC is composed of eight aaRSs: MRS; DRS; KRS; RRS; LRS; QRS; isoleucyl- (IRS); and glutamyl-prolyl-tRNA synthetase (EPRS), and 3 AIMPs: AIMP1/

p43; AIMP2/p38; and AIMP3/p18 (Fig. 1) [69]. The work undergone during the last decade on the mammalian MSC revealed that these complexes are reservoirs for releasable aaRSs used as regulatory and signaling factors [28,29]. Indeed, the human MRS can be released from the MSC, and relocates to the nucleolus to activate gene transcription [31]. In human macrophages, bi-functional EPRS has been shown to be released from the MSC following treatment with γ -interferon [70], thereby promoting translational silencing of a specific set of mRNAs that encode pro-inflammatory proteins [26]. Most of the MSC-participating aaRSs have been shown to be released in response to a specific stimulus and often to relocate to new subcellular compartments, where they exert a non-conventional role that is unrelated to protein synthesis [29]. The increased

number of aaRSs and AIMP3 building these MSCs in higher eukaryotes may be viewed as a mean to expand the diversity of the response repertoire to stimuli.

2.2. Anchoring domains found in MSC components

In metazoa, AIMP3s are the pivotal components that enable aaRSs to assemble into MSCs [71]. These assembly factors provide the protein-binding domains required for aaRS recruitment to MSCs and sometimes also contain RNA-binding domains that may be involved in MSC functions. As the composition and architecture of the metazoan MSCs is relatively well conserved, the presence of an MSC in any higher eukaryotic species can easily

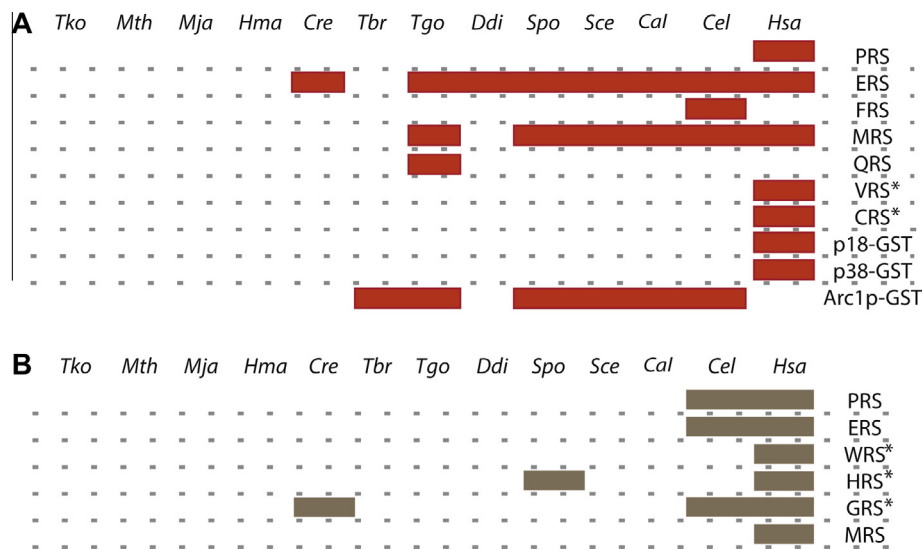


Fig. 2. Distribution of GST-like and WHEP protein-binding domains in the MSC components from 11 unicellular and 2 multicellular species. The organisms (abbreviated by their initials; top lane) that are displayed (one column for each organism) were selected because they have been shown to contain a MSC (all but *Ddi* and *Cre*), or out of phylogenetic interest. The presence of a GST-like (A) or WHEP domain (B) in one of the listed aaRS or AIMP protein (last columns in (A) and (B)) of the corresponding organism is symbolized by a red or gray box respectively. AaRSs listed in (A) or (B) correspond to those in which a GST-like or WHEP domain has been identified in at least one of the presented species (top lane). AIMP3s were divided into three categories: p18-GST and p38-GST which are only present in the Hsa MSC, and Arc1p/Arc1p-like protein present in Spo, Sce, Cal, Cel, Tgo and Tbr MSCs. AaRSs labeled with an asterisk (*) corresponds to those which have not been found in the MSC. Protein-binding domains were identified by PSI-Blast search using the protein sequence of the *S. cerevisiae* Arc1p GST-like domain (Accession: CAA64750.1 GI: 1620460) or that of the human EPRS GST-like domain (Accession: AAI36466.1, GI:223460490) or using the protein sequence of the human MRS WHEP domain (Accession: BAA95668.1, GI: 7804468) or that of the human EPRS WHEP domain (Accession: AAI36466.1, GI:223460490) with the NCBI protein database. Organisms studied were, in full name: *Tko*; *Thermococcus kodakarensis* (KOD1), *Mth*; *Methanothermobacter thermautotrophicus* (str Delta H); *Mja*; *Methanocaldococcus jannaschi* (DSM 2661), *Hma*; *Haloarcula marismortui* (ATCC 43049) *Ddi*; *Dictyostelium discoideum* (AX4), *Spo*; *Schizosaccharomyces pombe* (972 h-), *Sce*; *Saccharomyces cerevisiae* (S288C), *Cal*; *Candida albicans* (WO-1), *Cel*; *Caenorhabditis elegans*, *Hsa*; *Homo sapiens*, *Tbr*; *Trypanosoma brucei* (strain 927/4 GUTat10.1), *Tgo*; *Toxoplasma gondii* (strain ME49), *Cre*; *Chlamydomonas reinhardtii* (strain CC-1373).

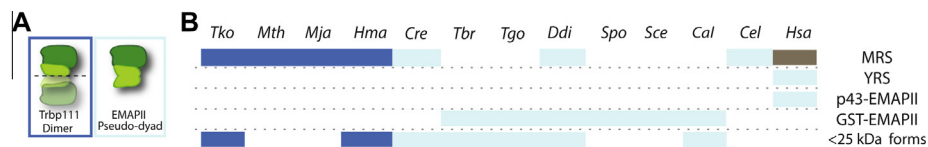


Fig. 3. EMAPII-like RNA-binding domains and their distribution among aaRSs and aaRS-related proteins. Panel A: two types of EMAPII-like domains have been found. These are the dimer-forming trbp111 type and the monomeric EMAPII type. The dimer-forming C-terminal part of trbp111 is shown in light green while the remainder of the monomer is shown in dark green. These colors are retained for the corresponding areas of EMAPII. The EMAPII type domain is homologous to trbp111 but has a C-terminal extension which folds into a structural mimic of the trbp111 dimer core, forming what has been called a pseudo-dyad. The dimer interface of trbp111 and the pseudodyad of EMAPII are both RNA-binding. Panel B: distribution of EMAPII-like domains in 11 unicellular and 2 multicellular organisms. The organisms (abbreviated by their initials; top lane) that are displayed (one column for each organism) were selected because they have been shown to contain a MSC (all but *Ddi* and *Cre*), or out of phylogenetic interest. Shown (last column) are five sorts of protein in which these domains were found (lines): small proteins estimated at 25 kDa or less mostly composed of a trbp111 or EMAPII domain, Arc1p-like GST-fused EMAPII proteins, and p43-like proteins. Dark blue bands indicate where trbp111 type domains were found, while light blue bands indicate where EMAPII type domains were found. Shown in dark gray for comparison is the RNA-binding WHEP insertion domain of *Hsa* MRS. Domains were identified by PSI-Blast using the protein sequence of *Aquifex aeolicus* trbp111 (Accession: 1PYB_D; GI:34810889) or of the *H. sapiens* p43 EMAPII domain (Accession: AAC60647.1, GI:833999) with the NCBI database. Organisms studied were, in full name: *Tko*; *Thermococcus kodakarensis* (KOD1), *Mth*; *Methanothermobacter thermautotrophicus* (str Delta H), *Mja*; *Methanocaldococcus jannaschi* (DSM 2661), *Hma*; *Haloarcula marismortui* (ATCC 43049), *Ddi*; *Dictyostelium discoideum* (AX4), *Spo*; *Schizosaccharomyces pombe* (972 h-), *Sce*; *Saccharomyces cerevisiae* (S288C), *Cal*; *Candida albicans* (WO-1), *Cel*; *Caenorhabditis elegans*, *Hsa*; *Homo sapiens*, *Tbr*; *Trypanosoma brucei* (strain 927/4 GUTat10.1), *Tgo*; *Toxoplasma gondii* (strain ME49), *Cre*; *Chlamydomonas reinhardtii* (strain CC-1373).

be verified by searching for structural homologues of AIMPs. On the contrary, the composition of MSCs that have been characterized in unicellular organisms is so diversified that the presence of an MSC in a given species cannot be easily established. The main reason for this is that unicellular MSCs do not systematically use AIMPs (or easily identified appended domains) to bind and aggregate aaRSs (Fig. 1). How these MSCs assemble and what protein-binding domains are used to build these particles has still not been extensively analysed. In this section, we list the domains known to mediate protein- or RNA-binding that have been identified in components of MSCs from unicellular organisms (Figs. 2 and 3).

2.2.1. Protein-binding domains

In order to form MSCs, aaRSs from eukaryotes interact with each other and/or with AIMPs through specific protein-binding domains such as GST-like domains or WHEP domains. Leucine zipper (LZ) are also involved for the formation of complexes in higher eukaryotes but have not yet been found in aaRSs from lower eukaryotes and are therefore not described in this section. UNE-X domains, unique domains specific to each aaRS, are not described in this section because they are not always protein- or RNA-binding domains [72].

2.2.1.1. GST-like domains. In lower eukaryotes, GST-like domains have been shown to mediate the formation of homo and heterodimers. For example, in the AME complex of the budding yeast *S. cerevisiae*, MRS and ERS associate with the anchoring protein AIMP Arc1p via GST-like domains of the N-terminal regions of all three proteins [19]. Among lower eukaryotes, GST-like domains have also been found in ERS and MRS from *S. pombe*, *Candida albicans* and *T. gondii*, in ERS of *Dictyostelium discoideum*, but also in Arc1p-like proteins from *S. pombe*, *C. albicans* and *D. discoideum* as well as in Tg p43 of *T. gondii* (Fig. 2). Interestingly, the GST-like domain of a *D. discoideum* Arc1p-like protein has a higher similarity with GST-like domains from metazoa (33% of identities and 61% of similarities with the GST-like domain of EPRS from *Homo sapiens*) than GST-like domains from yeast, in accordance with the proteome of *D. discoideum* [73]. The presence of GST-like domains in several aaRSs and anchoring Arc1p-like proteins in *C. albicans* and *D. discoideum* suggests that aaRSs might assemble into MSCs in these organisms. The identification of such MSCs will give us insights into not only protein-binding domains, but also the strategy used for MSC assembly (i.e. with an anchoring factor or with an aaRSs).

In metazoans, the number of aaRSs participating to MSCs is higher than in lower organisms (between 9 and 12 in *H. sapiens* depending on the counting and 8 in *C. elegans*), and concomitantly, the number of aaRSs harboring GST-like domains is also higher (4 for *H. sapiens* and for *C. elegans*). GST-like domains could then be seen as a means to strengthen interactions in MSCs, eventually leading to more diverse role like interactions with other partners to accomplish additional functions such as the interaction between the GST-like domain of human CRS with eEF-1H [74].

2.2.1.2. WHEP domains. The WHEP protein-binding domain has been reported to be strictly restricted to aaRSs of higher eukaryotes. This domain has been named after its presence in WRS, histidyl-tRNA synthetase (HRS) and EPRS and consists of 46 amino acids organized into a Helix-Turn-Helix (HTH) fold. WHEP domains are involved in protein-protein interactions, but also in RNA binding. They are mostly located in the aaRS N-terminus (as, for example, in HRS, GRS and WRS from *H. sapiens*) and less commonly in the aaRS C-terminus (MRS in *H. sapiens* and ERS in *C. elegans*). Exceptionally they can also occupy a central position like in human EPRS, in which 3 WHEP repeats link the ERS and the PRS fused sequences. In this fusion enzyme, WHEP repeats were first shown

to be able to bind RNA [75–77], but a recent study revealed that WHEP repeats bind proteins as well: for example, non-structural associated protein (NSAP1), associates with the third WHEP repeat of EPRS [78]. However, compared to the GST domain which is the major domain mediating MSC assembly, the use of the WHEP domain to anchor aaRSs in MSCs is limited to a few MSC-participating aaRSs. In addition, half of the WHEP-containing aaRSs do not participate to MSCs (GRS, WRS and HRS), suggesting that the presence of WHEP domains in aaRSs from a given species is not a strong indicator that they will assemble into a MSC.

In lower eukaryotes aaRSs, only one WHEP domain could be identified, namely in the *S. pombe* HRS (Fig. 2), and interestingly this aaRS is not part of the translosome [79]. However, it remains unclear whether this domain in *S. pombe* is involved in RNA-binding or protein-protein interactions. HRS from various organisms share low similarities in their N-terminal parts, and a study performed on *T. brucei* revealed that the N-terminal part of HRS forms a HTH domain [80], whereas in *H. sapiens*, this region harbors a WHEP domain. The finding that WHEP domains were acquired early in the history before the separation of plants and animals [81] re-opens the debate of how and why some aaRSs have gained and retained this domain.

2.2.1.3. Other unclassified protein-binding domains. In archaeal aaRSs, no authentic protein-binding domains like the GST-like domains or WHEP domains could be found (Fig. 2). However, it has been shown that *M. thermotrophicus* LRS interacts with KRS and PRS [82]. As mentioned previously (Section 2.1.1), KRS and PRS interact with the N- and C-terminal part of LRS, respectively. The C-terminal extension of PRS could mediate this interaction [83]. In contrast, in *H. sapiens* and *C. elegans*, the C-terminal part of LRS binds to the leucine-zipper (LZ) of RRS in their respective MSCs [84]. Further work is needed to characterize the domains or folds by which archaeal aaRSs participating in MSCs bind to each other.

2.2.2. tRNA-binding domains

EMAPII-like domains attracted the attention of researchers working in the field when it was discovered that this domain had a cytokine activity. The name EMAPII comes from: Endothelial Monocyte-Activating Polypeptide II, and corresponds to the cytokine activity that was first described for this domain [85]. It was later found that the precursor of EMAPII is the AIMP1/p43 of the MSC present in higher eukaryotes [21]. EMAPII-like domains were then also identified as individual and autonomous proteins, like Trbp111 in *Aquifex aeolicus*, and CsaA in *Bacillus subtilis*. Importantly, EMAPII is often found at the C-terminus of MRS in various organisms [86–88].

Structural studies showed that EMAPII-like domains can be divided into two categories: the shorter type resembling trbp111 (trbp111-like), which serves as a dimerization domain; and the longer EMAPII type, which has a C-terminal extension that prevents dimerization. This longer EMAPII type bears structural similarity to the interface of the dimerized trbp111 (Fig. 3A), which is also called “pseudo-dyad” and forms an RNA-binding site. Consequently, the presence of a trbp111-like domain in some bacterial and archaeal MRSs enables their dimerization. In higher eukaryotes, MRSs most probably cannot dimerize, as the trbp111-like domain is replaced by the pseudo-dyad type of EMAPII [89].

The appearance in eukaryotes of the pseudo-dyad type of EMAPII that is unable to dimerize coincides with the appearance of EMAPII-like domains in aaRSs other than MRS, as well as in AIMPs. In many of the organisms in which MSCs have been found, EMAPII-like domains seem to be present either in MRS or in AIMPs, including Arc1p and AIMP1/p43, but very rarely in both (Fig. 3B). This could be due to the transfer of the tRNA-binding role, within a

given MSC, from one protein to another. For example, in the yeast AME complex, EMAPII is located in the C-terminus of Arc1p and absent from MRS. If this EMAPII domain would be fused to the C-terminus to MRS instead this domain would likely also be able to bind tRNA in a similar manner as it does in the Arc1p scaffold [90]. However, the reason why EMAPII is located in Arc1p rather than in MRS is that Arc1p requires a tRNA binding domain to export tRNAs from the nucleus [91]. EMAPII domains are not only RNA-binding domains but also signaling entities, and EMAPII domains released from human AIMP1/p43 and YRS following cleavage have distinct functions [21,92]. This signaling role is not restricted to EMAPII domains of higher eukaryotes, as the unicellular parasite *Entamoeba histolytica*, overexpresses MRS and KRS in response to inflammatory signals, and the C-terminal EMAPII domains of these enzymes are released following cleavage to modulate host immune response [93]. Although no MSC has yet been identified in *Entamoeba*, similar signaling functions may eventually

be found for EMAPII-containing AIMP1s or MSC-participating aaRSs of other lower eukaryotes.

2.3. The diversity of assembly modes in MSCs from low complexity organisms

Whereas in higher eukaryotes there seems to be a conserved architecture and strategy to assemble MSCs, in unicellular organisms, the composition and assembly modes of MSCs is far more diverse (Fig. 1) and unpredictable. The diversity of the proteins (AIMP1s, aaRSs or unrelated proteins) and protein-binding domains mediating assembly of these complexes makes the validation of the presence of a MSC in a given species using bioinformatics unreliable. For example, MSCs in eukaryotes, especially the metazoa, assemble with the help of interaction domains that are found either on aaRSs or on AIMP1s. The extent to which this occurs is slightly lesser in the case of unicellular eukaryotes such as

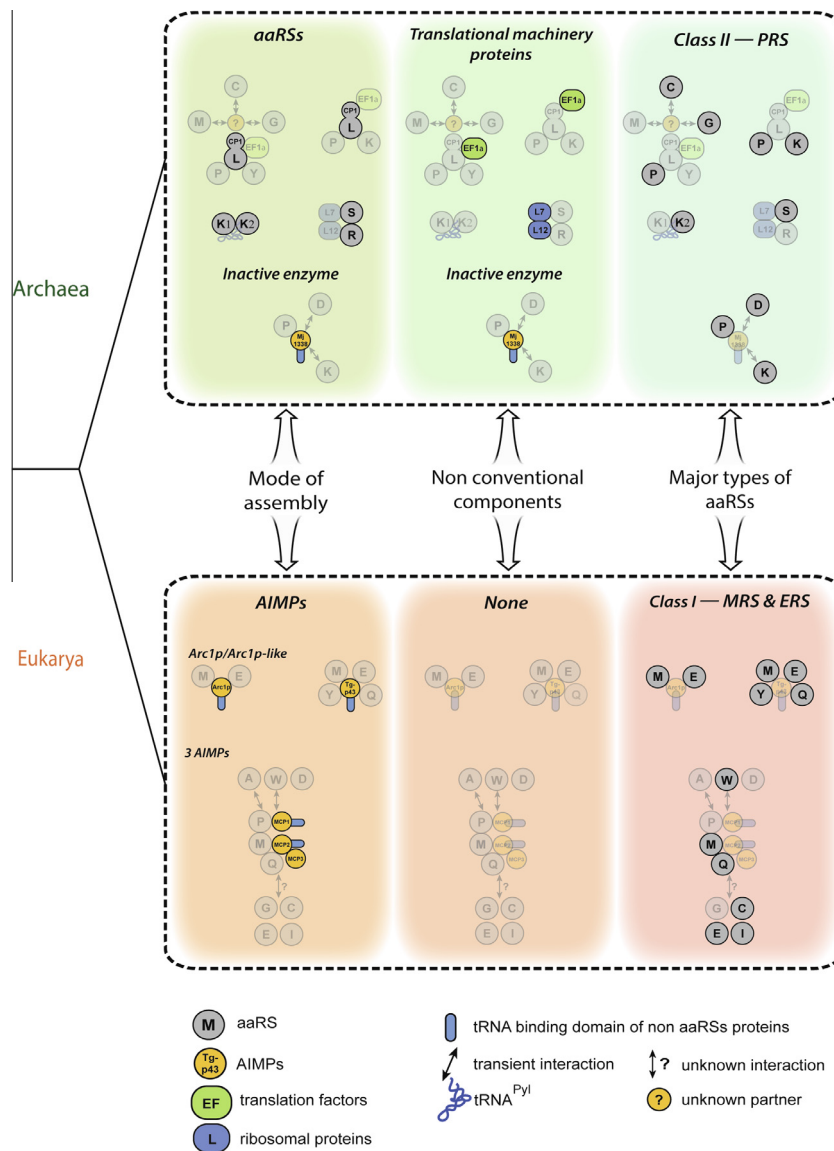


Fig. 4. Existence of two types of MSCs in low complexity organisms. Based on their assembly mode, the presence of non-conventional components and the most represented classes of aaRSs, MSCs from unicellular organisms segregate into either an archaeal- (top, green) or eukaryal-type (bottom, orange) complex. Shades (green or orange) highlight each of the 3 features (listed in the middle of the figure) that were used for the classification of the presented MSCs. Proteins mediating assembly of the MSCs, the non-conventional components and the aaRSs corresponding to most represented class of aaRSs participating to these MSCs are shown in bold, with the remaining components shaded.

S. cerevisiae and *T. brucei*, where well-defined MSC-associated domains like EMAPII, GST-like and WHEP (either in aaRSs or AIMP-Ps) are less commonly found. In archaea, the trend toward simplicity continues and aaRSs assemble into MSCs apparently without the help of any well-defined protein-binding domains. As such, the difficulty in predicting from genomic data whether or not an MSC may exist in a given organism may prove to be inversely proportional to its complexity. In the eukaryotes, the presence of WHEP domains on aaRSs, or the presence of genes encoding the EMAPII domain in conjunction with GST-like or EF-Tu-like domains are good indicators of the presence of an MSC. More robust experiments could then confirm this, as well as discover any novel or unusual components of an MSC (such as MCP3 from *T. brucei*). In prokaryotes, the scarcity of well-defined MSC-forming domains means that predicting the presence of MSCs may remain difficult, and highly dependent on classical experimental methods.

Despite this diversity, we attempted to classify MSCs from low complexity organisms by looking at three characteristic traits of MSCs: their mode of assembly and assembly factor, the presence or absence of non-conventional components; and the types of aaRSs that participate in these complexes. This analysis led us to propose that two different assembly modes and types of unicellular MSCs have emerged during evolution: the archaeal and eukaryal MSCs (Fig. 4).

The first distinctive feature between archaeal and lower eukaryote MSCs is their mode of assembly. In unicellular eukaryotes MSCs always assemble through protein–protein interaction domains that are present on AIMP-Ps. Importantly, AIMP-Ps from lower eukaryotes are structural homologs of the metazoan AIMP-Ps. By contrast, in all archaeal MSCs that have been described so far, complex formation is mediated by aaRSs, or other assembly factors, such as elongation factors or ribosomal proteins, but not by AIMP-Ps. It is still largely unknown what domains of these aaRSs, elongation factors or ribosomal proteins mediate the protein–protein interactions involved in MSC assembly. However, most of the archaeal species in which MSCs have been characterized are thermophiles; and in these organisms proteins have a tendency to assemble in complexes through the use of hydrophobic patches. This propensity to aggregate is often considered as an adaptation of these organisms to the extreme temperatures of their environment [94]. One possibility is that these hydrophobic patches could also be used to assemble archaeal MSCs.

The second idiosyncratic feature of the archaeal MSC is the presence of non-conventional components. Ribosomal proteins, translation factors (often elongation factors), or even metabolic enzymes are found in archaeal MSCs. These types of proteins are never found in eukaryal MSCs despite the fact that they are often indirectly linked to eukaryal MSCs. Indeed, a wide range of organisms has been shown to contain translatomes that gather several components of the translation machinery, including MSCs, in a huge mega-complex [67,95,96]. This association of components of the translational machinery with the MSC makes sense because it allows tRNA channeling, recycling and, by avoiding diffusion of aa-tRNA in the cell, improves the rate of mRNA codon reading.

The third major difference between archaeal and eukaryotic MSCs is a bias in aaRS composition. AaRSs segregate in two different classes, class 1 and class 2, mainly based on the fold of their aa- and ATP recognition domain [97]. In archaeal MSCs there are often more class 2 aaRSs than class 1 whereas in eukaryotes class 1 aaRSs are always the most abundant (Fig. 4). Moreover, PRS (a class 2 aaRS) is the most abundant aaRS found in archaeal MSCs (in 3 out of the 5 MSC), whereas MRS and ERS (class 1 aaRSs) are systematically present in the lower eukaryote MSCs.

The fact that PRS is the most abundant aaRSs in archaeal MSCs may be simply related to the properties of archaeal PRS. Indeed some of these enzymes, unlike their bacterial orthologues, are

deprived of editing domains and have an additional C-terminal extension that can thus be used for complex assembly [83,98,99].

The presence of MRS in MSCs from all unicellular eukaryotes may be related to its role as a sensor of oxidative stress. Indeed, it has been shown, from fungi to mammals, that MRS methionylates non-cognate tRNAs under oxidative stress. In *S. cerevisiae*, mismethionylated tRNAs can be used by translating ribosomes [60]. This mechanism is conserved in human cells [100]. tRNA mismethionylation leading to an increased content of methionine in the proteome protects the organism against oxidative through the scavenging of reactive oxygen species (ROS) [61]. As in unicellular eukaryotes MSC-bound MRS is always more efficient in charging tRNA with methionine, its recruitment in MSCs may be viewed as one way to boost the response to oxidative stress.

MSCs may also provide a link between metabolism and protein translation. Indeed, in *M. jannaschii*, aaRSs form a complex with the paralogue of a metabolic enzyme that is inactive but sensitive to H₂ concentrations [45,46], thereby linking aa-tRNA synthesis to methanogenesis. Likewise, AIMP-Ps such as the *S. cerevisiae* Arc1p and the human AIMP1/p43 seem to have a recurrent link with glucose metabolism or sensing. Indeed, Arc1p expression decreases upon glucose depletion [27], and AIMP1/p43 is a known cytokine involved in a wide range of functions including the regulation of glucose metabolism [101]. Noteworthy is the fact that the *S. cerevisiae* ERS (which also participates in the MSC) is linked to the glucose metabolism, and its mitochondrial import in respiring cell is essential [27].

To summarize, bacterial MSC has been identified only in *E. coli* and archaea and lower eukaryotes contain distinct MSCs, with differences in their modes of assembly, components and in the lineage of their aaRSs. The feeling that emerges from the studies made on these MSCs is that the major role for these archaeal and eukaryal MSCs would be to enhance aa-tRNA synthesis and to channel this molecule to the ribosome. However, without further work on the characterized MSCs from low complexity organisms and additional sampling of more species, the reason why these unicellular organisms contain MSCs will still remain mostly obscure.

Acknowledgements

The work was supported by the French National Program “Investissement d’Avenir” administered by the “Agence Nationale de la Recherche” (ANR) (“MitoCross” Laboratory of Excellence (Labex), funded as ANR-10-IDEX-0002-02), the University of Strasbourg and the CNRS. D.L. is supported by the “MitoCross” Labex (ANR-10-IDEX-0002-02) “Investissements d’Avenir” program, J.L.H. is recipient of an IDEX Unistra postdoctoral fellowship, G.B. is a recipient of a fellowship from the Ministère de l’Éducation Nationale, de la Recherche et de la Technologie and L.E. was supported by University of Strasbourg.

References

- [1] Chapeville, F., Lipmann, F., Von Ehrenstein, G., Weisblum, B., Ray Jr., W.J. and Benzer, S. (1962) On the role of soluble ribonucleic acid in coding for amino acids. *Proc. Natl. Acad. Sci. U.S.A.* 48, 1086–1092.
- [2] de Duve, C. (1988) Transfer RNAs: the second genetic code. *Nature* 333, 117–118.
- [3] Schimmel, P. and Ribas De Pouplana, L. (2000) Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem. Sci.* 25, 207–209.
- [4] Hoagland, M.B. (1955) An enzymic mechanism for amino acid activation in animal tissues. *Biochim. Biophys. Acta* 16, 288–289.
- [5] Hoagland, M.B., Keller, E.B. and Zamecnik, P.C. (1956) Enzymatic carboxyl activation of amino acids. *J. Biol. Chem.* 218, 345–358.
- [6] Ofengand, E.J., Dieckmann, M. and Berg, P. (1961) The enzymic synthesis of amino acyl derivatives of ribonucleic acid. III. Isolation of amino acid-acceptor ribonucleic acids from *Escherichia coli*. *J. Biol. Chem.* 236, 1741–1747.
- [7] Bandyopadhyay, A.K. and Deutscher, M.P. (1971) Complex of aminoacyl-transfer RNA synthetases. *J. Mol. Biol.* 60, 113–122.

- [8] Deutscher, M.P. (1974) Aminoacyl-tRNA synthetase complex from rat liver. *Methods Enzymol.* 29, 577–583.
- [9] Kern, D., Dietrich, A., Fasiolo, F., Renaud, M., Giege, R. and Ebel, J.P. (1977) The yeast aminoacyl-tRNA synthetases. Methodology for their complete or partial purification and comparison of their relative activities under various extraction conditions. *Biochimie* 59, 453–462.
- [10] Kern, D. and Lapointe, J. (1979) The twenty aminoacyl-tRNA synthetases from *Escherichia coli*. General separation procedure, and comparison of the influence of pH and divalent cations on their catalytic activities. *Biochimie* 61, 1257–1272.
- [11] Harris, C.L. (1987) An aminoacyl-tRNA synthetase complex in *Escherichia coli*. *J. Bacteriol.* 169, 2718–2723.
- [12] Yu, C.T. and Zamecnik, P.C. (1963) On the aminoacyl-RNA synthetase recognition sites of yeast and *E. Coli* transfer RNA. *Biochem. Biophys. Res. Commun.* 12, 457–463.
- [13] Schimmel, P. and Ribas de Pouplana, L. (1995) Transfer RNA: from minihelix to genetic code. *Cell* 81, 983–986.
- [14] Delarue, M. and Moras, D. (1993) The aminoacyl-transfer RNA-synthetase family – modules at work. *BioEssays* 15, 675–687.
- [15] Schimmel, P. and Ribas de Pouplana, L. (2001) Formation of two classes of tRNA synthetases in relation to editing functions and genetic code. *Cold Spring Harb. Symp. Quant. Biol.* 66, 161–166.
- [16] Becker, H.D., Roy, H., Moulinier, L., Mazauric, M.H., Keith, G. and Kern, D. (2000) *Thermus thermophilus* contains an eubacterial and an archaeobacterial aspartyl-tRNA synthetase. *Biochemistry* 39, 3216–3230.
- [17] Moulinier, L., Eiler, S., Eriani, G., Gangloff, J., Thierry, J.C., Gabriel, K., McClain, W.H. and Moras, D. (2001) The structure of an AspRS-tRNA(Asp) complex reveals a tRNA-dependent control mechanism. *EMBO J.* 20, 5290–5301.
- [18] Oshikane, H., Sheppard, K., Fukai, S., Nakamura, Y., Ishitani, R., Numata, T., Sherrer, R.L., Feng, L., Schmitt, E., Panvert, M., Blanquet, S., Mechulam, Y., Soll, D. and Nureki, O. (2006) Structural basis of RNA-dependent recruitment of glutamine to the genetic code. *Science* 312, 1950–1954.
- [19] Simader, H., Hothorn, M., Kohler, C., Basquin, J., Simos, G. and Suck, D. (2006) Structural basis of yeast aminoacyl-tRNA synthetase complex formation revealed by crystal structures of two binary sub-complexes. *Nucleic Acids Res.* 34, 3968–3979.
- [20] Simader, H., Hothorn, M. and Suck, D. (2006) Structures of the interacting domains from yeast glutamyl-tRNA synthetase and tRNA-aminoacylation and nuclear-export cofactor Arc1p reveal a novel function for an old fold. *Acta Crystallogr. D Biol. Crystallogr.* 62, 1510–1519.
- [21] Quevillon, S., Agou, F., Robinson, J.C. and Mirande, M. (1997) The p43 component of the mammalian multi-synthetase complex is likely to be the precursor of the endothelial monocyte-activating polypeptide II cytokine. *J. Biol. Chem.* 272, 32573–32579.
- [22] Cestari, I., Kalidas, S., Monnerat, S., Anupama, A., Phillips, M.A. and Stuart, K. (2013) A multiple aminoacyl-tRNA synthetase complex that enhances tRNA-aminoacylation in African trypanosomes. *Mol. Cell Biol.* 33, 4872–4888.
- [23] Godinic-Mikulcic, V., Jaric, J., Hausmann, C.D., Ibba, M. and Weygand-Durasevic, I. (2011) An archaeal tRNA-synthetase complex that enhances aminoacylation under extreme conditions. *J. Biol. Chem.* 286, 3396–3404.
- [24] Simos, G., Segref, A., Fasiolo, F., Hellmuth, K., Shevchenko, A., Mann, M. and Hurt, E.C. (1996) The yeast protein Arc1p binds to tRNA and functions as a cofactor for the methionyl- and glutamyl-tRNA synthetases. *EMBO J.* 15, 5437–5448.
- [25] Eswarappa, S.M. and Fox, P.L. (2013) Citric acid cycle and the origin of MARS. *Trends Biochem. Sci.* 38, 222–228.
- [26] Sampath, P., Mazumder, B., Seshadri, V., Gerber, C.A., Chavatte, L., Kinter, M., Ting, S.M., Dignam, J.D., Kim, S., Driscoll, D.M. and Fox, P.L. (2004) Noncanonical function of glutamyl-prolyl-tRNA synthetase: gene-specific silencing of translation. *Cell* 119, 195–208.
- [27] Frechin, M., Senger, B., Braye, M., Kern, D., Martin, R.P. and Becker, H.D. (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.
- [28] Ray, P.S., Arif, A. and Fox, P.L. (2007) Macromolecular complexes as depots for releasable regulatory proteins. *Trends Biochem. Sci.* 32, 158–164.
- [29] 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.
- [30] Hausmann, C.D. and Ibba, M. (2008) Aminoacyl-tRNA synthetase complexes: molecular multitasking revealed. *FEMS Microbiol. Rev.* 32, 705–721.
- [31] Ko, Y.G., Kang, Y.S., Kim, E.K., Park, S.G. and Kim, S. (2000) Nucleolar localization of human methionyl-tRNA synthetase and its role in ribosomal RNA synthesis. *J. Cell Biol.* 149, 567–574.
- [32] 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.
- [33] Lund, E. and Dahlberg, J.E. (1998) Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science* 282, 2082–2085.
- [34] Sarkar, S., Azad, A.K. and Hopper, A.K. (1999) Nuclear tRNA aminoacylation and its role in nuclear export of endogenous tRNAs in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14366–14371.
- [35] Hradec, J. and Dusek, Z. (1978) All factors required for protein synthesis are retained on heparin bound to Sepharose. *Biochem. J.* 172, 1–7.
- [36] Nathanson, L. and Deutscher, M.P. (2000) Active aminoacyl-tRNA synthetases are present in nuclei as a high molecular weight multienzyme complex. *J. Biol. Chem.* 275, 31559–31562.
- [37] Bailly, M., Blaise, M., Lorber, B., Becker, H.D. and Kern, D. (2007) The transamidosome: a dynamic ribonucleoprotein particle dedicated to prokaryotic tRNA-dependent asparagine biosynthesis. *Mol. Cell* 28, 228–239.
- [38] Fischer, F., Huot, J.L., Lorber, B., Diss, G., Hendrickson, T.L., Becker, H.D., Lapointe, J. and Kern, D. (2012) The asparagine-transamidosome from *Helicobacter pylori*: a dual-kinetic mode in non-discriminating aspartyl-tRNA synthetase safeguards the genetic code. *Nucleic Acids Res.* 40, 4965–4976.
- [39] Goldgur, Y. and Safro, M. (1994) Aminoacyl-tRNA synthetases from *Haloarcula marismortui*: an evidence for a multienzyme complex in a prokaryotic system. *Biochem. Mol. Biol. Int.* 32, 1075–1083.
- [40] Praetorius-Ibba, M., Rogers, T.E., Samson, R., Kelman, Z. and Ibba, M. (2005) Association between Archaeal prolyl- and leucyl-tRNA synthetases enhances tRNA(Pro) aminoacylation. *J. Biol. Chem.* 280, 26099–26104.
- [41] Hausmann, C.D., Praetorius-Ibba, M. and Ibba, M. (2007) An aminoacyl-tRNA synthetase: elongation factor complex for substrate channeling in archaeal translation. *Nucleic Acids Res.* 35, 6094–6102.
- [42] Hausmann, C.D. and Ibba, M. (2008) Structural and functional mapping of the archaeal multi-aminoacyl-tRNA synthetase complex. *FEBS Lett.* 582, 2178–2182.
- [43] Li, T., Li, Y., Guo, N., Wang, E. and Wang, Y. (1999) Discrimination of tRNA^{Leu} isoacceptors by the insertion mutant of *Escherichia coli* leucyl-tRNA synthetase. *Biochemistry* 38, 9084–9088.
- [44] Godinic-Mikulcic, V., Jaric, J., Greber, B.J., Franke, V., Hodnik, V., Anderlüh, G., Ban, N. and Weygand-Durasevic, I. (2014) Archaeal aminoacyl-tRNA synthetases interact with the ribosome to recycle tRNAs. *Nucleic Acids Res.*
- [45] Lipman, R.S., Chen, J., Evilia, C., Vitseva, O. and Hou, Y.M. (2003) Association of an aminoacyl-tRNA synthetase with a putative metabolic protein in archaea. *Biochemistry* 42, 7487–7496.
- [46] Afting, C., Kremmer, E., Brucker, C., Hochheimer, A. and Thauer, R.K. (2000) Regulation of the tRNA^{H2}-forming methylenetetrahydromethanopterin dehydrogenase (Hmd) and HmdIII in *Methanothermobacter marburgensis*. *Arch. Microbiol.* 174, 225–232.
- [47] Oza, J.P., Sowers, K.R. and Perona, J.J. (2012) Linking energy production and protein synthesis in hydrogenotrophic methanogens. *Biochemistry* 51, 2378–2389.
- [48] Raina, M., Elgamal, S., Santangelo, T.J. and Ibba, M. (2012) Association of a multi-synthetase complex with translating ribosomes in the archaeon *Thermococcus kodakarensis*. *FEBS Lett.* 586, 2232–2238.
- [49] James, C.M., Ferguson, T.K., Leykam, J.F. and Krzycki, J.A. (2001) The amber codon in the gene encoding the monomethylamine methyltransferase isolated from *Methanosarcina barkeri* is translated as a sense codon. *J. Biol. Chem.* 276, 34252–34258.
- [50] Blight, S.K., Larue, R.C., Mahapatra, A., Longstaff, D.G., Chang, E., Zhao, G., Kang, P.T., Green-Church, K.B., Chan, M.K. and Krzycki, J.A. (2004) Direct charging of tRNA^{CUA} with pyrrolysine in vitro and in vivo. *Nature* 431, 333–335.
- [51] Polycarpo, C., Ambrogelly, A., Ruan, B., Tumbula-Hansen, D., Ataide, S.F., Ishitani, R., Yokoyama, S., Nureki, O., Ibba, M. and Söll, D. (2003) Activation of the pyrrolysine suppressor tRNA requires formation of a ternary complex with class I and class II lysyl-tRNA synthetases. *Mol. Cell* 12, 287–294.
- [52] Kaminska, M., Havrylenko, S., Decottignies, P., Gillet, S., Le Marechal, P., Negrutskii, B. and Mirande, M. (2009) Dissection of the structural organization of the aminoacyl-tRNA synthetase complex. *J. Biol. Chem.* 284, 6053–6060.
- [53] Karanasios, E. and Simos, G. (2010) Building arks for tRNA: structure and function of the Arc1p family of non-catalytic tRNA-binding proteins. *FEBS Lett.* 584, 3842–3849.
- [54] Frechin, M., Kern, D., Martin, R.P., Becker, H.D. and Senger, B. (2010) Arc1p: anchoring, routing, coordinating. *FEBS Lett.* 584, 427–433.
- [55] Galani, K., Hurt, E. and Simos, G. (2005) The tRNA aminoacylation co-factor Arc1p is excluded from the nucleus by an Xpo1p-dependent mechanism. *FEBS Lett.* 579, 969–975.
- [56] Galani, K., Grosshans, H., Deinert, K., Hurt, E.C. and Simos, G. (2001) The intracellular location of two aminoacyl-tRNA synthetases depends on complex formation with Arc1p. *EMBO J.* 20, 6889–6898.
- [57] Graindorge, J.S., Senger, B., Tritsch, D., Simos, G. and Fasiolo, F. (2005) Role of Arc1p in the modulation of yeast glutamyl-tRNA synthetase activity. *Biochemistry* 44, 1344–1352.
- [58] Simos, G., Sauer, A., Fasiolo, F. and Hurt, E.C. (1998) A conserved domain within Arc1p delivers tRNA to aminoacyl-tRNA synthetases. *Mol. Cell* 1, 235–242.
- [59] Golinelli-Cohen, M.P. and Mirande, M. (2007) Arc1p is required for cytoplasmic confinement of synthetases and tRNA. *Mol. Cell. Biochem.* 300, 47–59.
- [60] Wilttrout, E., Goodenbour, J.M., Frechin, M. and Pan, T. (2012) Misacylation of tRNA with methionine in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 40, 10494–10506.
- [61] Luo, S. and Levine, R.L. (2009) Methionine in proteins defends against oxidative stress. *FASEB J.* 23, 464–472.
- [62] Vogt, W. (1995) Oxidation of methionyl residues in proteins: tools, targets, and reversal. *Free Radical Biol. Med.* 18, 93–105.
- [63] Sha, Z., Brill, L.M., Cabrera, R., Kleifeld, O., Scheliga, J.S., Glickman, M.H., Chang, E.C. and Wolf, D.A. (2009) The eIF3 interactome reveals the

- translasome, a supercomplex linking protein synthesis and degradation machineries. *Mol. Cell* 36, 141–152.
- [64] An, S. and Musier-Forsyth, K. (2005) Cys-tRNA(Pro) editing by *Haemophilus influenzae* YbaK via a novel synthetase-YbaK-tRNA ternary complex. *J. Biol. Chem.* 280, 34465–34472.
- [65] Ruan, B. and Soll, D. (2005) The bacterial YbaK protein is a Cys-tRNAPro and Cys-tRNA Cys deacylase. *J. Biol. Chem.* 280, 25887–25891.
- [66] van Rooyen, J.M., Murat, J.B., Hammoudi, P.M., Kieffer-Jaquinod, S., Coute, Y., Sharma, A., Pelloux, H., Belrhali, H. and Hakimi, M.A. (2014) Assembly of the novel five-component apicomplexan multi-aminoacyl-tRNA synthetase complex is driven by the hybrid scaffold protein Tg-p43. *PLoS One* 9, e89487.
- [67] Havrylenko, S., Legouis, R., Negrutskii, B. and Mirande, M. (2011) *Caenorhabditis elegans* evolves a new architecture for the multi-aminoacyl-tRNA synthetase complex. *J. Biol. Chem.* 286, 28476–28487.
- [68] Havrylenko, S., Legouis, R., Negrutskii, B. and Mirande, M. (2010) Methionyl-tRNA synthetase from *Caenorhabditis elegans*: a specific multidomain organization for convergent functional evolution. *Protein Sci.* 19, 2475–2484.
- [69] Park, S.G., Ewalt, K.L. and Kim, S. (2005) Functional expansion of aminoacyl-tRNA synthetases and their interacting factors: new perspectives on housekeepers. *Trends Biochem. Sci.* 30, 569–574.
- [70] 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.
- [71] Mirande, M., Kellermann, O. and Waller, J.P. (1982) Macromolecular complexes from sheep and rabbit containing seven aminoacyl-tRNA synthetases. II. Structural characterization of the polypeptide components and immunological identification of the methionyl-tRNA synthetase subunit. *J. Biol. Chem.* 257, 11049–11055.
- [72] Guo, M. and Yang, X.L. (2014) Architecture and metamorphosis. *Top. Curr. Chem.* 344, 89–118.
- [73] Eichinger, L., Pachebat, J.A., Glockner, G., Rajandream, M.A., Sugang, R., Berriman, M., Song, J., Olsen, R., Szafarski, K., Xu, Q., Tunggal, B., Kummerfeld, S., Madera, M., Konfortov, B.A., Rivero, F., Bankier, A.T., Lehmann, R., Hamlin, N., Davies, R., Gaudet, P., Fey, P., Pilcher, K., Chen, G., Saunders, D., Sodergren, E., Davis, P., Kerhornou, A., Nie, X., Hall, N., Anjard, C., Hemphill, L., Bason, N., Farbrother, P., Desany, B., Just, E., Morio, T., Rost, R., Churcher, C., Cooper, J., Haydock, S., van Driessche, N., Cronin, A., Goodhead, I., Muzny, D., Mourier, T., Pain, A., Lu, M., Harper, D., Lindsay, R., Hauser, H., James, K., Quiles, M., Madan Babu, M., Saito, T., Buchrieser, C., Wardroper, A., Felder, M., Thangavelu, M., Johnson, D., Knights, A., Loulseged, H., Mungall, K., Oliver, K., Price, C., Quail, M.A., Urushihara, H., Hernandez, J., Rabinowitsch, E., Steffen, D., Sanders, M., Ma, J., Kohara, Y., Sharp, S., Simmonds, M., Spiegler, S., Tivey, A., Sugano, S., White, B., Walker, D., Woodward, J., Winckler, T., Tanaka, Y., Shaulsky, G., Schleicher, M., Weinstock, G., Rosenthal, A., Cox, E.C., Chisholm, R.L., Gibbs, R., Loomis, W.F., Platzer, M., Kay, R.R., Williams, J., Dear, P.H., Noegel, A.A., Barrell, B. and Kuspa, A. (2005) The genome of the social amoeba *Dictyostelium discoideum*. *Nature* 435, 43–57.
- [74] Kim, J.E., Kim, K.H., Lee, S.W., Seol, W., Shiba, K. and Kim, S. (2000) An elongation factor-associating domain is inserted into human cysteinyl-tRNA synthetase by alternative splicing. *Nucleic Acids Res.* 28, 2866–2872.
- [75] Cahuzac, B., Berthonneau, E., Birlirakis, N., Guittet, E. and Mirande, M. (2000) A recurrent RNA-binding domain is appended to eukaryotic aminoacyl-tRNA synthetases. *EMBO J.* 19, 445–452.
- [76] Jeong, E.J., Hwang, G.S., Kim, K.H., Kim, M.J., Kim, S. and Kim, K.S. (2000) Structural analysis of multifunctional peptide motifs in human bifunctional tRNA synthetase: identification of RNA-binding residues and functional implications for tandem repeats. *Biochemistry* 39, 15775–15782.
- [77] Rho, S.B., Lee, J.S., Jeong, E.J., Kim, K.S., Kim, Y.G. and Kim, S. (1998) A multifunctional repeated motif is present in human bifunctional tRNA synthetase. *J. Biol. Chem.* 273, 11267–11273.
- [78] Jia, J., Arif, A., Ray, P.S. and Fox, P.L. (2008) WHEP domains direct noncanonical function of glutamyl-Prolyl tRNA synthetase in translational control of gene expression. *Mol. Cell* 29, 679–690.
- [79] Lee, S.W., Cho, B.H., Park, S.G. and Kim, S. (2004) Aminoacyl-tRNA synthetase complexes: beyond translation. *J. Cell Sci.* 117, 3725–3734.
- [80] Merritt, E.A., Arakaki, T.L., Gillespie, J.R., Larson, E.T., Kelley, A., Mueller, N., Napuli, A.J., Kim, J., Zhang, L., Verlinde, C.L., Fan, E., Zuckner, F., Buckner, F.S., van Voorhis, W.C. and Hol, W.G. (2010) Crystal structures of trypanosomal histidyl-tRNA synthetase illuminate differences between eukaryotic and prokaryotic homologs. *J. Mol. Biol.* 397, 481–494.
- [81] Shiba, K. (2002) Intron positions delineate the evolutionary path of a pervasively appended peptide in five human aminoacyl-tRNA synthetases. *J. Mol. Evol.* 55, 727–733.
- [82] Praetorius-Ibba, M., Hausmann, C.D., Paras, M., Rogers, T.E. and Ibba, M. (2007) Functional association between three archaeal aminoacyl-tRNA synthetases. *J. Biol. Chem.* 282, 3680–3687.
- [83] Ambrogelly, A., Kamtekar, S., Stathopoulos, C., Kennedy, D. and Soll, D. (2005) Asymmetric behavior of archaeal prolyl-tRNA synthetase. *FEBS Lett.* 579, 6017–6022.
- [84] Ling, C., Yao, Y.N., Zheng, Y.G., Wei, H., Wang, L., Wu, X.F. and Wang, E.D. (2005) The C-terminal appended domain of human cytosolic leucyl-tRNA synthetase is indispensable in its interaction with arginyl-tRNA synthetase in the multi-tRNA synthetase complex. *J. Biol. Chem.* 280, 34755–34763.
- [85] Kao, J., Ryan, J., Brett, G., Chen, J., Shen, H., Fan, Y.G., Godman, G., Familletti, P.C., Wang, F. and Pan, Y.C. (1992) Endothelial monocyte-activating polypeptide II. A novel tumor-derived polypeptide that activates host-response mechanisms. *J. Biol. Chem.* 267, 20239–20247.
- [86] Crepin, T., Schmitt, E., Blanquet, S. and Mechulam, Y. (2002) Structure and function of the C-terminal domain of methionyl-tRNA synthetase. *Biochemistry* 41, 13003–13011.
- [87] Kawaguchi, S., Muller, J., Linde, D., Kuramitsu, S., Shibata, T., Inoue, Y., Vassilyev, D.G. and Yokoyama, S. (2001) The crystal structure of the tCsaA protein: an export-related chaperone from *Thermus thermophilus*. *EMBO J.* 20, 562–569.
- [88] Swairjo, M.A., Morales, A.J., Wang, C.C., Ortiz, A.R. and Schimmel, P. (2000) Crystal structure of trbp111: a structure-specific tRNA-binding protein. *EMBO J.* 19, 6287–6298.
- [89] Renault, L., Kerjan, P., Pasqualato, S., Menetrey, J., Robinson, J.C., Kawaguchi, S., Vassilyev, D.G., Yokoyama, S., Mirande, M. and Cherfils, J. (2001) Structure of the EMAPII domain of human aminoacyl-tRNA synthetase complex reveals evolutionary dimer mimicry. *EMBO J.* 20, 570–578.
- [90] Karanasios, E., Simader, H., Panayotou, G., Suck, D. and Simos, G. (2007) Molecular determinants of the yeast Arc1p-aminoacyl-tRNA synthetase complex assembly. *J. Mol. Biol.* 374, 1077–1090.
- [91] Hellmuth, K., Lau, D.M., Bischoff, F.R., Kunzler, M., Hurt, E. and Simos, G. (1998) Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. *Mol. Cell Biol.* 18, 6374–6386.
- [92] Wakasugi, K. and Schimmel, P. (1999) Two distinct cytokines released from a human aminoacyl-tRNA synthetase. *Science* 284, 147–151.
- [93] Castro de Moura, M., Miro, F., Han, J.M., Kim, S., Celada, A. and Ribas de Pouplana, L. (2011) Entamoeba lysyl-tRNA synthetase contains a cytokine-like domain with chemokine activity towards human endothelial cells. *PLoS Negl. Trop Dis.* 5, e1398.
- [94] Vieille, C. and Zeikus, G.J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* 65, 1–43.
- [95] Reading, B.J., Williams, V.N., Chapman, R.W., Williams, T.I. and Sullivan, C.V. (2013) Dynamics of the striped bass (*Morone saxatilis*) ovary proteome reveal a complex network of the translasome. *J. Proteome Res.* 12, 1691–1699.
- [96] Kaminska, M., Havrylenko, S., Decottignies, P., Le Marechal, P., Negrutskii, B. and Mirande, M. (2009) Dynamic organization of aminoacyl-tRNA synthetase complexes in the cytoplasm of human cells. *J. Biol. Chem.* 284, 13746–13754.
- [97] 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.
- [98] Beuning, P.J. and Musier-Forsyth, K. (2001) Species-specific differences in amino acid editing by class II prolyl-tRNA synthetase. *J. Biol. Chem.* 276, 30779–30785.
- [99] Beuning, P.J. and Musier-Forsyth, K. (2000) Hydrolytic editing by a class II aminoacyl-tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8916–8920.
- [100] Netzer, N., Goodenbour, J.M., David, A., Dittmar, K.A., Jones, R.B., Schneider, J.R., Boone, D., Eves, E.M., Rosner, M.R., Gibbs, J.S., Embry, A., Dolan, B., Das, S., Hickman, H.D., Berglund, P., Bennink, J.R., Yewdell, J.W. and Pan, T. (2009) Innate immune and chemically triggered oxidative stress modifies translational fidelity. *Nature* 462, 522–526.
- [101] Park, S.G., Kang, Y.S., Kim, J.Y., Lee, C.S., Ko, Y.G., Lee, W.J., Lee, K.U., Yeom, Y.I. and Kim, S. (2006) Hormonal activity of AIM1/p43 for glucose homeostasis. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14913–14918.

Annexe 2



Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Nonconventional localizations of cytosolic aminoacyl-tRNA synthetases in yeast and human cells

Sylvain Debard^{a,1}, Gaétan Bader^{a,1}, Johan-Owen De Craene^{a,2}, Ludovic Enkler^b, Séverine Bär^a, Daphné Laporte^a, Philippe Hammann^c, Evelyne Myslinski^a, Bruno Senger^a, Sylvie Friant^a, Hubert Dominique Becker^{a,*}

^a Université de Strasbourg, CNRS, GMGM UMR 7156, F-67000 Strasbourg, France

^b IBMC-CNRS, Evolution des ARN non codants chez la levure, Architecture et Réactivité de l'ARN, 15 rue René Descartes, Université de Strasbourg, Strasbourg, France

^c Plateforme Protéomique Strasbourg-Esplanade, Institut de Biologie Moléculaire et Cellulaire, FRC 1589, Centre National de la Recherche Scientifique, Université de Strasbourg, 67084 Strasbourg, France

ARTICLE INFO

Article history:

Received 31 July 2016

Received in revised form 27 September 2016

Accepted 30 September 2016

Available online xxx

Keywords:

aaRS
tRNA
Yeast
Human
Microscopy
Fractionation
MTS
NLS

ABSTRACT

By definition, cytosolic aminoacyl-tRNA synthetases (aaRSs) should be restricted to the cytosol of eukaryotic cells where they supply translating ribosomes with their aminoacyl-tRNA substrates. However, it has been shown that other translationally-active compartments like mitochondria and plastids can simultaneously contain the cytosolic aaRS and its corresponding organellar ortholog suggesting that both forms do not share the same organellar function. In addition, a fair number of cytosolic aaRSs have also been found in the nucleus of cells from several species. Hence, these supposedly cytosolic-restricted enzymes have instead the potential to be multi-localized. As expected, in all examples that were studied so far, when the cytosolic aaRS is imported inside an organelle that already contains its *bona fide* corresponding organellar-restricted aaRSs, the cytosolic form was proven to exert a nonconventional and essential function. Some of these essential functions include regulating homeostasis and protecting against various stresses. It thus becomes critical to assess meticulously the subcellular localization of each of these cytosolic aaRSs to unravel their additional roles. With this objective in mind, we provide here a review on what is currently known about cytosolic aaRSs multi-compartmentalization and we describe all commonly used protocols and procedures for identifying the compartments in which cytosolic aaRSs relocate in yeast and human cells.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Contents

1. Introduction	00
2. Antibodies: A fundamental tool for aaRSs localization	00
3. Prediction and analysis of aaRSs with organellar import signal	00
3.1. Mitochondrial targeting sequence prediction	00
3.2. Nuclear localization signal prediction	00
4. Subcellular fractionation	00
4.1. Crude membranes fractionation from yeast cells	00
4.2. Purification of nuclei and preparation of nuclear protein extracts	00
4.2.1. Purification of nuclei from yeast cells	00
4.2.2. Purification of nuclei from human cells	00
4.3. Purification of mitochondria and preparation of mitochondrial protein extracts	00

* Corresponding author.

E-mail address: h.becker@unistra.fr (H.D. Becker).

¹ These authors contributed equally to this work.

² Present address: EA 2106, Biomolécules et Biotechnologies Végétales, Université François Rabelais de Tours, UFR Sciences et Techniques, Parc de Grandmont, 37200 Tours, France.

<http://dx.doi.org/10.1016/j.ymeth.2016.09.017>

1046-2023/© 2016 The Authors. Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

4.3.1.	Purification of mitochondria from yeast cells	00
4.3.2.	Obtention of mitoplasts from yeast mitochondria	00
4.3.3.	Purification of mitochondria from human cells	00
4.3.4.	Obtention of mitoplasts from human mitochondria	00
5.	Mass spectrometry analysis and identification of cytosolic aaRSs in yeast mitochondrial extracts	00
5.1.	Protein digestion solution	00
5.2.	Nano-liquid Chromatography – electrospray Ionization TripleTOF MS/MS Analysis	00
5.3.	Database search and data analysis	00
6.	Microscopy analysis and single cell localization	00
6.1.	Immunofluorescence on yeast cells	00
6.1.1.	Fixation steps	00
6.1.2.	Incubation with antibodies	00
6.2.	Immunofluorescence on human cells	00
7.	Concluding remarks	00
	Acknowledgements	00
	Appendix A. Supplementary data	00
	References	00

1. Introduction

Aminoacyl-tRNA synthetases (aaRSs) constitute a family of ubiquitous enzymes present in the three kingdoms of life and essentially known for transfer RNA (tRNA) aminoacylation. Indeed, they are required for the ligation of the 20 standard proteinogenic amino acids (aa) to their cognate tRNAs [1]. tRNA aminoacylation and protein synthesis occur in the cytosol of all organisms but unlike bacteria, eukaryotes have membrane-enclosed structures compartmentalizing their cytosol. Among these compartments, mitochondria in all eukaryotes and also chloroplasts in plants synthesize proteins by translating organellar mRNAs [2]. Hence, a compartmentalized set of the 20 aaRSs is expected in each translationally-active compartment. Since no gene encoding mitochondrial aaRSs was found in any of the mitochondrial DNA sequenced so far, the nuclear genome has to encode both full sets of cytosolic and organellar aaRSs. However, while this is true for the set of 20 cytosolic aaRSs, a full set of 20 additional separate genes encoding organellar aaRSs has never been found in any of the nuclear genomes sequenced so far [3–5]. Both sets of genes can usually be easily distinguished based on phylogenetic analyses. These studies show that a majority of the mitochondrial aaRSs are of bacterial descent but that the α -proteobacterial endosymbiotic origin of the mitochondrial aaRSs has largely been lost because of gene replacements, intranuclear gene duplications and divergence or horizontal gene transfers [6,7].

In plants, the evolution of the different aaRSs seems more complicated. Monocot and dicot have either at least 1 gene encoding a given aaRS for each protein-synthesizing compartment or 2 genes with one gene encoding the cytosolic and the second both the mitochondrial and the chloroplastic aaRS [8]. Surprisingly, both have one instance where only one gene seems to encode all 3 activities: isoleucyl-tRNA synthetase (IleRS) for monocot and glutamyl-tRNA synthetase (GlnRS) for dicot [5,8,9]. The same is true for diatoms and brown algae in which a single gene seems to encode all 3 arginyl-tRNA synthetases (ArgRSs) [5]. In many instances, the mechanism that allows expression of the organellar aaRS from a gene that also encodes the cytosolic isoform is unknown or ill defined [10,11]. For the sake of brevity, from now on we will use the three letters code for each amino acid mentioned.

Genes encoding several mitochondrial aaRSs are also missing in metazoan and fungal genomes and various strategies are used by these organisms to produce the mitochondrial aaRS from an apparently missing gene. Many of them are encoded by the same exact gene that encodes the cytosolic form, except that the mRNA that

will be translated into the mitochondrial isoform contains an additional 5'-extension for the mitochondrial targeting sequence (MTS). The synthesis of these 2 proteins from the same gene is achieved through different mechanisms: (i) alternative transcription start as shown for the yeast ValRS and HisRS mRNAs [12,13]; (ii) alternative splicing as shown for example for the human LysRS [14]; (iii) alternative translation initiation at two different AUG start codons within the same mRNA such as for the human GlyRS [15]; or (iv) alternative transcription start at a non-canonical start codon such as for the yeast AlaRS [16]. The complete list of cytosolic and organellar aaRSs generated from a single gene is presented in Table 1.

Eukaryotes, with the exception of a few parasites that possess a mitochondrial GlnRS encoded by the same gene as the cytosolic GlnRS, use a different strategy to compensate for the ubiquitous absence of the gene encoding mitochondrial GlnRS [41]. This widespread absence of a mitochondrial GlnRS is an illustration of the evolutionary origin of mitochondria that very likely originated from an α -proteobacterial endosymbiont that was generating glutamyl-tRNA^{Gln} (Gln-tRNA^{Gln}) not by direct charging of tRNA^{Gln} by a GlnRS but by a tRNA-dependent transamidation pathway, metazoans and fungi kept the endosymbiont's route for mitochondrial Gln-tRNA^{Gln} formation [42]. This tRNA-dependent transamidation pathway requires first glutamylation of mitochondrially-encoded tRNA^{Gln} (mttRNA^{Gln}) by a mitochondrial non-discriminating GluRS and subsequent transamidation of the charged glutamate into glutamine by a mitochondrial tRNA-dependent amidotransferase (AdT) [9,28,43,44]. However, the two enzymes that sustain this pathway seem to have species-specific features. In human cells, it is the mitochondrial GluRS that charges both the mttRNA^{Glu} and the mttRNA^{Gln} [44], whereas in the yeast *Saccharomyces cerevisiae* a pool of the cytosolic GluRS is imported into the mitochondria to carry out glutamylation of mttRNA^{Gln} [28]. Even if metazoans and fungi both belong to the opisthokonta group, their trimeric mitochondrial AdTs differ by one subunit, the classical bacterial-like GatC subunit found in metazoan AdTs is replaced by a fungi-specific GatF subunit [45].

In yeast, there is only one CysRS gene (Supp. Table 1) suggesting that it encodes both the cytosolic and mitochondrial isoforms. However, the mitochondrial CysRS has yet not been experimentally characterized. Along the same line, *S. cerevisiae* genome has two genes, *AIM10* and *YHRO20W* that encode for proteins showing similarities to prokaryotic (*AIM10*) and eukaryotic (*YHRO20W*) ProRSs (Supp. Table 1). While there are no experimental proofs that *AIM10* encodes a protein with ProRS activity, its deletion leads to a respiratory-deficient yeast cell showing that it required for

Table 1
List of dual-localized cytosolic aaRSs identified so far.

Organism	Localisation	aaRS	Techniques used	Mechanism	References
<i>Hs</i>	Mitochondrial	GlyRS	IF, CF	Alternative transcription start	[15,17]
		LysRS	FM	Alternative mRNA splicing	[14]
	Nuclear	LysRS	CM, CF, CE	Phosphorylation	[18,19]
		MetRS	FM, CE, IF, CF	Cell growth signal induced	[19,20]
		PheRS	IF	n.d.	[21]
		TrpRS	CF, IF, IEM	IFN gamma induced	[22–24]
		TyrRS	IF, CM, CF	tRNATyr binding	[25]
		TyrRS	CF, IF	Oxidative stress induced	[26]
		AlaRS	G, CF	Non-AUG alternative translation start	[16,27]
<i>Sc</i>	Mitochondrial	GluRS	FM, G, CF, CM	Arc1p released	[28,29]
		GlyRS1	G	Non-AUG alternative translation start	[30,31]
		HisRS	G, CF	Alternative transcription start	[12,32]
		ValRS	G	Alternative translation start	[13,33]
		MetRS	FM, CF, CM	Arc1p released	[29,34]
	Nuclear	TyrRS	FM, CF	NLS characterisation	[25,35]
		IleRS	IF, CF	Trans-splicing	[36,37]
		GluRS	IF, CF	Trans-splicing	[37,38]
		GlnRS	IF, CF	Trans-splicing	[37,38]
<i>Tb</i>	Mitochondrial	ProRS	IF	Trans-splicing	[37]
		AlaRS	IF	n.d.	[39]
		CysRS	FM, IF	Alternative mRNA splicing	[40]
		GlyRS	IF	n.d.	[39]
<i>Pf</i>	Apicoplast	ThrRS	IF	n.d.	[39]

aaRSs in bold are enzymes that relocalize *stricto sensu* from cytosol to the organellar compartment. *Hs*, *Homo sapiens*; *Sc*, *Saccharomyces cerevisiae*; *Tb*, *Trypanosoma brucei*; *Pf*, *Plasmodium falciparum*; CM, confocal microscopy; IF, immunofluorescence; CF, cell fractionation; FM, epifluorescence microscopy; CE, capillary electrophoresis; IEM, immuno-electron microscopy; G, genetic; NLS, nuclear localization signal; n.d., not determined.

mitochondria function [46]. The protein encoded by the *YHRO20W* gene has been shown to retain ProRS activity, *in vitro*, and to be essential for cell survival, suggesting that it might be the cytosolic

ProRS [47]. The mitoproteome done on purified *S. cerevisiae* mitochondria (performed as described in Section 5; Fig. 1A) shows that cytosolic CysRS is found in the mitoproteome, suggesting that the

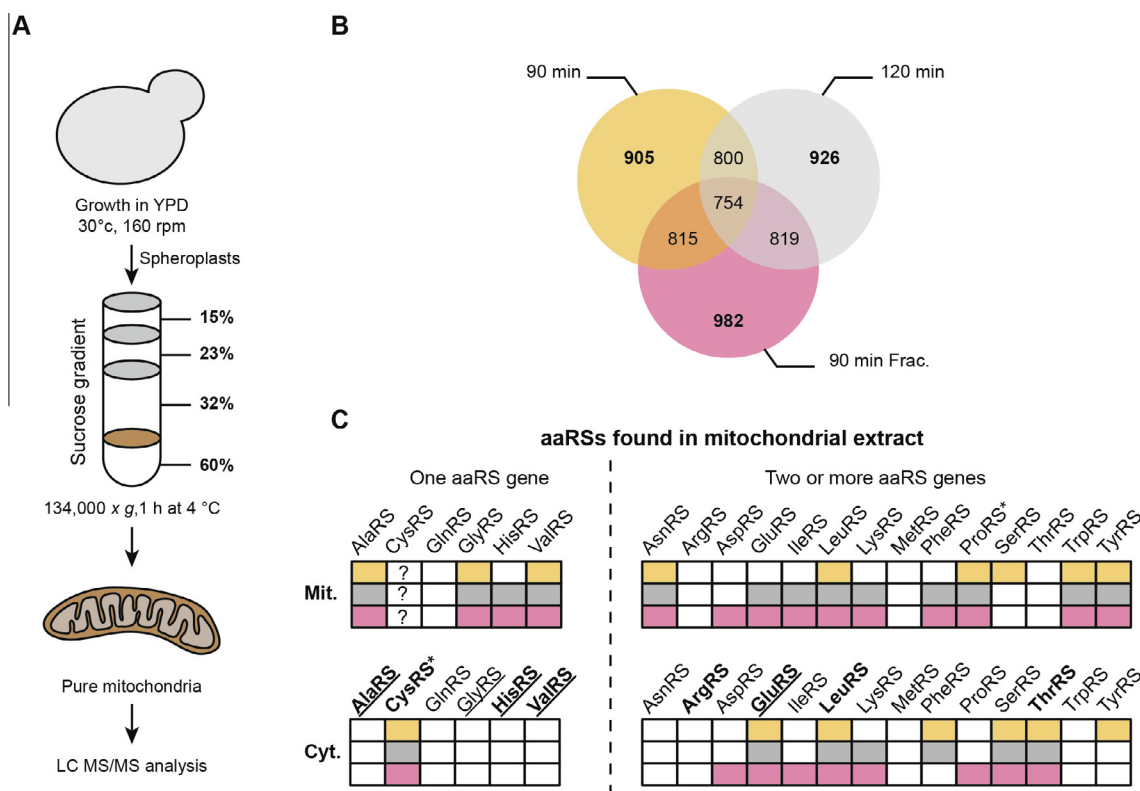


Fig. 1. Mitochondrial localization of cytosolic and mitochondrial aaRSs. (A) Schematized representation of yeast mitochondria preparation. The procedure we followed was that described by Meisinger and coworkers [48]. Mitochondria are recovered from the 60%–32% interphase. (B) LC MS/MS analysis of the mitoproteome. After separation, pure yeast mitochondria were subjected to LC MS/MS analysis with 3 different injection gradients (90 min, 120 min and 90 min Frac, see Section 5 for details). Number of total proteins identified in each samples are represented in bold. Proteins found shared between conditions are also showed. (C) Mitochondrial (top) and cytosolic (bottom) forms of aaRSs identified in mitochondrial extract by LC MS/MS analysis. Colors are corresponding to samples in panel B. Cytosolic aaRSs predicted to have an MTS by at least 1 predictor (see Table 5) are in bold. Yeast cytosolic aaRSs previously shown experimentally as mitochondria dual-localized are underlined (see Table 1). No mitochondrial specific CysRS identify in yeast. * indicates putative aaRSs. See Supplemental Table 1 for yeast aaRSs genes name.

missing mitochondrial CysRS could be compensated by mitochondrial import of the cytosolic ortholog (Fig. 1C). The *AIM10* gene product was found in the different mass-spectrometry analyses (Fig. 1C and Supp. Table 2), confirming its mitochondrial localization, but further experimental proofs that this protein is a functional ProRS are still needed. Interestingly, the *YHR020W* gene product was also found in one of the mass-spectrometry data set we obtained (Fig. 1C and Supp. Table 2), suggesting that the cytosolic ProRS might be dual-localized. In addition, depending on the sample injection time used prior to mass spectrometry, 6–8 additional other cytosolic aaRSs were found in the mitoproteome. Interestingly, each of these mitochondria-imported cytosolic aaRSs has a mitochondrial ortholog encoded by a separate gene. Nonetheless, the deletion of any of these genes encoding the *bona fide* mitochondrial aaRS leads to a respiratory deficiency (SGD website source). This means that the cytosolic aaRS orthologs, imported into mitochondria, cannot compensate for the loss of the mitochondrial ones and thus suggests that they have other roles than cognate $_{mt}$ tRNA aminoacylation.

Relocalization of cytosolic aaRSs is not restricted to compartments in which their presence is required to supply protein synthesis with aa-tRNAs. Indeed, some cytosolic aaRSs have been localized in the nucleus both in human and *S. cerevisiae* (Table 1). These and other proteins of the translation machinery were first observed while analyzing the protein content of *Xenopus* oocytes and has since also been observed in mammalian and yeast cells [20,29,49–51]. Their proposed function in the nucleus is that by charging only the mature tRNAs, these aaRSs are in charge of a quality-control step of the tRNA maturation process prior to their export to the cytoplasm [49,50]. More recent reports show other functions for these nuclear pools of human and yeast aaRSs that do not require their tRNA-charging capacities. For example, under oxidative stress a portion of the human TyrRS relocalizes to the nucleus where it activates E2F1, a transcription factor that up-regulates the expression of DNA damage repair genes [26]. Likewise, a pool of the human MetRS relocalizes into the nucleoli of

proliferating cells to regulate rRNA biogenesis [20]. In the yeast *S. cerevisiae*, a pool of MetRS relocalizes into the nucleus to regulate transcription of genes encoding subunits of respiratory complexes [29]. More broadly, a previous bioinformatics analysis identified 16 yeast cytosolic aaRSs as harboring a putative nuclear localization signal (NLS) suggesting that these enzymes could enter the nucleus to ensure additional functions other than translation [52]. However the MetRS, experimentally shown to relocalize into the nucleus was not identified by this bioinformatics study. This highlights the limits of the bioinformatics prediction algorithms to identify targeting sequences, hence the difficulties to predict in which additional compartments these aaRSs could eventually relocalize. However, the scientific community that studies these enzymes has come to the consensus that any cytosolic aaRSs could potentially be multi-localized.

Defining what we mean by multi-localized aaRSs is not an easy task given the variety of strategies that organisms use to relocalize the same protein in several compartments. In the present review, we will define a multi-localized aaRS as a cytosolic aaRS that is found in at least 2 different compartments and for which the organellar isoform is produced from the same gene as the cytosolic one through the various strategies that have been detailed in the previous paragraphs (see also Table 1).

The mechanisms regulating these multi-localizations have been mostly studied for aaRSs that belong to multi-synthetase complexes (MSCs) found in all eukaryotic species studied so far. They are composed of 2–9 cytosolic aaRSs interacting with 1–3 cytosolic anchors called aminoacyl-tRNA synthetase-interacting multifunctional proteins (AIMPs) [53–59]. It has been proposed and experimentally verified for some of the MSC-interacting cytosolic aaRSs that these MSCs are reservoirs for releasable aaRSs that can relocalize to other subcellular compartments to exert nonconventional functions [29,60,61].

In this review, we gather the various methods and procedures that have been used to address the localization of individual cytosolic aaRSs (Fig. 2). Some are biochemical (organelle purification),

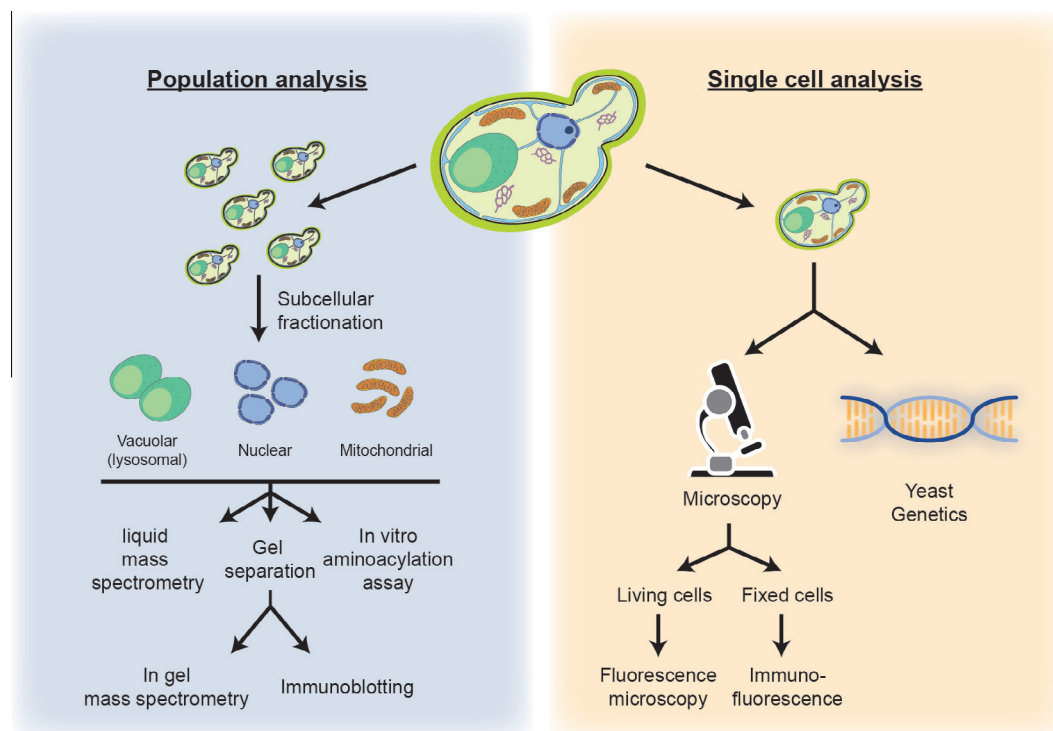


Fig. 2. Different strategies commonly used for localizing aaRSs in yeast cells. All techniques can be used on mammalian cells except those specific to yeast genetics.

genetic (isolation of mutations preventing mitochondrial import, not described in this review) and others are from cell biology (immunofluorescence on fixed cells or GFP-tagged proteins on living cells). The localization of each aaRS in various organisms together with the techniques that were used to localize them are described in Table 1. We also mention the predicted organellar import signals (MTS and NLS) that we found in yeast and human aaRSs. All the methods described in this review have their advantages and drawbacks, and the purpose of this review is to help researchers defining which techniques or procedures should be used to localize a given cytosolic aaRSs. We have focused mainly on techniques used on yeast and human cells although we mention work done on other organisms.

2. Antibodies: A fundamental tool for aaRSs localization

To ascertain the hitherto unreported localization of pools of cytosolic aaRSs pool in a compartment can be done by various methods, but immunodetection will be almost inevitable. Apart from genetics and tagging one's favorite aaRS with a fluorescent protein, all techniques rely on the use of antibodies because of their high degree of specificity and sensitivity. The key step in confirming the localization of a cytosolic aaRS in a compartment with accuracy and reproducibility is to choose the specific antibodies against the marker protein of the compartment of interest. The

number of antibodies currently available has expanded in the last years, however published articles often do not report critical parameters (such as the origin of antibodies, the immunogen used to raise the antibodies...), which sometimes compromises the reproducibility of the results. Hence there is a push to require for at least, minimal reporting standards about the use of antibodies [62,63]. In Table 2, we provide an exhaustive list of proteins that were used as compartment markers for both immunofluorescence and Western blot, in studies that report the subcellular localization of an aaRS. More controls used for yeast cells can be found in the *Current Protocols in Cell Biology* paper by Rieder and Emr [64]. Immunodetection of aaRSs, like for any other protein, has so far been done on either native and unmodified aaRS or on tagged enzymes. Table 3 lists the few antibodies that were raised against native aaRSs and that have been used to determine their subcellular localizations. Since antibodies directed against native aaRSs can unpredictably cross-react with other proteins and be less sensitive than expected, most of the immunodetections of aaRSs were performed on tagged-proteins using anti-tag antibodies that usually are highly specific and sensitive. The main limit of this approach is that adding a tag at one extremity of the aaRS can potentially influence the intracellular localization of the chimeric enzyme [65]. Table 4 lists all the tags that have been fused so far to specific aaRSs of different organisms, which extremity of the protein was tagged, which antibody and what approach were used to detect the tagged aaRSs.

Table 2

Proteins used as controls for compartment staining and for confirming organellar localization of aaRSs.

Compartment	Application	Organism	Protein	Antibody source	References
Cytosol	WB	<i>Hs</i>	Tubulin beta	n.a.	[18]
			Hsp90 alpha	c.a.	[20]
		<i>Sc</i>	Pgk1	c.a.	[29]
			Gut2	c.a.	[32]
			eEF1A	n.a.	[36]
Nucleus	WB	<i>Hs</i>	Lamin B1	c.a.	[18,20]
			YY1	c.a.	[20]
		<i>Sc</i>	Hta2	c.a.	[29]
			Nop1	c.a.	[29]
			Nucleolin	c.a.	[20]
Nucleolus	IF	<i>Hs</i>	Nucleolin	c.a.	[20]
Mitochondria	WB	<i>Sc</i>	Por1	c.a.	[29]
			Atp2	n.a.	[32]
		<i>Tb</i>	Rip1	n.a.	[32]
			Cit1	n.a.	[32]
			mHSP70	n.a.	[36,37]
Apicoplast	IF	<i>Tb</i>	ACP	n.a.	[39,40]
		<i>Pf</i>	ACP	h.	[39,40]

WB, western blot; IF, immunofluorescence; *Hs*, *Homo sapiens*; *Sc*, *Saccharomyces cerevisiae*; *Tb*, *Trypanosoma brucei*; *Pf*, *Plasmodium falciparum*; n.a., not available; c.a., commercially available; h., homemade.

Table 3

Antibodies used for localizing untagged aaRSs.

Organism	Application	aaRS	Antibody	aaRS antigen	References	
<i>Hs</i>	IF	ArgRS	h.: rabbit	His-tagged 72 N-terminal residues of <i>HsArgRS</i>	[20]	
		Glu-ProRS	h.: rabbit	His-tagged peptide from D677 to E884 of <i>HsEPRS</i>	[20]	
		GlnRS	h.: rabbit	His-tagged 236 N-terminal residues of <i>HsGlnRS</i>	[20]	
		LysRS	c.a.; rabbit	n.d	[18]	
		MetRS	h.: rabbit	Full-length denatured His-tagged <i>HsMetRS</i>	[20]	
		PheRS	h.; rabbit	Endogenous PheRS from sheep liver	[21]	
		TyrRS	h.: rabbit	n.d	[25,26]	
		WB	TyrRS	h.: rabbit	n.d	[25,26]
			TrpRS	h.: rabbit	Full-length <i>HsTrpRS</i>	[22]
			GluRS	c.a.; rabbit	His-tagged full-length <i>ScGluRS</i>	[29]
		<i>Sc</i>	WB	HisRS	h.; mouse	Full-length <i>ScHisRS</i>
MetRS	c.a.; rabbit			His-tagged full-length <i>ScMetRS</i>	[29]	
<i>Pf</i>	WB, IF	CysRS	h.; rabbit	KLH conjugate 14 C-terminal residues of <i>PfCysRS</i>	[40]	

WB, Western Blot; IF, Immunofluorescence; *Hs*, *Homo sapiens*; *Sc*, *Saccharomyces cerevisiae*; *Pf*, *Plasmodium falciparum*; n.d., not described; c.a., commercially available; h., homemade.

Table 4
Protein-tags used for localizing aaRSs.

Organism	Tag	Tagged aaRS	Tag position	Application	Antibody	References		
Hs	GFP	ArgRS	C-ter	CM		[66]		
		AsnRS	C-ter	CM		[66]		
		LysRS	C-ter; N-ter	CM, FM		[19,66]		
		MetRS	C-ter; N-ter	CM, FM		[19,66]		
Sc	V5 epitope	GlyRS	C-ter	IF, WB	c.a.	[15]		
	GFP	MetRS	C-ter; N-ter	FM, CM		[29,34]		
		Nter-GluRS	C-ter	CM		[28]		
		GluRS	N-ter	FM, CM		[34]		
		TyrRS	C-ter	FM		[35]		
	V5 epitope	GlnRS	C-ter	WB, FM		c.a.	[67]	
	6× His	AlaRS	C-ter	WB		c.a.	[27]	
	c-Myc	TyrRS	C-ter	WB		c.a.	[35]	
	Tb	V5 epitope	CysRS	C-ter		IF	c.a.	[37]
			GluRS	C-ter		IF	c.a.	[37]
GlyRS			C-ter	IF	c.a.	[37]		
IleRS			C-ter	IF	c.a.	[37]		
MetRS			C-ter	IF	c.a.	[37]		
ProRS			C-ter	IF	c.a.	[37]		
SerRS			C-ter	IF	c.a.	[37]		
ValRS			C-ter	IF	c.a.	[37]		
Pf			3× HA	IleRS	C-ter	IF, WB	n.d.	[36]
				AlaRS	C-ter	IF, WB	c.a.	[39]
				CysRS	C-ter	IF, WB	c.a.	[40]
			GFP	GlyRS	C-ter	IF, WB	c.a.	[40]
				ThrRS	C-ter	IF, WB	c.a.	[39]
	AlaRS	C-ter		IF, WB	c.a.	[39]		
	3× HA	GlyRS	C-ter	IF, WB	c.a.	[39]		
		ThrRS	C-ter	IF, WB	c.a.	[39]		

Hs, *Homo sapiens*; Sc, *Saccharomyces cerevisiae*; Tb, *Trypanosoma brucei*; Pf, *Plasmodium falciparum*; CM, confocal microscopy; FM, fluorescence microscopy; IF, immunofluorescence; WB, western blot; c.a., commercially available; n.d., not described.

3. Prediction and analysis of aaRSs with organellar import signal

Among the many bioinformatic predictors of protein localizations, we chose 3 to predict the presence of an MTS and 3 to predict NLSs [68]. We used these predictors regardless of the organismal origin of the aaRSs analyzed. Moreover, to gather as much information as possible we tried the “Euk-mPloc2.0” website (<http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/>) which predicts whether an aaRS might be localized to the mitochondria, the nucleus, the cytosol or extracellularly without predicting any targeting sequence of any kind [69].

3.1. Mitochondrial targeting sequence prediction

MTS were predicted using TPpred2.0 (<http://tppred2.biocomp.unibo.it/tppred2/>) [70], TargetP1.1 (<http://cbs.dtu.dk/services/TargetP/>) [71,72] and MitoFates (<http://mitf.cbrc.jp/MitoFates/cgi-bin/top.cgi>) [73]. Query sequences should be given in the FASTA format, and depending on the MTS predictor, between 1 and 20 sequences can be analyzed at once. These predictors give back a mitochondrial probability score, the MTS sequence, the putative cleavage site that follows the MTS and the N-terminal peptide resulting from the cleavage. Additionally, MitoFates gives the mitochondrial processing enzymes responsible for maturation of the imported proteins. (Supp. Tables 3a and 3b). The MitoFates predictor was used for fungal, metazoan and plant proteins, while TargetP1.1 was used for non-plant organisms with no cut-off and with cleavage site prediction.

Submitting the yeast cytosolic aaRSs sequences to these predictors resulted in the HisRS having a strong probability of harbouring an MTS, the Cys-, Glu- and Val-RS a good probability and the Arg-, Leu- and Thr-RS only a low probability (Table 5). The predictions

have been partially validated since the relocalization of cytosolic His-, Glu- and Val-RS has been experimentally described (Table 1). The good probability of CysRS of having an MTS suggests that it is the cytosolic CysRS that aminoacylates the mitochondrial-encoded tRNA^{Cys} since there is no gene encoding a mitochondrial CysRS.

For human aaRSs, only the mitochondrial LysRS and GlyRS are encoded by the same gene as the cytosolic form. Mitochondrial LysRS originates from alternative splicing [14], and the mitochondrial GlyRS via an alternative start site [15,17]. Analyzing the remaining human aaRSs by the predictors shows that only the CysRS might contain an MTS, however, this cytosolic aaRS has never been described to localize to mitochondria (Table 5).

3.2. Nuclear localization signal prediction

To identify NLS, we used NLStradamus (<http://www.moseslab.cs.toronto.ca/NLStradamus/>) [74], SeqNLS (<http://mleg.cse.sc.edu/seqNLS/>) [75], and NLS-Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) [76] predictors. For all predictors, we used default settings. We added a 60 amino-acids N- or C-terminal cut-off for the predictions by NLS-Mapper since it is the only one able to predict bipartite NLSs which are usually found in the 60 first or last amino acids of a sequence. We therefore only kept bipartite NLSs of aaRSs fitting this criteria. This program is better suited for prediction purposes since instead of a simple “nuclear or not nuclear” output, it ranks protein localization on a scale from exclusively nuclear (8–10), partially nuclear (7–8), dual-localized cytosol-nucleus (3–5) or exclusively cytosolic (1–2). Supplemental Tables 3a and 3b list the results obtained with these 3 predictors for human and yeast aaRSs.

Schimmel and Wang, in their first attempt at identifying NLSs in the 20 yeast cytosolic aaRSs using the program PSORT II, showed that Ala-, Gly-, Pro-, **Asn-**, **Asp-**, **Gln-**, **Glu-**, **His-**, **Ile-**, **Leu-**, **Lys-**, **Pheβ-**, **Ser-**, **Tyr-** and **Val-**RS harbor at least one SV40

Table 5Predictions of mitochondrial targeting and nuclear localization signals among *S. cerevisiae* and human cytosolic aaRSs.

Class	ArgRS	MTS predictors ¹						Localization predictor Euk-mPloc2.0		NLS predictors ²							
		TPpred2.0		TargetP1.1		MitoFates		Yeast	Human	NLStradamus		SeqNLS		NLS Mapper (Monopartite)		NLS Mapper (Bipartite)	
		Yeast	Human	Yeast	Human	Yeast	Human			Yeast	Human	Yeast	Human	Yeast	Human	Yeast	Human
Class I	ArgRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C/M	C	n.p.	n.p.	n.p.	1	n.p.	n.p.	1	1
	CysRS	0.771	0.776	n.p.	n.p.	0.539	n.p.	C	C	1	1	1	2	n.p.	n.p.	1	4
	GluRS	0.896		0.357		n.p.		C	C	1		2		n.p.	n.p.	1	
	GlnRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	1	n.p.	1	2	1	n.p.	n.p.	2
	IleRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	1	n.p.	2	2	1	1	1	2
	LeuRS	n.p.	n.p.	0.538	n.p.	n.p.	n.p.	C	C	2	1	2	2	n.p.	n.p.	1	2
	LysRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	1	1	2	1	n.p.	n.p.	2	2
	MetRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	n.p.	1	1	1	n.p.	1	1	3
	TrpRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	1	n.p.	1	1	1	n.p.	1	1
	TyrRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C/N	C/E	1	n.p.	2	2	1	1	2	2
Class II	ValRS	0.871	n.p.	n.p.	n.p.	0.996	n.p.	C/M	C	1	1	1	2	n.p.	n.p.	n.p.	1
	AlaRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C/M	C	n.p.	n.p.	2	1	n.p.	n.p.	1	1
	AsnRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	1	n.p.	1	1	n.p.	2	1	2
	AspRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	1	n.p.	2	1	n.p.	n.p.	2	1
	GlyRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	n.p.	n.p.	1	2	n.p.	n.p.	n.p.	1
	HisRS	0.556	n.p.	0.846	n.p.	0.996	n.p.	C/M	C/M	1	n.p.	1	n.p.	1	n.p.	3	2
	PheRS1	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	n.p.	n.p.	1	1	n.p.	1	n.p.	2
	PheRS2	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	n.p.	n.p.	1	2	1	n.p.	1	2
	ProRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	n.p.	n.p.	2		n.p.	n.p.	n.p.	
	SerRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	1	1	1	2	n.p.	1	1	1
ThrRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C/M	C	n.p.	1	2	2	1	n.p.	2	3	
	GluProRS	n.p.	n.p.	n.p.	n.p.	n.p.	n.p.	C	C	1		2	2	1	1	1	1

n.p.: not predicted; for NLS sequence details see [Supp. Table 3a](#).¹ MTS scores are probabilities of the considered protein to have an MTS.² NLS scores are numbers of NLS sequences identified in each aaRS protein. C, cytosolic; M, mitochondrial; N, nuclear; E, extracellular.

T-antigen-like NLS or bipartite NLS (aaRSs in bold type have been found both in our study and by Schimmel and Wang) [49,77]. Using the aforementioned websites, we could classify the yeast aaRSs in three groups, those with a strongly probability of having an NLS (**Asn**-, **Asp**-, **Cys**-, **Gln**-, **Glu**-, **His**-, **Ile**-, **Leu**-, **Lys**-, **Pheβ**-, **Ser**-, **Trp**- and **Tyr**-RS), those likely to have one (**Ala**-, **Met**-, **Thr**- and **Val**-RS) and those unlikely to have one (**Arg**-, **Gly**-, **Pheα**- and **Pro**-RS) (Table 5). We see a high overlap between the published results and our own suggesting the robustness of the methods used. It is noteworthy that the **MetRS** missing from the original list was predicted by NLS-Mapper and then by SeqNLS. Despite the fact that these bioinformatics studies highlight the presence of NLSs in many cytosolic aaRSs, only 2 have been experimentally confirmed: the **MetRS** [29] and the **TyrRS** (3 bioinformatically found NLSs and one confirmed [35]). This emphasizes the need to use different algorithms to identify NLSs and that they are not yet sufficiently accurate to not require experimental confirmation.

Similar NLS prediction analysis has been done on human cells (Table 5). Using the aforementioned websites, we could classify the human aaRSs in three groups, those with a strong probability of having an NLS (**Cys**-, **GluPro**-, **Leu**-, **Lys**-, **Met**-, **Pheβ**-, **Ser**-, **Thr**- and **Val**-RS), those likely to have one (**Ala**-, **Arg**-, **Asn**-, **Asp**-, **Gln**-, **Gly**-, **Ile**-, **Pheα**-, **Pheβ**-, **Trp**- and **Tyr**-RS) and the ones unlikely to have an NLS (**HisRS**).

The fact that NLS predictions for yeast and human aaRSs converge seem to reinforce the hypothesis that this localization is true and not artifactual. The one notable exception is **HisRS** which is most likely nuclear in yeast and most unlikely in human. Since the nuclear localizations of the **MetRS**, **TyrRS** and **IleRS** have been determined experimentally, we can speculate that either some of their functions are shared and would have been present in their common ancestor, or that they exert different roles suggesting that the functions were acquired after the opisthokonta splitting (Table 5).

4. Subcellular fractionation

Subcellular fractionation can be divided into 3 different techniques according to the precision of the separation and the organelle that one wants to purify: crude membranes fractionation, nuclear purification and mitochondrial purification.

4.1. Crude membranes fractionation from yeast cells

This technique allows the separation of membranes from the cytosol (S100) and yields 2 separate membrane fractions (P13 and P100). To analyze the subcellular distribution of yeast aaRSs we performed a subcellular fractionation adapted from the protocol published by Bonangelino and collaborators [78].

- Yeast cells are inoculated in the morning or the previous evening depending on the generation time in the appropriate medium (YPD or SC dropout) on a rotary shaker at 200 rpm at 30 °C.
- In the evening, the cells are diluted in 100 mL of the same medium to reach an OD_{600nm} of about 0.5 in the morning. 40 units of OD_{600nm} are transferred to a 50 mL Falcon tube. From here on out, all steps are performed on ice and centrifugations are performed at 4 °C.
- The cell suspension is centrifuged for 5 min at 5250g at 4 °C to pellet the cells. For cells grown on SC dropout medium, add 1 mL of YPD to facilitate the pelleting.
- Cells are washed twice in 20 mL of ice cold fractionation buffer (20 mM HEPES/KOH, pH 6.8, 50 mM CH₃COOK, 10 mM MgCl₂, 250 mM Sorbitol, 10 mM Na₂N₃).
- Cells are then resuspended in 800 μL of fractionation buffer (with protease inhibitor cocktail cOmplete ULTRA Tablets, EDTA-free from Roche™) and transferred into a 15 mL cold Cor-eX®II centrifuge tube (ref. 1-8441-15) containing 650 μL of glass beads (acid washed 0.25–0.5 mm diameter beads).

- The tube is vortexed 6 times for 30 s at full speed to lyse cells.
- The lysate is transferred to a 1.5 mL microfuge tube and centrifuged at 4 °C for 5 min at 500g in a fixed-angle rotor to clarify it.
- The S5 supernatant is carefully transferred to a fresh microfuge tube and centrifuged at 4 °C for 10 min at 13,000g in a fixed-angle rotor.
- The pellet (P13), which mainly contains the plasma membrane, the endoplasmic reticulum, the nucleus, the vacuole and mitochondria, is kept on ice until the end of the experiment.
- The supernatant (S13) is transferred in a 1.5 mL ultracentrifugation microfuge tube (Beckman Coulter® ref. 357448) for further centrifugation at 4 °C for 1 h at 100,000g in an ultracentrifuge using a fixed rotor.
- The supernatant (S100), which contains cytosol and vesicles, is carefully transferred to a fresh 1.5 mL microfuge tube. The pellet (P100) contains Golgi and endosomes.
- P13 and P100 are subsequently resuspended in fractionation buffer in the same volume as the S100 fraction.

4.2. Purification of nuclei and preparation of nuclear protein extracts

Purification of nuclei needs to be performed quickly to avoid passive diffusion of small proteins through nuclear pores. Moreover, the composition of buffers is critical for maintaining osmotic balance to prevent nuclear proteins from leaving the nucleoplasm fraction.

4.2.1. Purification of nuclei from yeast cells

This method is an adaptation from the original protocol of Stephanie E. Rieder and Scott D. Emr [79].

- Yeast cells from 2 L of culture in rich or synthetic medium, depending on auxotrophic markers or studied conditions, are harvested at 0.7 OD₆₀₀ and washed once with water to ensure that all traces of medium are eliminated.
- The wet pellet is weighed (approximately 7 g/2 L of culture). Pellet is resuspended in 100 mM Tris H₂SO₄, 10 mM DTT in a volume of 2 mL/g of wet cells and incubated 20 min at 30 °C with gentle agitation. The purpose of this step is to weaken the cell wall.
- The mixture is centrifuged 4 min at 3500g and the pellet is washed with 1.1 M sorbitol to keep osmotic pressure, 100 mM K₂HPO₄/KH₂PO₄, pH 7.4 using 7 mL of buffer per g of cells and resuspended with the same buffer.
- Zymolase 20T is added to the mixture (5 mg/g of pellet) and incubated during 45 min at 30 °C with gentle agitation (80 rpm), to digest the cell wall.

All following steps are performed at 4 °C, to avoid protein degradation. Approximately 3.5 g of wet cells should be settled by falcon 50 (two falcons 50 for 2 L of culture at 0.7 OD_{600nm}). Note that all the subsequent steps are detailed for the treatment of 1 falcon 50 out of the two necessary to treat the whole 2 L culture.

- The spheroplasts (cells without their wall) are centrifuged 4 min at 3500g, gently resuspended in 1.1 M sorbitol, 100 mM K₂HPO₄/KH₂PO₄, pH 7.4 to remove the zymolase and centrifuged again 4 min at 3500g.
- The pellet is resuspended in 20 mL 1.1 M sorbitol, 100 mM K₂HPO₄/KH₂PO₄, pH 7.4 and loaded on a 6 mL cushion containing 30 mM sorbitol, 5% (v/v) ficoll 400, with protease inhibitors. The spheroplasts are centrifuged 10 min at 4000g.
- The pellet is resuspended into 15 mL of ice cold ficoll 20% (v/v), 20 mM K₂HPO₄/KH₂PO₄, pH 7.4, 1 mM MgCl₂ and protease inhibitors, and the spheroplasts are broken by osmotic shock

in potter by 20 S in maximum 5 min. The ficoll keeps the osmotic pressure between the medium and the nuclei, to avoid nuclear protein from diffusing out of the nucleus.

- The lysate is incubated 10 min on ice to let nuclei cool down prior to subsequent centrifugation during 5 min at 13,000g.
- The supernatant is collected and immediately centrifuged again 10 min at 13,000g. At this step, nuclei cannot pellet because of the viscosity of the buffer.
- The supernatant is then loaded on a ficoll gradient containing 20 mM K₂HPO₄/KH₂PO₄ pH 7.4, 1 mM MgCl₂, protease inhibitors and ficoll 50% (v/v), ficoll 40% (v/v), ficoll 30% (v/v), 6.5 g each in 25 × 89–mm ultracentrifuge tube (Beckman Coulter®, Ultra Clear Tubes, 38.5 mL, ref. 344058). Then, the gradient is ultracentrifuged 1 h at 58,000g (Beckman, SW28 rotor, 18,000 rpm).
- After the centrifugation, the 20 and 30% layers containing all the cellular debris are removed. The 40% layer and the interface 40–50% containing the nuclei are harvested by pipetting. Nuclei are also present in the interface 30–40%, but this interface layer is not harvested because it contains too many non-nuclear contaminants.
- To remove the ficoll, the samples are diluted 10 times in 20 mM K₂HPO₄/KH₂PO₄, pH 7.4, 1 mM MgCl₂ and centrifuged 10 min at 10,000g to pellet nuclei (since the ficoll is diluted, nuclei are able to pellet at 3000g).
- Nuclei are finally resuspended in 500 µL of buffer containing 100 mM Tris HCl pH 7.8, 0.1 mM EDTA, 5 mM β-Mercaptoethanol, 1 mM benzamidine, protease inhibitors, and sonicated for 80 s (1 s on/1 s off, 30% amplitude, Bioblock Vibracell) in order to disrupt the nuclear envelope.
- After a chilling step on ice to avoid foaming, nuclei are re-sonicated for 30 s (1 s on/1 s off, 30% amplitude, Bioblock Vibracell).
- The nuclear extracts are ultra-centrifuged 30 min at 100,000g to pellet nuclear membranes. The supernatant contains soluble nuclear proteins.

4.2.2. Purification of nuclei from human cells

This protocol is adapted from Jason M. Dahlman and Denis C. Guttridge [80]. Perform all steps at room temperature unless otherwise specified. Volumes are determined for the preparation of 10 samples. All centrifugation was performed in a bench top microfuge. A minimum of 1 × 10⁶ cultured cells are used for this protocol.

- Cell culture media is removed from the plates.
- Cells are gently washed twice with 1 mL of PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄).
- After the addition of 1 mL of PBS, cells are removed from the polystyrene tissue culture dish using a cell scraper.
- Scraped cells (1 mL) are removed with a 1 mL pipette and placed in a 1.5 mL microcentrifuge tube.
- Cells are centrifuged at 250g for 5 min at 4 °C.
- After centrifugation, supernatants are removed and 5 pellet volumes of cytoplasmic extraction buffer (CEB: 10 mM HEPES-KOH, pH 7.6, 60 mM KCl, 1 mM EDTA, 0.25% (v/v) Tergitol-type NP-40 (Sigma-Aldrich), Complete protease inhibitor cocktail, EDTA-free (Roche)) are added to each sample.
- Cellular pellets are dissolved by flicking the tubes. It is important to minimize the amount of bubbles generated while flicking. After adding the cytoplasmic extraction buffer, let samples stand on ice for 1–3 min before resuspending and make sure that cellular pellets are completely dissolved.
- Resuspended pellets are incubated on ice for 5 min and then, centrifuged at 650g for 4 min at 4 °C.

- After centrifugation, supernatants, containing the cytoplasmic proteins, are carefully transferred to new 1.5 mL microcentrifuge tubes. While removing the supernatant, take special care not to disrupt the nuclear pellet, which now should take on an opaque appearance. The cytoplasmic extract can be either disposed of or stored at -80°C .
 - 100 μL of cytoplasmic wash buffer (CWB: 10 mM HEPES-KOH, pH 7.6, 60 mM KCl, 1 mM EDTA, Complete protease inhibitor cocktail, EDTA-free (Roche)) is added to the nuclear pellets to remove the excess of NP-40. The 1.5 mL microcentrifuge tubes are gently tapped to dislodge the nuclear pellets. There is no need to resuspend the pellets.
 - Centrifugation is repeated at 650g for 4 min at 4°C .
 - Supernatants are carefully discarded, once again taking care not to disturb the nuclear pellets.
 - 1 pellet volume of nuclear extraction buffer (NEB: 20 mM Tris HCl pH 8.0, 420 mM NaCl, 1.5 mM MgCl_2 , 0.2 mM EDTA, 25% (v/v) glycerol, Complete protease inhibitor cocktail, EDTA-free (Roche)) is added to the 1.5 mL microcentrifuge tubes containing the nuclear pellets and they are resuspended by flicking the tubes until they are completely resuspended. Try to minimize the amount of bubbles you generate while flicking.
 - The resuspended nuclear pellets are incubated on ice for 10 min and mixed by flicking every 2 min.
 - After the 10 min incubation, the resuspended nuclear pellets are centrifuged at 15,000g for 10 min at 4°C .
 - After centrifugation, supernatants, containing the nuclear proteins, are transferred to new pre-chilled 1.5 mL microcentrifuge tubes
 - Nuclear extracts are stored at -80°C .
- Samples are then 2-fold diluted with ice-cold homogenization buffer and the lysate is cleared by centrifugation at 1500g for 5 min to remove cell debris and nuclei.
 - The supernatant is centrifuged at 4000g for 5 min, the pellet is discarded and the supernatant is centrifuged again at 12,000g for 15 min to isolate the mitochondrial fraction.
 - The supernatant is discarded and the mitochondrial pellet is resuspended in 0.5 to 1 mL of SEM buffer (250 mM sucrose, 1 mM EDTA, 10 mM MOPS-KOH, pH 7.2) and the protein concentration is measured using the Bradford method and adjusted to 5 mg/mL.

In order to obtain highly purified mitochondria, firstly prepare sucrose step gradients in Beckman centrifuge tubes (Beckman Coulter[®], Ultra Clear Tubes, 13.2 mL, ref. 344059):

- 1.5 mL 60% (w/v) sucrose diluted in EM buffer (1 mM EDTA, 10 mM MOPS-KOH, pH 7.2) is loaded at the bottom of the centrifuge tube.
- Without disturbing the phases, 4 mL 32% (w/v), 1.5 mL 23% (w/v), and 1.5 mL 15% (w/v) sucrose/EM are carefully pipetted stepwise and the tubes are kept at 4°C .
- Samples are homogenized again in a glass Teflon homogenizer (10–15 strokes) and poured (0.2 to 1 mL) to the sucrose gradient.
- The clarified samples are then centrifuged (Beckman, SW41 Ti rotor) at 134,000g for 1 h at 4°C and the highly purified mitochondria are collected with a Pasteur pipette from the 60–32% sucrose interface.
- Mitochondrial samples are diluted 2-fold in SEM buffer and centrifuged at 12,000g for 15 min. Repeat this step to wash the purified mitochondria.
- Pellet can be frozen at -80°C .

4.3. Purification of mitochondria and preparation of mitochondrial protein extracts

4.3.1. Purification of mitochondria from yeast cells

To analyze the mitochondrial localization of aaRSs, we performed a mitochondria purification (described below) adapted from the protocol published by Meisinger and collaborators [48].

- All steps are performed at 4°C unless otherwise specified. To get enough mitochondria it is recommended to use at least 2–3 L of yeast cells grown to log-phase.
- Yeast cells are pelleted by centrifugation at 3000g for 5 min and resuspended in distilled water.
- Cells are pelleted again as for nuclei preparation, the wet weight is measured and the pellet is resuspended at a concentration of 2 mL/g of pellet in prewarmed at 30°C DTT buffer (100 mM Tris- H_2SO_4 , pH 9.4, 10 mM dithiothreitol (DTT)).
- After 20 min incubation in a 30°C shaker at 80 rpm, cells are pelleted at 3000g for 5 min and resuspended in 7 mL/g of zymolyase buffer (1.2 M sorbitol, 20 mM potassium phosphate, pH 7.4), pelleted again at 3000g for 5 min and resuspended at a concentration of 7 mL/g of cells in zymolyase buffer containing 3 mg Zymolyase 20T (Seikagaku Kogyo Co., Tokyo, Japan) per gram wet weight.
- After incubation for 45 min in 30°C shaker at 80 rpm, cells are pelleted at 3000g for 5 min and resuspended in 7 mL/g of fresh zymolyase buffer, pelleted again at 3000g for 5 min and resuspended (6.5 mL/g wet weight) in ice-cold homogenization buffer (0.6 M sorbitol, 10 mM Tris HCl, pH 7.4, 1 mM EDTA, 0.2% (w/v) BSA (Sigma-Aldrich), and 1 mM of freshly prepared PMSF).
- From this step, it is very important to maintain the lysate, buffers and rotors at 4°C to prevent proteolysis.
- Samples are poured in a glass Teflon homogenizer and spheroplasts are homogenized with 15 S.

For extraction of soluble mitochondrial proteins, mitochondria are disrupted by sonication (see below) otherwise they are resuspended in SEM buffer and the concentration is adjusted as required.

- Mitochondria pellets are resuspended in 1 vol of lysis buffer (50 mM Tris-HCl, pH 7.2, 1.1 M Sorbitol, 2 mM EDTA; 5 mM β -mercaptoethanol).
- Sonication (4 run of 10 S with a 2 s duration time, amplitude 25%; Vibracell) is performed with a small probe (2 mm diameter) on ice.
- Broken mitochondria are then centrifuged for one hour at 105,000g at 4°C to eliminate unbroken mitochondria and disrupted membranes.
- The supernatant which is recovered, corresponds to soluble mitochondrial proteins and can be frozen at -80°C .

4.3.2. Obtention of mitoplasts from yeast mitochondria

For generation of mitoplasts, digitonin is used to permeabilize the outer mitochondrial membrane of yeast mitochondria.

- Mitochondria are suspended at a concentration of 1 mg/ml in SEM buffer.
- 250 μg of mitochondria are incubated for 25 min on ice in 250 μL of SEM buffer containing 0.05% to 0.2% digitonin.
- The resulting mitoplasts are harvested by spinning at 14,000 rpm for 10 min and the supernatant is saved for further analysis.
- The mitoplast containing pellet is washed once again with SEM buffer.
- To determine the efficiency of digitonin treatment, 100 μg of mitochondria are treated as above with digitonin, centrifuged,

and the pellet and supernatant fractions separated by SDS-PAGE and transferred to a polyvinylidene difluoride membrane (Immobilon P, Millipore). Immunoblots are performed using antibodies specific for cytochrome *c* peroxidase (CCPO) (intermembrane space), TIM23 (inner membrane), delta-1-pyrroline-5-carboxylate dehydrogenase (Put2p) (matrix) and porin (outer membrane).

4.3.3. Purification of mitochondria from human cells

This protocol is adapted from Jason M. Dahlman and Denis C. Guttridge [80] and has been designed to extract mitochondria from HeLa cells. Volumes are adapted to cells grown in 5 flasks of 58 cm². All steps are performed on ice or at 4 °C.

- Cell culture growth media from desired cells is removed.
- Cells are gently washed twice with 5 mL of PBS (see Section 4.2.2).
- Cells are removed from the polystyrene tissue culture dish by adding 1 mL of PBS and the cells are scraped gently with a cell scraper.
- Cells are centrifuged for 5 min at 250g at room temperature
- The pellet is resuspended in 1 mL of BSA buffer (0.6 M sorbitol, 1 mM EDTA, 20 mM Hepes-KOH, pH 7.6, 300 mM NaCl, Complete protease inhibitor cocktail, EDTA-free (Roche) and 0.3% (v/v) of BSA (1 mg/mL)).
- Cells are then mechanically broken on ice by repeated (at least 80 times) aspiration/backflow using a 1 mL syringe with a G23 needle.
- The lysate is centrifuged twice at 1200g for 6 min at 4 °C.
- The supernatant is then centrifuged at 16,000g for 40 min at 4 °C.
- Pellet can be frozen at –80 °C.

4.3.4. Obtention of mitoplasts from human mitochondria

For generation of mitoplasts, digitonin is used to permeabilize the outer mitochondrial membrane of human mitochondria.

- Mitochondria-containing pellet is resuspended in 300 µL BSA buffer (0.6 M sorbitol, 1 mM EDTA, 10 mM Hepes-KOH, pH 6.7 and 0.3% (v/v) of BSA (1 mg/mL)) containing 0.2% digitonin (50 µL of a 1 mg/mL stock solution) and incubated 15 min at room temperature.
- 1 mL of BSA buffer is added and the mixture is centrifuged 20 min at 16,000g at 4 °C.
- The resulting pellet is washed two times with 1 mL BSA buffer without resuspending the mitoplasts.
- The final mitoplasts-containing pellet can be resuspended in various types of buffers depending on the subsequent analyses that will be performed (TRIzol (Invitrogen) for RNA extraction, Laemmli for SDS-PAGE analysis...) or stored at –80 °C as a wet pellet.

5. Mass spectrometry analysis and identification of cytosolic aARs in yeast mitochondrial extracts

We performed yeast mitochondria purification as explained in the Section 4.3.1 and recovered pure mitochondria at the 32–60% interphase from a cushions gradient (Fig. 1A). In order to identify mitochondrial proteins, mitochondria were disrupted (see Section 4.3.1) and soluble proteins were subjected to mass spectrometry (see protocol below). For analysis, we carried out 3 different injections, two of 90 min (one regular and one using restricted mass ranges of 400–800 *m/z* and 800–1250 *m/z*, termed Frac.) and one of 120 min. These features allowed us to identify 905 (90 min.), 982 (90 min. Frac.) and 926 (120 min) mitochondrial proteins respectively (Fig. 1B). We describe here the detailed procedure

for protein digestion and mass spectrometry analysis as well as data analysis.

5.1. Protein digestion solution

- Proteins from mitochondrial extracts are resuspended in 50 mM ammonium bicarbonate.
- After a reduction-alkylation step (5 mM DTT and 10 mM iodoacetamide), proteins are digested overnight with 100 ng of sequencing-grade trypsin (Promega).
- After centrifugation at 12,000g, the supernatants are collected in glass inserts and vacuum dried.

5.2. Nano-liquid Chromatography – electrospray Ionization TripleTOF MS/MS Analysis

- Before injection, dried peptides are resuspended in 15 µL of 0.1% (v/v) formic acid. One third of each sample was injected on a NanoLC-2DPlus system (nanoFlexChiP module; Eksigent, ABSciex, Concord, Ontario, Canada) coupled to a TripleTOF 5600 mass spectrometer (ABSciex) operating in positive mode.
- Peptides are loaded with a trap and elute configuration on C18 reverse-phase columns (ChiP C-18 precolumn 300 µm ID × 5 mm ChromXP and ChiP C-18 analytical column 75 µm ID × 15 cm ChromXP; Eksigent).
- Peptides are eluted by using a 5–40% gradient of 0.1% (v/v) formic acid in acetonitrile for 90 or 120 min at a 300 nL/minute flow rate. The TripleTOF 5600 was operated in high-sensitivity data-dependant acquisition mode with Analyst software (v1.6, ABSciex) on a 400–1250 *m/z* range.
- To extend the sensitivity, two new injections are carried out using restricted mass ranges, 400–800 *m/z* and 800–1250 *m/z*. An external calibration is performed before each sample by monitoring 10 peptides of a beta-galactosidase trypsin digest. A discovery “Top20” method is used: up to 20 of the most intense multiply-charged ions (2+ to 5+) are selected for CID fragmentation, with a cycle time of 3.3 s.

5.3. Database search and data analysis

Raw data are first converted to Mascot Generic File format (.mgf) and searched against the yeast *S. cerevisiae* database supplemented by a decoy database (reverse sequences). The database search algorithm used is Mascot (version 2.2, Matrix Science, London, UK) through the ProteinScape 3.1 package (Bruker Daltonics, Leipzig, Germany). Peptide modifications allowed during the search are: N-acetyl (protein), carbamidomethylation(C) and oxidation (M). Mass tolerances in MS and MS/MS is set to 20 ppm and 0.5 Da, respectively. Two trypsin missed cleavages sites is allowed. Peptide identifications obtained from Mascot is validated with a protein FDR < 1%, using the Protein Assessment tool from ProteinScape. Identified proteins are assessed by spectral count.

The aARs identified in the mitochondrial extract analyzed by three different injections methods are presented in Fig. 1C and Supp. Table 2. Excepting the Arg-, Met- and Thr-RS (note that there are no genes encoding the mitochondrial CysRS and GlnRS in yeast), we found all the mitochondrial aARs in our assay. The absence of these three mitochondrial aARs could be due to a strong binding to the mitochondrial membrane that resulted in their loss during the protein extract preparation. Indeed, it has previously been shown that a functional aAR could be membrane-anchored [81]. In addition to the mitochondrial aARs, we identified 10 cytosolic ones (Asp-, Glu-, Ile-, Leu-, Lys-, Phe-, Pro, Ser-, Thr- and Tyr-RS) of which only the cytosolic GluRS has been shown experimentally to be dual-localized [28,29]. Since there is no gene identified for the yeast mitochondrial CysRS, we assume

Table 6

List of markers for specific compartments in mammalian cells and their respective antibodies.

Compartment	Protein marker*	Antibody
Nucleus	Lamin B1	Abcam ab16048
Plasma membrane	Cadherin	Abcam ab6528
Endoplasmic Reticulum	Calnexin	Abcam ab22595
	Calreticulin	Abcam ab2907
	GRP-78	Abcam ab21685
	GM130	Abcam ab31561
Golgi	Giantin	Abcam ab80864
	TGN46	Abcam ab2809
	COX IV-1	Abcam ab14744
Mitochondria	Cytochrome c	Abcam ab13575
	Early endosome antigen 1	Abcam ab2900
Endosomes	Ras related protein Rab-5A	Abcam ab18211
	Ras related protein Rab-7a	Abcam ab126712
	LAMP-1	Abcam ab24170
Lysosomes	LAMP-2	Abcam ab18529
	MAP1A/MAP1B LC3 A	Abcam ab52768
Autophagosomes	Catalase	Abcam ab16771

* Protein marker names are recommended names or short names from the UniProt database.

that the cytosolic form would also be mitochondrial; and we, indeed, found it in the mitochondrial fraction we analyzed. It is now important to confirm the mitochondrial localization of these cytosolic aaRSs and then determine their function in this other compartment.

6. Microscopy analysis and single cell localization

Unlike subcellular fractionation, microscopic analysis allows the study of aaRSs localization at the single cell scale. For more precision, a confocal microscope is needed to set only one focal plan and prevent accumulation of cytosolic fluorescence around the organelle. If the aaRS is fused to GFP, analysis can be performed directly on living cells, and aaRSs used with a GFP-tag are listed in Table 4. To control the aaRS localization on living cells, the need of compartment markers is essential, like DAPI and Mitotracker for nuclear and mitochondrial compartment, respectively. For yeast cells, some additional fluorescent dyes can also be used for fluorescent microscopy on living cells, as CellTracker Blue CMAC (ThermoFischer Scientific, ref. C2110) that is specific for vacuolar lumen (DAPI filter), or FM4-64 (ThermoFischer Scientific, ref. T3166) used in time lapse experiments to stain the plasma membrane, the endosomes and the vacuolar membrane (RFP filter) [82]. For more precision, protein markers can be targeted by specific antibodies: frequently used markers for mammalian intracellular compartments are listed in Table 6. As GFP can possibly perturb the aaRS localization due to its large size, smaller tags (listed in Table 4) have been used for immunofluorescence (IF) study. For IF, a fixation step is essential to allow antibodies penetration in the sample. Next sections present immunofluorescence protocols for both yeast and human cells.

6.1. Immunofluorescence on yeast cells

This technique allows the observation by fluorescent microscopy of the aaRSs of interest on fixed yeast cells.

6.1.1. Fixation steps

- Cells are grown overnight in 20 mL of either complete or synthetic medium (YPD or SC dropout medium) to about 1×10^7 cells/mL ($OD_{600nm} = 0.4-0.6$); 10 mL culture gives 6 samples.
- 1.25 mL of 1 M K_2HPO_4/KH_2PO_4 , pH 6.5 and 1.25 mL of 37% (v/v) formaldehyde are added to 10 mL of overnight culture

(10×10^7 cells) and incubated for 2 h at room temperature with gentle shaking every 30 min.

- Cells are centrifuged for 5 min at 3000g, washed 3 times with 10 mL SP buffer (1.2 M sorbitol, 0.1 M KPO_4 , pH 6.5) and resuspended in 1 mL SP/ME (SP, 20 mM β -mercaptoethanol).
- 10 μ L Zymolyase 20T (5 mg/mL) is added and the mixture is incubated 45 min at 30 °C (gentle shaking every 10 min); the amount of Zymolyase needed depends on the strain.
- Cells are centrifuged for 5 min at 2000g, washed 3 times with 3 mL SP buffer and resuspended in 0.1 mL SP; after the digestion of the cell wall, the resuspension has to be very gentle to avoid lysis.
- Cells are stored at 4 °C (can be kept for 1–2 weeks).

6.1.2. Incubation with antibodies

The following steps are done on a slide for immunofluorescence (HTC or epoxy, single use slides, diameter 6 mm) at room temperature.

- Slides are washed with ethanol 96% (v/v).
- 20 μ L of poly-Lysine (0.1% (w/v) in H_2O) is added to each well and incubated for 1 min.
- The liquid is removed, slides are dried and washed 3 times with 20 μ L H_2O ; all washing steps are done by adding one drop (with Pasteur pipette) and removing with vacuum without touching the slide.
- Slides are placed in a humid chamber (Petri dish with H_2O -soaked paper), 15 μ L of fixed yeast cells are added to each well and incubated for 30 min.
- The excess of liquid is removed very gently and slides are washed twice with PBS 1 \times (see Section 4.2.2).
- Slides are then incubated for 5 min with 15 μ L of PBT (PBS 1 \times , Triton X-100 0.1% (v/v), NaN_3 0.1% (w/v), BSA 1% (w/v)).
- After removing the excess of liquid, slides are incubated for 30 min at room temperature with 15 μ L of primary antibody diluted in PBT, according to manufacturer indications.
- Slides are washed 10 times with PBT, followed by the incubation for 30 min in the dark and at room temperature with 15 μ L of secondary antibodies (either Cy-conjugated or Alexa Fluor[®]-conjugated; ThermoFischer Scientific) diluted in PBT.
- Slides are washed 10 times with PBS.
- Optional: Nuclei can be stained with 15 μ L of DAPI (10 μ g/mL) and incubated 5 min in the dark. 5 washes with PBS are needed to avoid non-specific background fluorescence.
- Finally, 3 μ L of Vectashield[®] (Vector laboratories) are added and slides are covered with coverslip; fix coverslip on all ends with nail polish.
- Slides can be stored at 4 °C in the dark (can be kept for months) or directly observed by fluorescence microscopy.

6.2. Immunofluorescence on human cells

This technique allows observing the subcellular localization of the aaRSs of interest in fixed mammalian cells on a single cell scale with a fluorescence microscope.

- Cells are grown in complete medium, mostly DMEM (D5796 Sigma) or MEM (M2279 Sigma), supplemented with 10% (v/v) Fetal Calf Serum (FCS).
- On day one of the experiment, the medium of the culture flask is removed, cells are washed with PBS (see Section 4.2.2.) and detached using trypsin (T4049 Sigma).
- Cells are then spotted on 10-well spotslides (Marienfeld Superior, #1216650) to a 70% confluency (for most cell types 4000–5000 cells per spot), let to attach and grown overnight in a cell culture incubator at 37 °C, 5% CO_2 .

- After removing the medium, cells are fixed by adding 50 μ L of 4% (v/v) formaldehyde (F8775 Sigma) in PBS to the spots and incubated for 20 min at room temperature.
- If Mitotracker (ThermoFisher Scientific, M7512) is used to detect mitochondria, Mitotracker staining has to be done on live cells prior to fixation. Therefore, cells are incubated for 15–45 min at 37 °C with Mitotracker diluted in pre-warmed medium (37 °C) to a concentration of 100–500 nM. Cells are then washed once in medium and once in PBS and fixed as described above.
- Cells are then washed once with PBS and permeabilized with 0.2% (v/v) Triton-X 100 (Sigma-Aldrich) in PBS for 10 min at room temperature.
- After one more wash with PBS, unspecific sites are blocked by incubation with 20% (v/v) FCS in PBS for one hour at RT.
- Cells are then incubated with the antibodies raised against the aaRS of interest and marker proteins of the different subcellular compartments in PBS-2% (v/v) FCS for two hours at room temperature or overnight at 4 °C.
- After washing five times with PBS, the fluorophore-conjugated secondary antibodies diluted in PBS-2% (v/v) FCS are added for one hour at room temperature. DAPI can be added to the secondary antibody solution at a concentration of 300 nM.
- Finally, after two final washes in PBS, a drop of pre-warmed (50 °C) elvanol (Poly(vinyl alcohol) – 341584 Sigma-Aldrich) is added to each spot and the slide sealed with a coverslip. Cells are then visualized using a confocal laser scanning microscope and fluorescence can be quantified with the ImageJ software.

7. Concluding remarks

Localizing multi-compartmental aaRSs cannot be done using a one-for-all technique. For example, using fluorescence microscopy to simultaneously visualize in a single cell, the organellar and cytosolic fractions of the same aaRS (or protein in general) is complicated since the fluorescence of the organellar pool will be masked by the fluorescence of the cytosolic portion which is usually more abundant. So far no tools have been developed to overcome this difficulty and the use of confocal microscopy is usually not sufficient when one wants to ascertain the localization of a minor organellar pool of a cytosolic aaRS. Moreover, like for any other protein putatively organellar, it is currently very difficult to distinguish between a luminal localization of an aaRS and a peripheral localization of the protein on the cytosolic side of the organelle. Being able to make these distinctions is extremely important to be able to study the nonconventional roles of these multi-localized aaRSs and identify the cellular processes to which the organellar pools of these cytosolic aaRSs participate. Also, the possibility that several cytosolic aaRSs could be peripherally associated to membrane compartments other than the mitochondria and the nucleus by interacting either with membrane proteins or directly to lipids has been overlooked so far. For all these reasons, one still has to turn to the isolation of membranes and afterwards to organelle purification to be able to ascertain the relocation inside a given compartment of a fraction of a cytosolic aaRS (summarized in Table 1).

Another difficulty to overcome when working with dual-localized aaRSs is that the cytosolic pool of these proteins is always essential for viability, rendering the study of the functions of the organellar pools of the same aaRS very difficult. In this respect, yeast is a very convenient model especially when one wants to analyze the mitochondrial activity of a cytosolic aaRSs. Since these organelles are only essential for the respiratory metabolism and not when yeast is fermenting, mutations of the cytosolic aaRS gene that lead to only a respiratory deficiency (fermentative growth is unaffected) are targeting residues essential for the mitochondrial

activity of the dual-localized aaRS without affecting its cytosolic role. This simple phenotypic screen was successfully used to identify the cryptic and nonconventional MTS of yeast cytosolic GluRS which participates to the mitochondrial transamidation pathway upon mitochondrial import [28]. It was also used to show that mitochondrial forms of His-, Val- and GlyRSs are generated through an alternative translational start [13,31,32].

If confirming that a pool of a cytosolic aaRS can relocate inside an organelle is difficult with the current technologies, identifying the import signals to study the dynamics of these proteins is even more arduous. Import signal predictors still have a low consensual predictive accuracy for these dual-localized aaRSs, mainly because they often contain nonconventional import signals. Identifying the organellar roles of dual-localized aaRSs will necessitate the reinterpretation of previous yeast mutant screens as well as establishing new mutant screens. In humans, the systematic sequencing of genes encoding cytosolic aaRSs of patients with mitochondrial-related diseases and in depth analysis of the localization, protein interactants and activity of the mutant aaRS is essential for identifying the organellar functions of dual-localized aaRSs. By doing so, it was recently shown that the mutation in the cytosolic GlyRS gene that causes type 2D Charcot-Marie-Tooth disease changes the protein binding activity of GlyRS that interferes with a signaling pathway essential for motor neuron survival [83].

These findings also raise many questions, the most important being how cells regulate the distribution of multi-compartmental cytosolic aaRS between the cytosol and the different organelles. One mechanism by which cells manage the subcellular distribution of the various pools of a cytosolic aaRS is by forming the MSCs and triggering release of each cytosolically-anchored aaRS by post-translational modification of the aaRS. Since so far, a maximum of 9 cytosolic aaRSs were found participating to MSCs, there are probably others posttranslational modifications, or other mechanisms that participate in regulating the dynamic compartmentalization of the remaining 11 potentially multi-localized aaRSs. Another important issue concerning the dynamical relocation of these aaRSs is whether a portion of an existing pool is redistributed or a fraction of the neosynthesized proteins will specifically be deviated on demand towards the organelle? Finally, while in this review we focused on the relocation of cytosolic aaRSs within the cell, it has to be noted that several human aaRSs can be secreted in the extracellular medium where they also accomplish nonconventional functions [84,85].

Acknowledgements

The work was supported by the French National Program Investissement d'Avenir administered by the "Agence Nationale de la Recherche" (ANR-France), "MitoCross" Laboratory of Excellence (Labex), funded as ANR-10-IDEX-0002-02 (to H.D.B, J-O.D.C), the University of Strasbourg (H.D.B, E.M), the CNRS (B.S, S.F), the INSERM and the Université de Strasbourg through the IDEX 2015 Attractivité (to S.B), the Agence Nationale de la Recherche – ANR-13-BSV2-0004 to S.F and AFM-Téléthon – AFM-SB/CP/2013-0133/16551 (to S.F), the Ministère de l'Éducation Nationale, de la Recherche et de l'Enseignement Supérieur (G.B, S.D), the Association pour la Recherche sur le Cancer (D.L) and the ANR [ANR-10-LABX-0036_NETRNA] (L.E, P.H).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ymeth.2016.09.017>.

References

- [1] M. Ibba, D. Söll, Aminoacyl-tRNA synthesis, *Annu. Rev. Biochem.* 69 (2000) 617–650.
- [2] D.R. Smith, P.J. Keeling, Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes, *Proc. Natl. Acad. Sci. U.S.A.* 112 (2015) 10177–10184.
- [3] M. Sissler, J. Pütz, F. Fasiolo, C. Florentz, Mitochondrial aminoacyl-tRNA synthetases, *madame Curie Biosci. Database* [Internet], Austin Landes Biosci. 2000–2013.
- [4] D. Diodato, D. Ghezzi, V. Tiranti, The mitochondrial aminoacyl tRNA synthetases: genes and syndromes, *Int. J. Cell Biol.* 2014 (2014) 787956.
- [5] J.L. Huot, L. Enkler, C. Megel, L. Karim, D. Laporte, H.D. Becker, A.-M. Duchêne, M. Sissler, L. Maréchal-Drouard, Idiosyncrasies in decoding mitochondrial genomes, *Biochimie* 100 (2014) 95–106.
- [6] J.R. Brown, D. Gentry, J.A. Becker, K. Ingraham, D.J. Holmes, M.J. Stanhope, Horizontal transfer of drug-resistant aminoacyl-transfer-RNA synthetases of anthrax and Gram-positive pathogens, *EMBO Rep.* 4 (2003) 692–698.
- [7] B. Brindéfalk, J. Viklund, D. Larsson, M. Thollesson, S.G.E. Andersson, Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases, *Mol. Biol. Evol.* 24 (2007) 743–756.
- [8] A.-M. Duchêne, A. Giritch, B. Hoffmann, V. Cognat, D. Lancelin, N.M. Peeters, M. Zaepfel, L. Maréchal-Drouard, I.D. Small, Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*, *Proc. Natl. Acad. Sci. U.S.A.* 102 (2005) 16484–16489.
- [9] C. Pujol, M. Bailly, D. Kern, L. Marechal-Drouard, H. Becker, A.-M. Duchene, Dual-targeted tRNA-dependent amidotransferase ensures both mitochondrial and chloroplastic Gln-tRNA_{Gln} synthesis in plants, *Proc. Natl. Acad. Sci. U.S.A.* 105 (2008) 6481–6485.
- [10] Y. Hirakawa, F. Burki, P.J. Keeling, Dual targeting of aminoacyl-tRNA synthetases to the mitochondrion and complex plastid in chlorarachniophytes, *J. Cell Sci.* 125 (2012) 6176–6184.
- [11] G.H. Gile, D. Moog, C.H. Slamovits, U.-G. Maier, J.M. Archibald, Dual organellar targeting of aminoacyl-tRNA synthetases in diatoms and cryptophytes, *Genome Biol. Evol.* 7 (2015) 1728–1742.
- [12] G. Natsoulis, F. Hilger, G.R. Fink, The HTS1 gene encodes both the cytoplasmic and mitochondrial histidine tRNA synthetases of *S. cerevisiae*, *Cell* 46 (1986) 235–243.
- [13] B. Chatton, P. Walter, J.P. Ebel, F. Lacroute, F. Fasiolo, The yeast VAS1 gene encodes both mitochondrial and cytoplasmic valyl-tRNA synthetases, *J. Biol. Chem.* 263 (1988) 52–57.
- [14] E. Tokkunova, H. Park, J. Xia, M.P. King, E. Davidson, 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 (2000) 35063–35069.
- [15] J. Alexandrova, C. Paulus, J. Rudinger-Thirion, F. Jossinet, M. Frugier, Elaborate uORF/IRE features control expression and localization of human glycyl-tRNA synthetase, *RNA Biol.* 6286 (2015) 1301–1313.
- [16] H.-L. Tang, L.-S. Yeh, N.-K. Chen, T. Ripmaster, P. Schimmel, C.-C. Wang, Translation of a yeast mitochondrial tRNA synthetase initiated at redundant non-AUG Codons, *J. Biol. Chem.* 279 (2004) 49656–49663.
- [17] C.-I. Chien, Y.-W. Chen, Y.-H. Wu, C.-Y. Chang, T.-L. Wang, C.-C. Wang, Functional substitution of a eukaryotic glycyl-tRNA synthetase with an evolutionarily unrelated bacterial cognate enzyme, *PLoS One* 9 (2014) e94659.
- [18] Y. Ofir-Birin, P. Fang, S.P. Bennett, H.-M. Zhang, J. Wang, I. Rachmin, R. Shapiro, J. Song, A. Dagan, J. Pozo, S. Kim, A.G. Marshall, P. Schimmel, X.-L. Yang, H. Nechushtan, E. Razin, M. Guo, Structural switch of lysyl-tRNA synthetase between translation and transcription, *Mol. Cell.* 49 (2013) 30–42.
- [19] N. Gunasekera, S.W. Lee, S. Kim, K. Musier-Forsyth, E. Arriaga, Nuclear localization of aminoacyl-tRNA synthetases using single-cell capillary electrophoresis laser-induced fluorescence analysis, *Anal. Chem.* 76 (2004) 4741–4746.
- [20] Y.G. Ko, Y.S. Kang, E.K. Kim, S.G. Park, S. Kim, Nucleolar localization of human methionyl-tRNA synthetase and its role in ribosomal RNA synthesis, *J. Cell Biol.* 149 (2000) 567–574.
- [21] M. Mirande, D. Le Corre, D. Louvard, H. Reggio, J.-P. Pailliez, J.-P. Waller, Association of an aminoacyl-tRNA synthetase complex and of phenylalanyl-tRNA synthetase with the cytoskeletal framework fraction from mammalian cells, *Exp. Cell Res.* 156 (1985) 91–102.
- [22] M. Sajish, Q. Zhou, S. Kishi, D.M. Valdez, M. Kapoor, M. Guo, S. Lee, S. Kim, X.-L. Yang, P. Schimmel, Trp-tRNA synthetase bridges DNA-PKcs to PARP-1 to link IFN- γ and p53 signaling, *Nat. Chem. Biol.* 8 (2012) 547–554.
- [23] V.I. Popenko, N.E. Cherny, S.F. Beresten, J.L. Ivanova, V.V. Filonenko, L.L. Kisselev, Immunoelectron microscopic location of tryptophanyl-tRNA synthetase in mammalian, prokaryotic and archaeobacterial cells, *Eur. J. Cell Biol.* 62 (1993) 248–258.
- [24] E.L. Paley, V.N. Baranov, N.M. Alexandrova, L.L. Kisselev, Tryptophanyl-tRNA synthetase in cell lines resistant to tryptophan analogs, *Exp. Cell Res.* 195 (1991) 66–78.
- [25] G. Fu, T. Xu, Y. Shi, N. Wei, X.-L. Yang, TRNA-controlled nuclear import of a human tRNA synthetase, *J. Biol. Chem.* 287 (2012) 9330–9334.
- [26] N. Wei, Y. Shi, L.N. Truong, K.M. Fisch, T. Xu, E. Gardiner, G. Fu, Y.-S.O. Hsu, S. Kishi, A.I. Su, X. Wu, X.-L. Yang, Oxidative stress diverts tRNA synthetase to nucleus for protection against DNA damage, *Mol. Cell.* 56 (2014) 323–332.
- [27] H.-Y. Huang, H.-L. Tang, H.-Y. Chao, L.-S. Yeh, C.-C. Wang, An unusual pattern of protein expression and localization of yeast alanyl-tRNA synthetase isoforms, *Mol. Microbiol.* 60 (2006) 189–198.
- [28] M. Frechin, B. Senger, M. Braye, D. Kern, R.P. Martin, H.D. Becker, Yeast mitochondrial Gln-tRNA_{Gln} is generated by a GatFAB-mediated transamidation pathway involving Arc1p-controlled subcellular sorting of cytosolic GluRS, *Genes Dev.* 23 (2009) 1119–1130.
- [29] M. Frechin, L. Enkler, E. Tetaud, D. Laporte, B. Senger, C. Blancard, P. Hammann, G. Bader, S. Clauder-Münster, L.M. Steinmetz, R.P. Martin, J.-P. di Rago, H.D. Becker, Expression of nuclear and mitochondrial genes encoding ATP synthase is synchronized by disassembly of a multisynthetase complex, *Mol. Cell.* 56 (2014) 763–776.
- [30] K.-J. Chang, C.-C. Wang, Translation initiation from a naturally occurring non-AUG codon in *Saccharomyces cerevisiae*, *J. Biol. Chem.* 279 (2004) 13778–13785.
- [31] R.J. Turner, M. Lovato, P. Schimmel, One of two genes encoding glycyl-tRNA synthetase in *Saccharomyces cerevisiae* provides mitochondrial and cytoplasmic functions, *J. Biol. Chem.* 275 (2000) 27681–27688.
- [32] M.I. Chiu, T.L. Mason, G.R. Fink, HTS1 encodes both the cytoplasmic and mitochondrial histidyl-tRNA synthetase of *Saccharomyces cerevisiae*: mutations alter the specificity of compartmentation, *Genetics* 132 (1992) 987–1001.
- [33] C.-C. Wang, K.-J. Chang, H.-L. Tang, C.-J. Hsieh, P. Schimmel, Mitochondrial form of a tRNA synthetase can be made bifunctional by manipulating its leader peptide, *Biochemistry* 42 (2003) 1646–1651.
- [34] K. Galani, H. Grosshans, K. Deinert, A.C. Hurt, G. Simos, The intracellular location of two aminoacyl-tRNA synthetases depends on complex formation with Arc1p, *EMBO J.* 20 (2001) 6889–6898.
- [35] A.K. Azad, D.R. Stanford, S. Sarkar, A.K. Hopper, Role of nuclear pools of aminoacyl-tRNA synthetases in tRNA nuclear export, *Mol. Biol. Cell.* 12 (2001) 1381–1392.
- [36] J. Rettig, Y. Wang, A. Schneider, T. Ochsenreiter, Dual targeting of isoleucyl-tRNA synthetase in *Trypanosoma brucei* is mediated through alternative trans-splicing, *Nucleic Acids Res.* 40 (2012) 1299–1306.
- [37] I. Cestari, S. Kalidas, S. Monnerat, A. Anupama, M.A. Phillips, K. Stuart, A multiple aminoacyl-tRNA synthetase complex that enhances tRNA-aminoacylation in African trypanosomes, *Mol. Cell Biol.* 33 (2013) 4872–4888.
- [38] J. Rinehart, E.K. Horn, D. Wei, D. Soll, A. Schneider, Non-canonical eukaryotic glutaminyl- and glutamyl-tRNA synthetases form mitochondrial aminoacyl-tRNA in *Trypanosoma brucei*, *J. Biol. Chem.* 279 (2004) 1161–1166.
- [39] K.E. Jackson, J.S. Pham, M. Kwek, N.S. De Silva, S.M. Allen, C.D. Goodman, G.I. McFadden, L. Ribas de Pouplana, S.A. Ralph, Dual targeting of aminoacyl-tRNA synthetases to the apicoplast and cytosol in *Plasmodium falciparum*, *Int. J. Parasitol.* 42 (2012) 177–186.
- [40] J.S. Pham, R. Sakaguchi, L.M. Yeoh, N.S. De Silva, G.I. McFadden, Y.-M. Hou, S.A. Ralph, A dual-targeted aminoacyl-tRNA synthetase in *Plasmodium falciparum* charges cytosolic and apicoplast tRNACys, *Biochem. J.* 458 (2014) 513–523.
- [41] M. Frechin, A.-M. Duchêne, H.D. Becker, Translating organellar glutamine codons: a case by case scenario?, *RNA Biol.* 6 (2009) 31–34.
- [42] P. López-García, D. Moreira, Open questions on the origin of eukaryotes, *Trends Ecol. Evol.* 30 (2015) 697–708.
- [43] A. Schön, C.G. Kannangara, S. Cough, D. Söll, Protein biosynthesis in organelles requires misaminoacylation of tRNA, *Nature* 331 (1988) 187–190.
- [44] A. Nagao, T. Suzuki, T. Katoh, Y. Sakaguchi, T. Suzuki, Biogenesis of glutaminyl-tRNA_{Gln} in human mitochondria, *Proc. Natl. Acad. Sci. U.S.A.* 106 (2009) 16209–16214.
- [45] Y. Arais, J.L. Huot, T. Sekiguchi, M. Frechin, F. Fischer, L. Enkler, B. Senger, R. Ishitani, H.D. Becker, O. Nureki, Crystal structure of *Saccharomyces cerevisiae* mitochondrial GatFAB reveals a novel subunit assembly in tRNA-dependent amidotransferases, *Nucleic Acids Res.* 42 (2014) 6052–6063.
- [46] S. Merz, B. Westermann, Genome-wide deletion mutant analysis reveals genes required for respiratory growth, mitochondrial genome maintenance and mitochondrial protein synthesis in *Saccharomyces cerevisiae*, *Genome Biol.* 10 (2009) R95.
- [47] S. Hati, B. Ziervogel, J. Sternjohn, F.-C. Wong, M.C. Nagan, A.E. Rosen, P.G. Siliciano, J.W. Chihade, K. Musier-Forsyth, Pre-transfer editing by class II prolyl-tRNA synthetase: role of aminoacylation active site in “selective release” of noncognate amino acids *, *J. Biochem.* 281 (2006) 27862–27872.
- [48] C. Meisinger, T. Sommer, N. Pfanner, Purification of *saccharomyces cerevisiae* mitochondria devoid of microsomal and cytosolic contaminations, *Anal. Biochem.* 287 (2000) 339–342.
- [49] E. Lund, J.E. Dahlberg, Proofreading and aminoacylation of tRNAs before export from the nucleus, *Science* 282 (1998) 2082–2085.
- [50] S. Sarkar, A.K. Azad, A.K. Hopper, Nuclear tRNA aminoacylation and its role in nuclear export of endogenous tRNAs in *Saccharomyces cerevisiae*, *Proc. Natl. Acad. Sci. U.S.A.* 96 (1999) 14366–14371.
- [51] J. Dostie, F. Lejbkovic, N. Sonenberg, Nuclear eukaryotic initiation factor 4E (eIF4E) colocalizes with splicing factors in speckles, *J. Cell Biol.* 148 (2000) 239–247.
- [52] P. Schimmel, C.-C. Wang, Getting tRNA synthetases into the nucleus, *Trends Biochem. Sci.* 24 (1999) 127–128.
- [53] A.K. Bandyopadhyay, M.P. Deutscher, Complex of aminoacyl-transfer RNA synthetases, *J. Mol. Biol.* 60 (1971) 113–122.
- [54] P. Kerjan, C. Cerini, M. Sémériva, M. Mirande, The multienzyme complex containing nine aminoacyl-tRNA synthetases is ubiquitous from *Drosophila* to mammals, *Biochim. Biophys. Acta* 1199 (1994) 293–297.

- [55] G. Simos, A. Segref, F. Fasiolo, K. Hellmuth, A. Shevchenko, M. Mann, E.C. Hurt, The yeast protein Arc1p binds to tRNA and functions as a cofactor for the methionyl- and glutamyl-tRNA synthetases, *EMBO J.* 15 (1996) 5437–5448.
- [56] S. Quevillon, M. Mirande, The p18 component of the multisynthetase complex shares a protein motif with the β and γ subunits of eukaryotic elongation factor 1, *FEBS Lett.* 395 (1996) 63–67.
- [57] S. Quevillon, J.-C. Robinson, E. Berthonneau, M. Siatecka, M. Mirande, Macromolecular assemblage of aminoacyl-tRNA synthetases: identification of protein-protein interactions and characterization of a core protein, *J. Mol. Biol.* 285 (1999) 183–195.
- [58] S.S. Kim, S.Y. Hur, Y.R. Kim, N.J. Yoo, S.H. Lee, Expression of AIMP1, 2 and 3, the scaffolds for the multi-tRNA synthetase complex, is downregulated in gastric and colorectal cancer, *Tumori J.* 97 (2011) 380–385.
- [59] D. Laporte, J.L. Huot, G. Bader, L. Enkler, B. Senger, H.D. Becker, Exploring the evolutionary diversity and assembly modes of multi-aminoacyl-tRNA synthetase complexes: Lessons from unicellular organisms, *FEBS Lett.* 588 (2014) 4268–4278.
- [60] B.S. Negrutskii, R. Stapulionis, M.P. Deuschert, Supramolecular organization of the mammalian translation system, *Biochemistry* 91 (1994) 964–968.
- [61] P.S. Ray, A. Arif, P.L. Fox, Macromolecular complexes as depots for releasable regulatory proteins, *Trends Biochem. Sci.* 32 (2007) 158–164.
- [62] C.B. Saper, A guide to the perplexed on the specificity of antibodies, *J. Histochem. Cytochem.* 57 (2009) 1–5.
- [63] J.E. Gilda, R. Ghosh, J.X. Cheah, T.M. West, S.C. Bodine, A.V. Gomes, Western blotting inaccuracies with unverified antibodies: need for a western blotting minimal reporting standard (WBMRS), *PLoS One* 10 (2015) e0135392.
- [64] S.E. Rieder, S.D. Emr, Overview of subcellular fractionation procedures for the yeast *Saccharomyces cerevisiae*, *Curr. Protoc. Cell Biol.* Chapter 3 (2001) Unit 3.7.
- [65] E. Palmer, T. Freeman, Investigation into the use of C- and N-terminal GFP fusion proteins for subcellular localization studies using reverse transfection microarrays, *Comp. Funct. Genomics* 5 (2004) 342–353.
- [66] M. Kaminska, S. Havrylenko, P. Decottignies, P. Le Maréchal, B. Negrutskii, M. Mirande, Dynamic organization of aminoacyl-tRNA synthetase complexes in the cytoplasm of human cells, *J. Biol. Chem.* 284 (2009) 13746–13754.
- [67] J. Rinehart, B. Krett, M.A.T. Rubio, J.D. Alfonso, D. Söll, *Saccharomyces cerevisiae* imports the cytosolic pathway for Gln-tRNA synthesis into the mitochondrion, *Genes Dev.* 19 (2005) 583–592.
- [68] P. Dönnies, A. Höglund, Predicting protein subcellular localization: past, present, and future, *Genomics Proteomics Bioinformatics* 2 (2004) 209–215.
- [69] K.-C. Chou, H.-B. Shen, A new method for predicting the subcellular localization of eukaryotic proteins with both single and multiple sites: Euk-mPloc 2.0, *PLoS One* 5 (2010) e9931.
- [70] C. Savojardo, P.L. Martelli, P. Fariselli, R. Casadio, TPpred2: improving the prediction of mitochondrial targeting peptide cleavage sites by exploiting sequence motifs, *Bioinformatics* 30 (2014) 2973–2974.
- [71] O. Emanuelsson, H. Nielsen, S. Brunak, G. von Heijne, Predicting subcellular localization of proteins based on their N-terminal amino acid sequence, *J. Mol. Biol.* 300 (2000) 1005–1016.
- [72] H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne, Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites, *Protein Eng.* 10 (1997) 1–6.
- [73] Y. Fukasawa, J. Tsuji, S.-C. Fu, K. Tomii, P. Horton, K. Imai, MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites, *Mol. Cell. Proteomics* 14 (2015) 1113–1126.
- [74] A.N. Nguyen Ba, A. Pogoutse, N. Provart, A.M. Moses, NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction, *BMC Bioinformatics* 10 (2009) 202.
- [75] J. Lin, J. Hu, SeqNLS: nuclear localization signal prediction based on frequent pattern mining and linear motif scoring, *PLoS One* 8 (2013) e76864.
- [76] S. Kosugi, M. Hasebe, M. Tomita, H. Yanagawa, Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs, *Proc. Natl. Acad. Sci. U.S.A.* 106 (2009) 10171–10176.
- [77] K. Nakai, P. Horton, PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization, *Trends Biochem. Sci.* 24 (1999) 34–36.
- [78] C.J. Bonangelino, N.L. Catlett, L.S. Weisman, Vac7p, a novel vacuolar protein, is required for normal vacuole inheritance and morphology, *Mol. Cell. Biol.* 17 (1997) 6847–6858.
- [79] S.E. Rieder, S.D. Emr, Isolation of subcellular fractions from the yeast *Saccharomyces cerevisiae*, *Curr. Protoc. Cell Biol.* Chapter 3 (2001) Unit 3.8.
- [80] J.M. Dahlman, D.C. Guttridge, Detection of NF- κ B activity in skeletal muscle cells by electrophoretic mobility shift analysis, *Methods Mol. Biol.* 798 (2012) 505–516.
- [81] E. Olmedo-Verd, J. Santamaría-Gómez, J.A.G. Ochoa de Alda, L. Ribas de Pouplana, I. Luque, Membrane anchoring of aminoacyl-tRNA synthetases by convergent acquisition of a novel protein domain, *J. Biol. Chem.* 286 (2011) 41057–41068.
- [82] T.A. Vida, S.D. Emr, A new vital stain for visualizing vacuolar membrane dynamics and endocytosis in yeast, *J. Cell Biol.* 128 (1995) 779–792.
- [83] W. He, G. Bai, H. Zhou, N. Wei, N.M. White, J. Lauer, H. Liu, Y. Shi, C.D. Dumitru, K. Lettieri, V. Shubayev, A. Jordanova, V. Guerguelcheva, P.R. Griffin, R.W. Burgess, S.L. Pfaff, X.-L. Yang, CMT2D neuropathy is linked to the neomorphic binding activity of glycyl-tRNA synthetase, *Nature* 526 (2015) 710–714.
- [84] M.C. Park, T. Kang, D. Jin, J.M. Han, S.B. Kim, Y.J. Park, K. Cho, Y.W. Park, M. Guo, W. He, X.-L. Yang, P. Schimmel, S. Kim, Secreted human glycyl-tRNA synthetase implicated in defense against ERK-activated tumorigenesis, *Proc. Natl. Acad. Sci. U.S.A.* 109 (2012) E640–E647.
- [85] Z. Wei, Z. Xu, X. Liu, W.-S. Lo, F. Ye, C.-F. Lau, F. Wang, J.J. Zhou, L.A. Nangle, X.-L. Yang, M. Zhang, P. Schimmel, Alternative splicing creates two new architectures for human tyrosyl-tRNA synthetase, *Nucleic Acids Res.* 44 (2016) 1247–1255.

Bibliographie

- Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nishikawa, S., Endo, T., and Kohda, D. (2000). Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* *100*, 551-560.
- Aebi, M., Kirchner, G., Chen, J.Y., Vijayraghavan, U., Jacobson, A., Martin, N.C., and Abelson, J. (1990). Isolation of a temperature-sensitive mutant with an altered tRNA nucleotidyltransferase and cloning of the gene encoding tRNA nucleotidyltransferase in the yeast *Saccharomyces cerevisiae*. *The Journal of biological chemistry* *265*, 16216-16220.
- Ahel, I., Korencic, D., Ibba, M., and Soll, D. (2003). Trans-editing of mischarged tRNAs. *Proceedings of the National Academy of Sciences of the United States of America* *100*, 15422-15427.
- Ahn, H.C., Kim, S., and Lee, B.J. (2003). Solution structure and p43 binding of the p38 leucine zipper motif: coiled-coil interactions mediate the association between p38 and p43. *FEBS letters* *542*, 119-124.
- Ai, H.W., Henderson, J.N., Remington, S.J., and Campbell, R.E. (2006). Directed evolution of a monomeric, bright and photostable version of *Clavularia cyan* fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochem J* *400*, 531-540.
- Akins, R.A., and Lambowitz, A.M. (1987). A protein required for splicing group I introns in *Neurospora* mitochondria is mitochondrial tyrosyl-tRNA synthetase or a derivative thereof. *Cell* *50*, 331-345.
- Alexandrova, J., Paulus, C., Rudinger-Thirion, J., Jossinet, F., and Frugier, M. (2015). Elaborate uORF/IRES features control expression and localization of human glycyl-tRNA synthetase. *RNA Biol* *12*, 1301-1313.
- Algire, M.A., Maag, D., Savio, P., Acker, M.G., Tarun, S.Z., Jr., Sachs, A.B., Asano, K., Nielsen, K.H., Olsen, D.S., Phan, L., *et al.* (2002). Development and characterization of a reconstituted yeast translation initiation system. *Rna* *8*, 382-397.
- Alkalaeva, E.Z., Pisarev, A.V., Frolova, L.Y., Kisselev, L.L., and Pestova, T.V. (2006). In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. *Cell* *125*, 1125-1136.
- Altman, R. (1890). *Die Elementarorganismen und ihre Beziehungen zu den Zellen*. In Leipzig, Veit & Comp, pp. pp. 28 and 145.
- An, S., and Musier-Forsyth, K. (2005). Cys-tRNA(Pro) editing by *Haemophilus influenzae* YbaK via a novel synthetase.YbaK.tRNA ternary complex. *The Journal of biological chemistry* *280*, 34465-34472.
- Anand, M., Balar, B., Ulloque, R., Gross, S.R., and Kinzy, T.G. (2006). Domain and nucleotide dependence of the interaction between *Saccharomyces cerevisiae* translation elongation factors 3 and 1A. *The Journal of biological chemistry* *281*, 32318-32326.
- Anand, M., Chakraborty, K., Marton, M.J., Hinnebusch, A.G., and Kinzy, T.G. (2003). Functional interactions between yeast translation eukaryotic elongation factor (eEF) 1A and eEF3. *The Journal of biological chemistry* *278*, 6985-6991.
- Andersen, C.B., Becker, T., Blau, M., Anand, M., Halic, M., Balar, B., Mielke, T., Boesen, T., Pedersen, J.S., Spahn, C.M., *et al.* (2006). Structure of eEF3 and the mechanism of transfer RNA release from the E-site. *Nature* *443*, 663-668.
- Anderson, J., Phan, L., Cuesta, R., Carlson, B.A., Pak, M., Asano, K., Bjork, G.R., Tamame, M., and Hinnebusch, A.G. (1998). The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes & development* *12*, 3650-3662.

- Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., *et al.* (1981). Sequence and organization of the human mitochondrial genome. *Nature* *290*, 457-465.
- Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H., and Kurland, C.G. (1998). The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* *396*, 133-140.
- Araiso, Y., Huot, J.L., Sekiguchi, T., Frechin, M., Fischer, F., Enkler, L., Senger, B., Ishitani, R., Becker, H.D., and Nureki, O. (2014). Crystal structure of *Saccharomyces cerevisiae* mitochondrial GatFAB reveals a novel subunit assembly in tRNA-dependent amidotransferases. *Nucleic acids research* *42*, 6052-6063.
- Arif, A., Jia, J., Moodt, R.A., DiCorleto, P.E., and Fox, P.L. (2011). Phosphorylation of glutamyl-prolyl tRNA synthetase by cyclin-dependent kinase 5 dictates transcript-selective translational control. *Proceedings of the National Academy of Sciences of the United States of America* *108*, 1415-1420.
- 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. *Molecular cell* *35*, 164-180.
- Arif, A., Terenzi, F., Potdar, A.A., Jia, J., Sacks, J., China, A., Halawani, D., Vasu, K., Li, X., Brown, J.M., *et al.* (2017). EPRS is a critical mTORC1-S6K1 effector that influences adiposity in mice. *Nature* *542*, 357-361.
- Asano, K., Phan, L., Valasek, L., Schoenfeld, L.W., Shalev, A., Clayton, J., Nielsen, K., Donahue, T.F., and Hinnebusch, A.G. (2001). A multifactor complex of eIF1, eIF2, eIF3, eIF5, and tRNA(i)Met promotes initiation complex assembly and couples GTP hydrolysis to AUG recognition. *Cold Spring Harb Symp Quant Biol* *66*, 403-415.
- Axelrod, D., Koppel, D.E., Schlessinger, J., Elson, E., and Webb, W.W. (1976). Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys J* *16*, 1055-1069.
- Azad, A.K., Stanford, D.R., Sarkar, S., and Hopper, A.K. (2001). Role of nuclear pools of aminoacyl-tRNA synthetases in tRNA nuclear export. *Molecular biology of the cell* *12*, 1381-1392.
- Bacia, K., Kim, S.A., and Schwille, P. (2006). Fluorescence cross-correlation spectroscopy in living cells. *Nat Methods* *3*, 83-89.
- Bader, G., Debard, S., De Craene, J.O., Enkler, L., Bar, S., Laporte, D., Hammann, P., Myslinski, E., Senger, B., Friant, S., *et al.* (2017). Nonconventional localizations of cytosolic aminoacyl-tRNA synthetases in yeast and human cells. *Methods* *113*, 91-104.
- Baehler, P.J., Biondi, R.M., van Bemmelen, M., Veron, M., and Reymond, C.D. (2002). Random insertion of green fluorescent protein into the regulatory subunit of cyclic adenosine monophosphate-dependent protein kinase. *Methods Mol Biol* *183*, 57-68.
- Bailly, M., Blaise, M., Lorber, B., Becker, H.D., and Kern, D. (2007). The transamidosome: a dynamic ribonucleoprotein particle dedicated to prokaryotic tRNA-dependent asparagine biosynthesis. *Molecular cell* *28*, 228-239.
- Bakowska-Zywicka, K., Kasprzyk, M., and Twardowski, T. (2016a). tRNA-derived short RNAs bind to *Saccharomyces cerevisiae* ribosomes in a stress-dependent manner and inhibit protein synthesis in vitro. *FEMS yeast research* *16*.
- Bakowska-Zywicka, K., Mleczko, A.M., Kasprzyk, M., Machtel, P., Zywicki, M., and Twardowski, T. (2016b). The widespread occurrence of tRNA-derived fragments in *Saccharomyces cerevisiae*. *FEBS Open Bio* *6*, 1186-1200.
- Balaban, R.S., Nemoto, S., and Finkel, T. (2005). Mitochondria, oxidants, and aging. *Cell* *120*, 483-495.

- Baleva, M., Gowher, A., Kamenski, P., Tarassov, I., Entelis, N., and Masquida, B. (2015). A Moonlighting Human Protein Is Involved in Mitochondrial Import of tRNA. *Int J Mol Sci* *16*, 9354-9367.
- Bandyopadhyay, A.K., and Deutscher, M.P. (1971). Complex of aminoacyl-transfer RNA synthetases. *Journal of molecular biology* *60*, 113-122.
- Banerjee, R., Gladkova, C., Mapa, K., Witte, G., and Mokranjac, D. (2015). Protein translocation channel of mitochondrial inner membrane and matrix-exposed import motor communicate via two-domain coupling protein. *Journal Article* *4*, e11897.
- Barbarese, E., Koppel, D.E., Deutscher, M.P., Smith, C.L., Ainger, K., Morgan, F., and Carson, J.H. (1995). Protein translation components are colocalized in granules in oligodendrocytes. *Journal of cell science* *108 (Pt 8)*, 2781-2790.
- Barrientos, A., Zambrano, A., and Tzagoloff, A. (2004). Mss51p and Cox14p jointly regulate mitochondrial Cox1p expression in *Saccharomyces cerevisiae*. *The EMBO journal* *23*, 3472-3482.
- Barthelme, D., Dinkelaker, S., Albers, S.V., Londei, P., Ermler, U., and Tampe, R. (2011). Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1. *Proceedings of the National Academy of Sciences of the United States of America* *108*, 3228-3233.
- Baudendistel, N., Muller, G., Waldeck, W., Angel, P., and Langowski, J. (2005). Two-hybrid fluorescence cross-correlation spectroscopy detects protein-protein interactions in vivo. *Chemphyschem* *6*, 984-990.
- Baum, D.A., and Baum, B. (2014). An inside-out origin for the eukaryotic cell. *BMC Biol* *12*, 76.
- Beach, D.L., Salmon, E.D., and Bloom, K. (1999). Localization and anchoring of mRNA in budding yeast. *Current biology* : *CB* *9*, 569-578.
- Becker, H.D., and Kern, D. (1998). *Thermus thermophilus*: a link in evolution of the tRNA-dependent amino acid amidation pathways. *Proceedings of the National Academy of Sciences of the United States of America* *95*, 12832-12837.
- Becker, H.D., Min, B., Jacobi, C., Raczniak, G., Pelaschier, J., Roy, H., Klein, S., Kern, D., and Soll, D. (2000). The heterotrimeric *Thermus thermophilus* Asp-tRNA(Asn) amidotransferase can also generate Gln-tRNA(Gln). *FEBS letters* *476*, 140-144.
- Becker, L., Bannwarth, M., Meisinger, C., Hill, K., Model, K., Krimmer, T., Casadio, R., Truscott, K.N., Schulz, G.E., Pfanner, N., *et al.* (2005). Preprotein translocase of the outer mitochondrial membrane: reconstituted Tom40 forms a characteristic TOM pore. *Journal of molecular biology* *353*, 1011-1020.
- Becker, T., Wenz, L.S., Kruger, V., Lehmann, W., Muller, J.M., Goroncy, L., Zufall, N., Lithgow, T., Guiard, B., Chacinska, A., *et al.* (2011). The mitochondrial import protein Mim1 promotes biogenesis of multispanning outer membrane proteins. *The Journal of cell biology* *194*, 387-395.
- Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011). The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* *334*, 1524-1529.
- Ben-Shem, A., Jenner, L., Yusupova, G., and Yusupov, M. (2010). Crystal structure of the eukaryotic ribosome. *Science* *330*, 1203-1209.
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA particles in living yeast. *Molecular cell* *2*, 437-445.
- Bhangoo, M.K., Tzankov, S., Fan, A.C., Deigaard, K., Thomas, D.Y., and Young, J.C. (2007). Multiple 40-kDa heat-shock protein chaperones function in Tom70-dependent mitochondrial import. *Molecular biology of the cell* *18*, 3414-3428.

- Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., *et al.* (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453-1462.
- Bockler, S., and Westermann, B. (2014). ER-mitochondria contacts as sites of mitophagosome formation. *Autophagy* 10, 1346-1347.
- Bonfils, G., Jaquenoud, M., Bontron, S., Ostrowicz, C., Ungermann, C., and De Virgilio, C. (2012). Leucyl-tRNA synthetase controls TORC1 via the EGO complex. *Molecular cell* 46, 105-110.
- Bonnefoy, N., and Fox, T.D. (2007). Directed alteration of *Saccharomyces cerevisiae* mitochondrial DNA by biolistic transformation and homologous recombination. *Methods Mol Biol* 372, 153-166.
- Brejc, K., Sixma, T.K., Kitts, P.A., Kain, S.R., Tsien, R.Y., Ormo, M., and Remington, S.J. (1997). Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* 94, 2306-2311.
- Brix, J., Rudiger, S., Bukau, B., Schneider-Mergener, J., and Pfanner, N. (1999). Distribution of binding sequences for the mitochondrial import receptors Tom20, Tom22, and Tom70 in a presequence-carrying preprotein and a non-cleavable preprotein. *The Journal of biological chemistry* 274, 16522-16530.
- Brunel, C., Romby, P., Moine, H., Caillet, J., Grunberg-Manago, M., Springer, M., Ehresmann, B., and Ehresmann, C. (1993). Translational regulation of the *Escherichia coli* threonyl-tRNA synthetase gene: structural and functional importance of the thrS operator domains. *Biochimie* 75, 1167-1179.
- Buddha, M.R., Keery, K.M., and Crane, B.R. (2004a). An unusual tryptophanyl tRNA synthetase interacts with nitric oxide synthase in *Deinococcus radiodurans*. *Proceedings of the National Academy of Sciences of the United States of America* 101, 15881-15886.
- Buddha, M.R., Tao, T., Parry, R.J., and Crane, B.R. (2004b). Regioselective nitration of tryptophan by a complex between bacterial nitric-oxide synthase and tryptophanyl-tRNA synthetase. *The Journal of biological chemistry* 279, 49567-49570.
- Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D., *et al.* (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273, 1058-1073.
- Burgess, H.M., Richardson, W.A., Anderson, R.C., Salaun, C., Graham, S.V., and Gray, N.K. (2011). Nuclear relocalisation of cytoplasmic poly(A)-binding proteins PABP1 and PABP4 in response to UV irradiation reveals mRNA-dependent export of metazoan PABPs. *Journal of cell science* 124, 3344-3355.
- Cabantous, S., Nguyen, H.B., Pedelacq, J.D., Koraichi, F., Chaudhary, A., Ganguly, K., Lockard, M.A., Favre, G., Terwilliger, T.C., and Waldo, G.S. (2013). A new protein-protein interaction sensor based on tripartite split-GFP association. *Sci Rep* 3, 2854.
- Cabantous, S., Pedelacq, J.D., Mark, B.L., Naranjo, C., Terwilliger, T.C., and Waldo, G.S. (2005a). Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from *Mycobacterium tuberculosis*. *J Struct Funct Genomics* 6, 113-119.
- Cabantous, S., Terwilliger, T.C., and Waldo, G.S. (2005b). Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nature biotechnology* 23, 102-107.
- Causier, B., and Davies, B. (2002). Analysing protein-protein interactions with the yeast two-hybrid system. *Plant Mol Biol* 50, 855-870.

- Cavalier-Smith, T. (1975). Electron and light microscopy of gametogenesis and gamete fusion in *Chlamydomonas reinhardtii*. *Protoplasma* *86*, 1-18.
- Cavener, D.R., and Ray, S.C. (1991). Eukaryotic start and stop translation sites. *Nucleic acids research* *19*, 3185-3192.
- Cestari, I., Kalidas, S., Monnerat, S., Anupama, A., Phillips, M.A., and Stuart, K. (2013). A multiple aminoacyl-tRNA synthetase complex that enhances tRNA-aminoacylation in African trypanosomes. *Molecular and cellular biology* *33*, 4872-4888.
- Chacinska, A., Pfannschmidt, S., Wiedemann, N., Kozjak, V., Sanjuan Szklarz, L.K., Schulze-Specking, A., Truscott, K.N., Guiard, B., Meisinger, C., and Pfanner, N. (2004). Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. *The EMBO journal* *23*, 3735-3746.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994). Green fluorescent protein as a marker for gene expression. *Science* *263*, 802-805.
- Chamberlain, J.R., Lee, Y., Lane, W.S., and Engelke, D.R. (1998). Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP. *Genes & development* *12*, 1678-1690.
- Chang, K.J., and Wang, C.C. (2004). Translation initiation from a naturally occurring non-AUG codon in *Saccharomyces cerevisiae*. *The Journal of biological chemistry* *279*, 13778-13785.
- Chang, K.M., and Hendrickson, T.L. (2009). Recognition of tRNA^{Gln} by *Helicobacter pylori* GluRS2--a tRNA^{Gln}-specific glutamyl-tRNA synthetase. *Nucleic acids research* *37*, 6942-6949.
- Chatton, B., Walter, P., Ebel, J.P., Lacroute, F., and Fasiolo, F. (1988). The yeast VAS1 gene encodes both mitochondrial and cytoplasmic valyl-tRNA synthetases. *The Journal of biological chemistry* *263*, 52-57.
- Chen, Q., Vazquez, E.J., Moghaddas, S., Hoppel, C.L., and Lesnefsky, E.J. (2003). Production of reactive oxygen species by mitochondria: central role of complex III. *The Journal of biological chemistry* *278*, 36027-36031.
- Chen, Y., Beck, A., Davenport, C., Chen, Y., Shattuck, D., and Tavtigian, S.V. (2005). Characterization of TRZ1, a yeast homolog of the human candidate prostate cancer susceptibility gene ELAC2 encoding tRNAse Z. *BMC Mol Biol* *6*, 12.
- Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., *et al.* (2012). *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic acids research* *40*, D700-705.
- Chien, C.I., Chen, Y.W., Wu, Y.H., Chang, C.Y., Wang, T.L., and Wang, C.C. (2014). Functional substitution of a eukaryotic glycyl-tRNA synthetase with an evolutionarily unrelated bacterial cognate enzyme. *PLoS one* *9*, e94659.
- Chinault, A.C., Tan, K.H., Hassur, S.M., and Hecht, S.M. (1977). Initial position of aminoacylation of individual *Escherichia coli*, yeast, and calf liver transfer RNAs. *Biochemistry* *16*, 766-776.
- Choi, J.W., Um, J.Y., Kundu, J.K., Surh, Y.J., and Kim, S. (2009). Multidirectional tumor-suppressive activity of AIMP2/p38 and the enhanced susceptibility of AIMP2 heterozygous mice to carcinogenesis. *Carcinogenesis* *30*, 1638-1644.
- Claude, A., and Fullam, E.F. (1945). An Electron Microscope Study of Isolated Mitochondria : Method and Preliminary Results. *J Exp Med* *81*, 51-62.
- Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G., and Ward, W.W. (1993). Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry* *32*, 1212-1218.

- Commans, S., and Bock, A. (1999). Selenocysteine inserting tRNAs: an overview. *FEMS microbiology reviews* 23, 335-351.
- Cooper, K.F., Scarnati, M.S., Krasley, E., Mallory, M.J., Jin, C., Law, M.J., and Strich, R. (2012). Oxidative-stress-induced nuclear to cytoplasmic relocalization is required for Not4-dependent cyclin C destruction. *Journal of cell science* 125, 1015-1026.
- Copela, L.A., Fernandez, C.F., Sherrer, R.L., and Wolin, S.L. (2008). Competition between the Rex1 exonuclease and the La protein affects both Trf4p-mediated RNA quality control and pre-tRNA maturation. *Rna* 14, 1214-1227.
- Corti, O., Hampe, C., Koutnikova, H., Darios, F., Jacquier, S., Prigent, A., Robinson, J.C., Pradier, L., Ruberg, M., Mirande, M., *et al.* (2003). The p38 subunit of the aminoacyl-tRNA synthetase complex is a Parkin substrate: linking protein biosynthesis and neurodegeneration. *Hum Mol Genet* 12, 1427-1437.
- Cramer, A., Whitehorn, E.A., Tate, E., and Stemmer, W.P. (1996). Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nature biotechnology* 14, 315-319.
- Crick, F.H. (1968). The origin of the genetic code. *Journal of molecular biology* 38, 367-379.
- Cubitt, A.B., Heim, R., Adams, S.R., Boyd, A.E., Gross, L.A., and Tsien, R.Y. (1995). Understanding, improving and using green fluorescent proteins. *Trends in biochemical sciences* 20, 448-455.
- Curnow, A.W., Ibba, M., and Soll, D. (1996). tRNA-dependent asparagine formation. *Nature* 382, 589-590.
- Cusack, S., Berthet-Colominas, C., Hartlein, 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 Å. *Nature* 347, 249-255.
- Czirok, A., Zach, J., Kozel, B.A., Mecham, R.P., Davis, E.C., and Rongish, B.J. (2006). Elastic fiber macro-assembly is a hierarchical, cell motion-mediated process. *J Cell Physiol* 207, 97-106.
- Dai, D.F., Chiao, Y.A., Marcinek, D.J., Szeto, H.H., and Rabinovitch, P.S. (2014). Mitochondrial oxidative stress in aging and healthspan. *Longev Healthspan* 3, 6.
- Daigle, N., and Ellenberg, J. (2007). LambdaN-GFP: an RNA reporter system for live-cell imaging. *Nat Methods* 4, 633-636.
- Dalton, T.P., Shertzer, H.G., and Puga, A. (1999). Regulation of gene expression by reactive oxygen. *Annu Rev Pharmacol Toxicol* 39, 67-101.
- de Zamaroczy, M., and Bernardi, G. (1986). The primary structure of the mitochondrial genome of *Saccharomyces cerevisiae*--a review. *Gene* 47, 155-177.
- Deinert, K., Fasiolo, F., Hurt, E.C., and Simos, G. (2001). Arc1p organizes the yeast aminoacyl-tRNA synthetase complex and stabilizes its interaction with the cognate tRNAs. *The Journal of biological chemistry* 276, 6000-6008.
- Denney, R.M. (1977). Detection and partial purification of rapidly sedimenting forms of aminoacyl-transfer ribonucleic acid synthetases from human placenta. *Archives of biochemistry and biophysics* 183, 156-167.
- Dhahbi, J.M., Spindler, S.R., Atamna, H., Boffelli, D., and Martin, D.I. (2014). Deep Sequencing of Serum Small RNAs Identifies Patterns of 5' tRNA Half and YRNA Fragment Expression Associated with Breast Cancer. *Biomark Cancer* 6, 37-47.
- Dias, J., Renault, L., Perez, J., and Mirande, M. (2013). Small-angle X-ray solution scattering study of the multi-aminoacyl-tRNA synthetase complex reveals an elongated and multi-armed particle. *The Journal of biological chemistry* 288, 23979-23989.

- Dimmer, K.S., Papic, D., Schumann, B., Sperl, D., Krumpke, K., Walther, D.M., and Rapaport, D. (2012). A crucial role for Mim2 in the biogenesis of mitochondrial outer membrane proteins. *Journal of cell science* *125*, 3464-3473.
- Dopf, J., and Horiagon, T.M. (1996). Deletion mapping of the *Aequorea victoria* green fluorescent protein. *Gene* *173*, 39-44.
- Drose, S., and Brandt, U. (2008). The mechanism of mitochondrial superoxide production by the cytochrome bc1 complex. *The Journal of biological chemistry* *283*, 21649-21654.
- Duchene, A.M., Giritch, A., Hoffmann, B., Cognat, V., Lancelin, D., Peeters, N.M., Zaepfel, M., Marechal-Drouard, L., and Small, I.D. (2005). Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* *102*, 16484-16489.
- Duchene, A.M., Peeters, N., Dietrich, A., Cosset, A., Small, I.D., and Wintz, H. (2001). Overlapping destinations for two dual targeted glycyl-tRNA synthetases in *Arabidopsis thaliana* and *Phaseolus vulgaris*. *The Journal of biological chemistry* *276*, 15275-15283.
- Dynes, J.L., and Steward, O. (2007). Dynamics of bidirectional transport of Arc mRNA in neuronal dendrites. *J Comp Neurol* *500*, 433-447.
- Ehresmann, B., Moine, H., Romby, P., Springer, M., Grunberg-Manago, M., Ebel, J.P., and Ehresmann, C. (1988). Secondary structure of the *Escherichia coli* translational operator of threonyl-tRNA synthetase and relationship to its function. *Gene* *72*, 187-188.
- Elbaz-Alon, Y., Rosenfeld-Gur, E., Shinder, V., Futerman, Anthony H., Geiger, T., and Schuldiner, M. (2014). A Dynamic Interface between Vacuoles and Mitochondria in Yeast. *Developmental Cell* *30*, 95-102.
- Endres, M., Neupert, W., and Brunner, M. (1999). Transport of the ADP/ATP carrier of mitochondria from the TOM complex to the TIM22.54 complex. *The EMBO journal* *18*, 3214-3221.
- 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.
- Esser, K., Tursun, B., Ingenhoven, M., Michaelis, G., and Pratje, E. (2002). A novel two-step mechanism for removal of a mitochondrial signal sequence involves the mAAA complex and the putative rhomboid protease Pcp1. *Journal of molecular biology* *323*, 835-843.
- Feinberg, E.H., Vanhove, M.K., Bendesky, A., Wang, G., Fetter, R.D., Shen, K., and Bargmann, C.I. (2008). GFP Reconstitution Across Synaptic Partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* *57*, 353-363.
- Fernandez-Murray, J.P., and McMaster, C.R. (2006). Identification of novel phospholipid binding proteins in *Saccharomyces cerevisiae*. *FEBS letters* *580*, 82-86.
- Fernie, A.R., Carrari, F., and Sweetlove, L.J. (2004). Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. *Current opinion in plant biology* *7*, 254-261.
- Fersht, A.R., and Dingwall, C. (1979). Evidence for the double-sieve editing mechanism in protein synthesis. Steric exclusion of isoleucine by valyl-tRNA synthetases. *Biochemistry* *18*, 2627-2631.
- Feyder, S., De Craene, J.O., Bar, S., Bertazzi, D.L., and Friant, S. (2015). Membrane trafficking in the yeast *Saccharomyces cerevisiae* model. *Int J Mol Sci* *16*, 1509-1525.
- Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* *340*, 245-246.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., *et al.* (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* *269*, 496-512.

Foury, F., Roganti, T., Lecrenier, N., and Purnelle, B. (1998). The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae*. *FEBS letters* *440*, 325-331.

Frechin, M., Duchene, A.M., and Becker, H.D. (2009a). Translating organellar glutamine codons: a case by case scenario? *RNA Biol* *6*, 31-34.

Frechin, M., Enkler, L., Tetaud, E., Laporte, D., Senger, B., Blancard, C., Hammann, P., Bader, G., Clauder-Munster, S., Steinmetz, L.M., *et al.* (2014). Expression of nuclear and mitochondrial genes encoding ATP synthase is synchronized by disassembly of a multisynthetase complex. *Molecular cell* *56*, 763-776.

Frechin, M., Kern, D., Martin, R.P., Becker, H.D., and Senger, B. (2010). Arc1p: anchoring, routing, coordinating. *FEBS letters* *584*, 427-433.

Frechin, M., Senger, B., Braye, M., Kern, D., Martin, R.P., and Becker, H.D. (2009b). Yeast mitochondrial Gln-tRNA(Gln) is generated by a GatFAB-mediated transamidation pathway involving Arc1p-controlled subcellular sorting of cytosolic GluRS. *Genes & development* *23*, 1119-1130.

Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M., and Kisselev, L. (1996). Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. *Rna* *2*, 334-341.

Fu, G., Xu, T., Shi, Y., Wei, N., and Yang, X.L. (2012). tRNA-controlled nuclear import of a human tRNA synthetase. *The Journal of biological chemistry* *287*, 9330-9334.

Fukui, H., Hanaoka, R., and Kawahara, A. (2009). Noncanonical activity of seryl-tRNA synthetase is involved in vascular development. *Circ Res* *104*, 1253-1259.

Fusco, D., Accornero, N., Lavoie, B., Shenoy, S.M., Blanchard, J.M., Singer, R.H., and Bertrand, E. (2003). Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Current biology : CB* *13*, 161-167.

Galani, K., Grosshans, H., Deinert, K., Hurt, E., and Simos, G. (2001). The intracellular location of two aminoacyl-tRNA synthetases depends on complex formation with Arc1p. *The EMBO journal* *20*, 6889-6898.

Galani, K., Hurt, E., and Simos, G. (2005). The tRNA aminoacylation co-factor Arc1p is excluded from the nucleus by an Xpo1p-dependent mechanism. *FEBS letters* *579*, 969-975.

Galluzzi, L., Kepp, O., and Kroemer, G. (2012). Mitochondria: master regulators of danger signalling. *Nature reviews Molecular cell biology* *13*, 780-788.

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., and Weissman, J.S. (2003). Global analysis of protein expression in yeast. *Nature* *425*, 737-741.

Ghosh, I., Hamilton, A., and Regan, L. (2000). Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *J Am Chem Soc* *122*, 5658-5659.

Glick, B.S., Brandt, A., Cunningham, K., Muller, S., Hallberg, R.L., and Schatz, G. (1992). Cytochromes c1 and b2 are sorted to the intermembrane space of yeast mitochondria by a stop-transfer mechanism. *Cell* *69*, 809-822.

Godar, D.E., Godar, D.E., Garcia, V., Jacobo, A., Aebi, U., and Yang, D.C. (1988). Structural organization of the multienzyme complex of mammalian aminoacyl-tRNA synthetases. *Biochemistry* *27*, 6921-6928.

- Godinic-Mikulcic, V., Jaric, J., Greber, B.J., Franke, V., Hodnik, V., Anderluh, G., Ban, N., and Weygand-Durasevic, I. (2014). Archaeal aminoacyl-tRNA synthetases interact with the ribosome to recycle tRNAs. *Nucleic acids research* 42, 5191-5201.
- Godinic, V., Mocibob, M., Rocak, S., Ibba, M., and Weygand-Durasevic, I. (2007). Peroxin Pex21p interacts with the C-terminal noncatalytic domain of yeast seryl-tRNA synthetase and forms a specific ternary complex with tRNA(Ser). *The FEBS journal* 274, 2788-2799.
- Goksoyr, J. (1967). Evolution of eucaryotic cells. *Nature* 214, 1161.
- Goldgur, Y., and Safro, M. (1994). Aminoacyl-tRNA synthetases from *Haloarcula marismortui*: an evidence for a multienzyme complex in a procaryotic system. *Biochem Mol Biol Int* 32, 1075-1083.
- Golinelli-Cohen, M.P., and Mirande, M. (2007). Arc1p is required for cytoplasmic confinement of synthetases and tRNA. *Molecular and cellular biochemistry* 300, 47-59.
- Golinelli-Cohen, M.P., Zakrzewska, A., and Mirande, M. (2004). Complementation of yeast Arc1p by the p43 component of the human multisynthetase complex does not require its association with yeast MetRS and GluRS. *Journal of molecular biology* 340, 15-27.
- Gonczarowska-Jorge, H., Zahedi, R.P., and Sickmann, A. (2017). The proteome of baker's yeast mitochondria. *Mitochondrion* 33, 15-21.
- Gorner, W., Durchschlag, E., Martinez-Pastor, M.T., Estruch, F., Ammerer, G., Hamilton, B., Ruis, H., and Schuller, C. (1998). Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. *Genes & development* 12, 586-597.
- Gowher, A., Smirnov, A., Tarassov, I., and Entelis, N. (2013). Induced tRNA import into human mitochondria: implication of a host aminoacyl-tRNA-synthetase. *PLoS one* 8, e66228.
- Graindorge, J.S., Senger, B., Tritch, D., Simos, G., and Fasiolo, F. (2005). Role of Arc1p in the modulation of yeast glutamyl-tRNA synthetase activity. *Biochemistry* 44, 1344-1352.
- Gray, M.W. (2014). The pre-endosymbiont hypothesis: a new perspective on the origin and evolution of mitochondria. *Cold Spring Harb Perspect Biol* 6.
- Griesbeck, O., Baird, G.S., Campbell, R.E., Zacharias, D.A., and Tsien, R.Y. (2001). Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *The Journal of biological chemistry* 276, 29188-29194.
- Grosshans, H., Hurt, E., and Simos, G. (2000). An aminoacylation-dependent nuclear tRNA export pathway in yeast. *Genes & Dev* 14, 830-840.
- Gu, W., Hurto, R.L., Hopper, A.K., Grayhack, E.J., and Phizicky, E.M. (2005). Depletion of *Saccharomyces cerevisiae* tRNA(His) guanylyltransferase Thg1p leads to uncharged tRNA^{His} with additional m(5)C. *Molecular and cellular biology* 25, 8191-8201.
- Gu, W., Jackman, J.E., Lohan, A.J., Gray, M.W., and Phizicky, E.M. (2003). tRNA^{His} maturation: an essential yeast protein catalyzes addition of a guanine nucleotide to the 5' end of tRNA^{His}. *Genes & development* 17, 2889-2901.
- Gunasekera, N., Lee, S.W., Kim, S., Musier-Forsyth, K., and Arriaga, E. (2004). Nuclear localization of aminoacyl-tRNA synthetases using single-cell capillary electrophoresis laser-induced fluorescence analysis. *Analytical chemistry* 76, 4741-4746.
- Guo, M., and Schimmel, P. (2013). Essential nontranslational functions of tRNA synthetases. *Nat Chem Biol* 9, 145-153.
- Hamanaka, R.B., and Chandel, N.S. (2010). Mitochondrial reactive oxygen species regulate cellular signaling and dictate biological outcomes. *Trends in biochemical sciences* 35, 505-513.

- 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.
- Han, J.M., Park, B.J., Park, S.G., Oh, Y.S., Choi, S.J., Lee, S.W., Hwang, S.K., Chang, S.H., Cho, M.H., and Kim, S. (2008). AIMP2/p38, the scaffold for the multi-tRNA synthetase complex, responds to genotoxic stresses via p53. *Proceedings of the National Academy of Sciences of the United States of America* *105*, 11206-11211.
- Hao, B., Gong, W., Ferguson, T.K., James, C.M., Krzycki, J.A., and Chan, M.K. (2002). A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* *296*, 1462-1466.
- Harris, C.L. (1987). An aminoacyl-tRNA synthetase complex in *Escherichia coli*. *Journal of bacteriology* *169*, 2718-2723.
- Harris, C.L. (1990). High-molecular-weight forms of aminoacyl-tRNA synthetases and tRNA modification enzymes in *Escherichia coli*. *Journal of bacteriology* *172*, 1798-1803.
- Hausmann, C.D., and Ibba, M. (2008). Structural and functional mapping of the archaeal multi-aminoacyl-tRNA synthetase complex. *FEBS letters* *582*, 2178-2182.
- Hausmann, C.D., Praetorius-Ibba, M., and Ibba, M. (2007). An aminoacyl-tRNA synthetase:elongation factor complex for substrate channeling in archaeal translation. *Nucleic acids research* *35*, 6094-6102.
- Havrylenko, S., Legouis, R., Negrutskii, B., and Mirande, M. (2010). Methionyl-tRNA synthetase from *Caenorhabditis elegans*: a specific multidomain organization for convergent functional evolution. *Protein science : a publication of the Protein Society* *19*, 2475-2484.
- Havrylenko, S., Legouis, R., Negrutskii, B., and Mirande, M. (2011). *Caenorhabditis elegans* evolves a new architecture for the multi-aminoacyl-tRNA synthetase complex. *The Journal of biological chemistry* *286*, 28476-28487.
- Heim, R., Cubitt, A.B., and Tsien, R.Y. (1995). Improved green fluorescence. *Nature* *373*, 663-664.
- Heim, R., Prasher, D.C., and Tsien, R.Y. (1994). Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* *91*, 12501-12504.
- Heinemann, I.U., Soll, D., and Randau, L. (2010). Transfer RNA processing in archaea: unusual pathways and enzymes. *FEBS letters* *584*, 303-309.
- Hellmuth, K., Lau, D.M., Bischoff, F.R., Kunzler, M., Hurt, E., and Simos, G. (1998). Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. *Molecular and cellular biology* *18*, 6374-6386.
- Herbert, C.J., Labouesse, M., Dujardin, G., and Slonimski, P.P. (1988). The NAM2 proteins from *S. cerevisiae* and *S. douglasii* are mitochondrial leucyl-tRNA synthetases, and are involved in mRNA splicing. *The EMBO journal* *7*, 473-483.
- Herzog, W., Muller, K., Huisken, J., and Stainier, D.Y. (2009). Genetic evidence for a noncanonical function of seryl-tRNA synthetase in vascular development. *Circ Res* *104*, 1260-1266.
- Hinnebusch, A.G. (1997). Translational regulation of yeast GCN4. A window on factors that control initiator-tRNA binding to the ribosome. *The Journal of biological chemistry* *272*, 21661-21664.
- Hinnebusch, A.G. (2005). Translational regulation of GCN4 and the general amino acid control of yeast. *Annual review of microbiology* *59*, 407-450.
- Honda, S., Loher, P., Shigematsu, M., Palazzo, J.P., Suzuki, R., Imoto, I., Rigoutsos, I., and Kirino, Y. (2015). Sex hormone-dependent tRNA halves enhance cell proliferation in breast and prostate cancers. *Proceedings of the National Academy of Sciences of the United States of America* *112*, E3816-3825.

- Hopper, A.K., and Phizicky, E.M. (2003). tRNA transfers to the limelight. *Genes & development* *17*, 162-180.
- Horst, M., Oppliger, W., Rospert, S., Schonfeld, H.J., Schatz, G., and Azem, A. (1997). Sequential action of two hsp70 complexes during protein import into mitochondria. *The EMBO journal* *16*, 1842-1849.
- Hou, Y.M., and Schimmel, P. (1988). A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature* *333*, 140-145.
- Houten, S.M., and Wanders, R.J. (2010). A general introduction to the biochemistry of mitochondrial fatty acid beta-oxidation. *Journal of inherited metabolic disease* *33*, 469-477.
- Huang, H.Y., Tang, H.L., Chao, H.Y., Yeh, L.S., and Wang, C.C. (2006). An unusual pattern of protein expression and localization of yeast alanyl-tRNA synthetase isoforms. *Molecular microbiology* *60*, 189-198.
- Hyun, S.I., Maruri-Avidal, L., and Moss, B. (2015). Topology of Endoplasmic Reticulum-Associated Cellular and Viral Proteins Determined with Split-GFP. *Traffic* *16*, 787-795.
- Ibba, M., and Soll, D. (2000). Aminoacyl-tRNA synthesis. *Annu Rev Biochem* *69*, 617-650.
- Iborra, F., Dorizzi, M., and Labouesse, J. (1973). Tryptophanyl-transfer ribonucleic-acid synthetase from beef pancreas. Ligand binding and dissociation equilibrium between the active dimeric and inactive monomeric structures. *Eur J Biochem* *39*, 275-282.
- Ieva, R., Schrempp, S.G., Opalinski, L., Wollweber, F., Hoss, P., Heisswolf, A.K., Gebert, M., Zhang, Y., Guiard, B., Rospert, S., *et al.* (2014). Mgr2 functions as lateral gatekeeper for preprotein sorting in the mitochondrial inner membrane. *Molecular cell* *56*, 641-652.
- Imlay, J.A. (2008). Cellular defenses against superoxide and hydrogen peroxide. *Annual review of biochemistry* *77*, 755-776.
- Inouye, S., and Tsuji, F.I. (1994). Aequorea green fluorescent protein. Expression of the gene and fluorescence characteristics of the recombinant protein. *FEBS letters* *341*, 277-280.
- Jackson, K.E., Pham, J.S., Kwek, M., De Silva, N.S., Allen, S.M., Goodman, C.D., McFadden, G.I., Ribas de Pouplana, L., and Ralph, S.A. (2012). Dual targeting of aminoacyl-tRNA synthetases to the apicoplast and cytosol in *Plasmodium falciparum*. *Int J Parasitol* *42*, 177-186.
- Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature reviews Molecular cell biology* *11*, 113-127.
- Jahn, D. (1992). Complex formation between glutamyl-tRNA synthetase and glutamyl-tRNA reductase during the tRNA-dependent synthesis of 5-aminolevulinic acid in *Chlamydomonas reinhardtii*. *FEBS letters* *314*, 77-80.
- Jaramillo, A.M., Weil, T.T., Goodhouse, J., Gavis, E.R., and Schupbach, T. (2008). The dynamics of fluorescently labeled endogenous gurken mRNA in *Drosophila*. *Journal of cell science* *121*, 887-894.
- Jayaraman, S., Haggie, P., Wachter, R.M., Remington, S.J., and Verkman, A.S. (2000). Mechanism and cellular applications of a green fluorescent protein-based halide sensor. *The Journal of biological chemistry* *275*, 6047-6050.
- Jindal, S., Dudani, A.K., Singh, B., Harley, C.B., and Gupta, R.S. (1989). Primary structure of a human mitochondrial protein homologous to the bacterial and plant chaperonins and to the 65-kilodalton mycobacterial antigen. *Molecular and cellular biology* *9*, 2279-2283.
- Johnson, E.C., and Kent, S.B. (2006). Studies on the insolubility of a transmembrane peptide from signal peptide peptidase. *J Am Chem Soc* *128*, 7140-7141.
- Johnston, S.A., Anziano, P.Q., Shark, K., Sanford, J.C., and Butow, R.A. (1988). Mitochondrial transformation in yeast by bombardment with microprojectiles. *Science* *240*, 1538-1541.

Jones, D.H., McMillan, A.J., Fersht, A.R., and Winter, G. (1985). Reversible dissociation of dimeric tyrosyl-tRNA synthetase by mutagenesis at the subunit interface. *Biochemistry* 24, 5852-5857.

Jores, T., Klinger, A., Gross, L.E., Kawano, S., Flinner, N., Duchardt-Ferner, E., Wohnert, J., Kalbacher, H., Endo, T., Schleiff, E., *et al.* (2016). Characterization of the targeting signal in mitochondrial beta-barrel proteins. *Nat Commun* 7, 12036.

Kaddoum, L., Magdeleine, E., Waldo, G.S., Joly, E., and Cabantous, S. (2010). One-step split GFP staining for sensitive protein detection and localization in mammalian cells. *Biotechniques* 49, 727-728, 730, 732 *passim*.

Kaishima, M., Ishii, J., Matsuno, T., Fukuda, N., and Kondo, A. (2016). Expression of varied GFPs in *Saccharomyces cerevisiae*: codon optimization yields stronger than expected expression and fluorescence intensity. *Sci Rep* 6, 35932.

Kalderon, D., Richardson, W.D., Markham, A.F., and Smith, A.E. (1984). Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* 311, 33-38.

Kaminska, M., Deniziak, M., Kerjan, P., Barciszewski, J., and Mirande, M. (2000). A recurrent general RNA binding domain appended to plant methionyl-tRNA synthetase acts as a cis-acting cofactor for aminoacylation. *The EMBO journal* 19, 6908-6917.

Kaminska, M., Havrylenko, S., Decottignies, P., Gillet, S., Le Marechal, P., Negrutskii, B., and Mirande, M. (2009a). Dissection of the structural organization of the aminoacyl-tRNA synthetase complex. *The Journal of biological chemistry* 284, 6053-6060.

Kaminska, M., Havrylenko, S., Decottignies, P., Le Marechal, P., Negrutskii, B., and Mirande, M. (2009b). Dynamic Organization of Aminoacyl-tRNA Synthetase Complexes in the Cytoplasm of Human Cells. *The Journal of biological chemistry* 284, 13746-13754.

Kamiyama, D., Sekine, S., Barsi-Rhyne, B., Hu, J., Chen, B., Gilbert, L.A., Ishikawa, H., Leonetti, M.D., Marshall, W.F., Weissman, J.S., *et al.* (2016). Versatile protein tagging in cells with split fluorescent protein. *Nat Commun* 7, 11046.

Kang, P.J., Ostermann, J., Shilling, J., Neupert, W., Craig, E.A., and Pfanner, N. (1990). Requirement for hsp70 in the mitochondrial matrix for translocation and folding of precursor proteins. *Nature* 348, 137-143.

Kang, T., Kwon, N.H., Lee, J.Y., Park, M.C., Kang, E., Kim, H.H., Kang, T.J., and Kim, S. (2012). AIMP3/p18 controls translational initiation by mediating the delivery of charged initiator tRNA to initiation complex. *Journal of molecular biology* 423, 475-481.

Kapasi, P., Chaudhuri, S., Vyas, K., Baus, D., Komar, A.A., Fox, P.L., Merrick, W.C., and Mazumder, B. (2007). L13a blocks 48S assembly: role of a general initiation factor in mRNA-specific translational control. *Molecular cell* 25, 113-126.

Karanasios, E., Boleti, H., and Simos, G. (2008). Incorporation of the Arc1p tRNA-binding domain to the catalytic core of MetRS can functionally replace the yeast Arc1p-MetRS complex. *Journal of molecular biology* 381, 763-771.

Karanasios, E., Simader, H., Panayotou, G., Suck, D., and Simos, G. (2007). Molecular determinants of the yeast Arc1p-aminoacyl-tRNA synthetase complex assembly. *Journal of molecular biology* 374, 1077-1090.

Kelemen, B.R., Klink, T.A., Behlke, M.A., Eubanks, S.R., Leland, P.A., and Raines, R.T. (1999). Hypersensitive substrate for ribonucleases. *Nucleic acids research* 27, 3696-3701.

Kellermann, O., Brevet, A., Tonetti, H., and Waller, J.P. (1979). Macromolecular complexes of aminoacyl-tRNA synthetases from eukaryotes. 1. Extensive purification and characterization of the

high-molecular-weight complex(es) of seven aminoacyl-tRNA synthetases from sheep liver. *Eur J Biochem* 99, 541-550.

Kellermann, O., Tonetti, H., Brevet, A., Mirande, M., Pailliez, J.P., and Waller, J.P. (1982). Macromolecular complexes from sheep and rabbit containing seven aminoacyl-tRNA synthetases. I. Species specificity of the polypeptide composition. *The Journal of biological chemistry* 257, 11041-11048.

Kemper, C., Habib, S.J., Engl, G., Heckmeyer, P., Dimmer, K.S., and Rapaport, D. (2008). Integration of tail-anchored proteins into the mitochondrial outer membrane does not require any known import components. *Journal of cell science* 121, 1990-1998.

Kerjan, P., Cerini, C., Semeriva, M., and Mirande, M. (1994). The multienzyme complex containing nine aminoacyl-tRNA synthetases is ubiquitous from *Drosophila* to mammals. *Biochimica et biophysica acta* 1199, 293-297.

Khoshnevis, S., Neumann, P., and Ficner, R. (2010). Crystal structure of the RNA recognition motif of yeast translation initiation factor eIF3b reveals differences to human eIF3b. *PLoS one* 5.

Kim, E., Kim, S.H., Kim, S., and Kim, T.S. (2006). The novel cytokine p43 induces IL-12 production in macrophages via NF-kappaB activation, leading to enhanced IFN-gamma production in CD4+ T cells. *J Immunol* 176, 256-264.

Kim, H.S., Hoja, U., Stolz, J., Sauer, G., and Schweizer, E. (2004a). Identification of the tRNA-binding protein Arc1p as a novel target of in vivo biotinylation in *Saccharomyces cerevisiae*. *The Journal of biological chemistry* 279, 42445-42452.

Kim, J.S., and Raines, R.T. (1993). Ribonuclease S-peptide as a carrier in fusion proteins. *Protein science : a publication of the Protein Society* 2, 348-356.

Kim, K.J., Park, M.C., Choi, S.J., Oh, Y.S., Choi, E.C., Cho, H.J., Kim, M.H., Kim, S.H., Kim, D.W., Kim, S., *et al.* (2008). Determination of three-dimensional structure and residues of the novel tumor suppressor AIMP3/p18 required for the interaction with ATM. *The Journal of biological chemistry* 283, 14032-14040.

Kim, S.A., Heinze, K.G., Bacia, K., Waxham, M.N., and Schwille, P. (2005). Two-photon cross-correlation analysis of intracellular reactions with variable stoichiometry. *Biophys J* 88, 4319-4336.

Kim, S.A., Heinze, K.G., Waxham, M.N., and Schwille, P. (2004b). Intracellular calmodulin availability accessed with two-photon cross-correlation. *Proceedings of the National Academy of Sciences of the United States of America* 101, 105-110.

Kim, T., Park, S.G., Kim, J.E., Seol, W., Ko, Y.G., and Kim, S. (2000). Catalytic peptide of human glutamyl-tRNA synthetase is essential for its assembly to the aminoacyl-tRNA synthetase complex. *The Journal of biological chemistry* 275, 21768-21772.

Kim, Y.E., Kim, Y.N., Kim, J.A., Kim, H.M., and Jung, Y. (2015). Green fluorescent protein nanopolygons as monodisperse supramolecular assemblies of functional proteins with defined valency. *Nat Commun* 6, 7134.

Kimelman, A., Levy, A., Sberro, H., Kidron, S., Leavitt, A., Amitai, G., Yoder-Himes, D.R., Wurtzel, O., Zhu, Y., Rubin, E.M., *et al.* (2012). A vast collection of microbial genes that are toxic to bacteria. *Genome research* 22, 802-809.

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. *Nature structural & molecular biology* 11, 149-156.

Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S., and Ban, N. (2011). Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science* 334, 941-948.

- Knapp, G., Ogden, R.C., Peebles, C.L., and Abelson, J. (1979). Splicing of yeast tRNA precursors: structure of the reaction intermediates. *Cell* 18, 37-45.
- Ko, H.S., von Coelln, R., Sriram, S.R., Kim, S.W., Chung, K.K., Pletnikova, O., Troncoso, J., Johnson, B., Saffary, R., Goh, E.L., *et al.* (2005). Accumulation of the authentic parkin substrate aminoacyl-tRNA synthetase cofactor, p38/JTV-1, leads to catecholaminergic cell death. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 25, 7968-7978.
- Ko, Y.G., Kang, Y.S., Kim, E.K., Park, S.G., and Kim, S. (2000). Nucleolar localization of human methionyl-tRNA synthetase and its role in ribosomal RNA synthesis. *The Journal of cell biology* 149, 567-574.
- Ko, Y.G., Kim, E.Y., Kim, T., Park, H., Park, H.S., Choi, E.J., and Kim, S. (2001). Glutamine-dependent antiapoptotic interaction of human glutaminyl-tRNA synthetase with apoptosis signal-regulating kinase 1. *The Journal of biological chemistry* 276, 6030-6036.
- Koch, J.R., and Schmid, F.X. (2014). Mia40 is optimized for function in mitochondrial oxidative protein folding and import. *ACS chemical biology* 9, 2049-2057.
- Koehler, C., Round, A., Simader, H., Suck, D., and Svergun, D. (2013). Quaternary structure of the yeast Arc1p-aminoacyl-tRNA synthetase complex in solution and its compaction upon binding of tRNAs. *Nucleic acids research* 41, 667-676.
- Kolkman, A., Olsthoorn, M.M., Heeremans, C.E., Heck, A.J., and Slijper, M. (2005). Comparative proteome analysis of *Saccharomyces cerevisiae* grown in chemostat cultures limited for glucose or ethanol. *Molecular & cellular proteomics : MCP* 4, 1-11.
- Kolupaeva, V.G., de Breyne, S., Pestova, T.V., and Hellen, C.U. (2007). In vitro reconstitution and biochemical characterization of translation initiation by internal ribosomal entry. *Methods in enzymology* 430, 409-439.
- Koonin, E.V., and Novozhilov, A.S. (2009). Origin and evolution of the genetic code: the universal enigma. *IUBMB life* 61, 99-111.
- Koppel, D.E., Axelrod, D., Schlessinger, J., Elson, E.L., and Webb, W.W. (1976). Dynamics of fluorescence marker concentration as a probe of mobility. *Biophys J* 16, 1315-1329.
- Kornmann, B., Currie, E., Collins, S.R., Schuldiner, M., Nunnari, J., Weissman, J.S., and Walter, P. (2009). An ER-mitochondria tethering complex revealed by a synthetic biology screen. *Science* 325, 477-481.
- Kornmann, B., Osman, C., and Walter, P. (2011). The conserved GTPase Gem1 regulates endoplasmic reticulum-mitochondria connections. *Proceedings of the National Academy of Sciences of the United States of America* 108, 14151-14156.
- Kozak, M. (1978). How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell* 15, 1109-1123.
- Kozel, B.A., Rongish, B.J., Czirok, A., Zach, J., Little, C.D., Davis, E.C., Knutsen, R.H., Wagenseil, J.E., Levy, M.A., and Mecham, R.P. (2006). Elastic fiber formation: a dynamic view of extracellular matrix assembly using timer reporters. *J Cell Physiol* 207, 87-96.
- Krebs, H.A., and Johnson, W.A. (1937). Metabolism of ketonic acids in animal tissues. *Biochem J* 31, 645-660.
- Kruger, M.K., Pedersen, S., Hagervall, T.G., and Sorensen, M.A. (1998). The modification of the wobble base of tRNA^{Glu} modulates the translation rate of glutamic acid codons in vivo. *Journal of molecular biology* 284, 621-631.
- Krumpe, K., Frumkin, I., Herzig, Y., Rimon, N., Ozbalci, C., Brugger, B., Rapaport, D., and Schuldiner, M. (2012). Ergosterol content specifies targeting of tail-anchored proteins to mitochondrial outer membranes. *Molecular biology of the cell* 23, 3927-3935.

- Kufel, J., and Tollervey, D. (2003). 3'-processing of yeast tRNA^{Trp} precedes 5'-processing. *Rna* 9, 202-208.
- Kulzer, S., Petersen, W., Baser, A., Mandel, K., and Przyborski, J.M. (2013). Use of self-assembling GFP to determine protein topology and compartmentalisation in the *Plasmodium falciparum*-infected erythrocyte. *Mol Biochem Parasitol* 187, 87-90.
- Kuner, T., and Augustine, G.J. (2000). A genetically encoded ratiometric indicator for chloride: capturing chloride transients in cultured hippocampal neurons. *Neuron* 27, 447-459.
- Kurland, C.G., and Andersson, S.G. (2000). Origin and evolution of the mitochondrial proteome. *Microbiology and molecular biology reviews : MMBR* 64, 786-820.
- Kutik, S., Stojanovski, D., Becker, L., Becker, T., Meinecke, M., Kruger, V., Prinz, C., Meisinger, C., Guiard, B., Wagner, R., *et al.* (2008). Dissecting membrane insertion of mitochondrial beta-barrel proteins. *Cell* 132, 1011-1024.
- Kwon, N.H., Kang, T., Lee, J.Y., Kim, H.H., Kim, H.R., Hong, J., Oh, Y.S., Han, J.M., Ku, M.J., Lee, S.Y., *et al.* (2011). Dual role of methionyl-tRNA synthetase in the regulation of translation and tumor suppressor activity of aminoacyl-tRNA synthetase-interacting multifunctional protein-3. *Proceedings of the National Academy of Sciences of the United States of America* 108, 19635-19640.
- Lange, S., Katayama, Y., Schmid, M., Burkacky, O., Brauchle, C., Lamb, D.C., and Jansen, R.P. (2008). Simultaneous transport of different localized mRNA species revealed by live-cell imaging. *Traffic* 9, 1256-1267.
- LaRiviere, F.J., Wolfson, A.D., and Uhlenbeck, O.C. (2001). Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science* 294, 165-168.
- Lederberg, J. (1952). Cell genetics and hereditary symbiosis. *Physiol Rev* 32, 403-430.
- Lee, C.M., Sedman, J., Neupert, W., and Stuart, R.A. (1999). The DNA helicase, Hmi1p, is transported into mitochondria by a C-terminal cleavable targeting signal. *The Journal of biological chemistry* 274, 20937-20942.
- Lee, E.Y., Lee, H.C., Kim, H.K., Jang, S.Y., Park, S.J., Kim, Y.H., Kim, J.H., Hwang, J., Kim, J.H., Kim, T.H., *et al.* (2016). Infection-specific phosphorylation of glutamyl-prolyl tRNA synthetase induces antiviral immunity. *Nat Immunol* 17, 1252-1262.
- Lee, Y.N., Nechushtan, H., Figov, N., and Razin, E. (2004). The function of lysyl-tRNA synthetase and Ap4A as signaling regulators of MITF activity in FcepsilonRI-activated mast cells. *Immunity* 20, 145-151.
- Lee, Y.S., Shibata, Y., Malhotra, A., and Dutta, A. (2009). A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes & development* 23, 2639-2649.
- Lenaz, G., Baracca, A., Barbero, G., Bergamini, C., Dalmonte, M.E., Del Sole, M., Faccioli, M., Falasca, A., Fato, R., Genova, M.L., *et al.* (2010). Mitochondrial respiratory chain super-complex I-III in physiology and pathology. *Biochimica et biophysica acta* 1797, 633-640.
- Lill, R., and Muhlenhoff, U. (2005). Iron-sulfur-protein biogenesis in eukaryotes. *Trends in biochemical sciences* 30, 133-141.
- Ling, C., Yao, Y.N., Zheng, Y.G., Wei, H., Wang, L., Wu, X.F., and Wang, E.D. (2005). The C-terminal appended domain of human cytosolic leucyl-tRNA synthetase is indispensable in its interaction with arginyl-tRNA synthetase in the multi-tRNA synthetase complex. *The Journal of biological chemistry* 280, 34755-34763.
- Ling, J., O'Donoghue, P., and Soll, D. (2015). Genetic code flexibility in microorganisms: novel mechanisms and impact on physiology. *Nat Rev Microbiol* 13, 707-721.

- Lipman, R.S., Chen, J., Evilia, C., Vitseva, O., and Hou, Y.M. (2003). Association of an aminoacyl-tRNA synthetase with a putative metabolic protein in archaea. *Biochemistry* *42*, 7487-7496.
- Lippincott-Schwartz, J., Altan-Bonnet, N., and Patterson, G.H. (2003). Photobleaching and photoactivation: following protein dynamics in living cells. *Nature cell biology Suppl*, S7-14.
- Livet, J., Weissman, T.A., Kang, H., Draft, R.W., Lu, J., Bennis, R.A., Sanes, J.R., and Lichtman, J.W. (2007). Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* *450*, 56-62.
- Llopis, J., McCaffery, J.M., Miyawaki, A., Farquhar, M.G., and Tsien, R.Y. (1998). Measurement of cytosolic, mitochondrial, and Golgi pH in single living cells with green fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America* *95*, 6803-6808.
- Luo, S., and Levine, R.L. (2009). Methionine in proteins defends against oxidative stress. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* *23*, 464-472.
- Magliery, T.J., Wilson, C.G., Pan, W., Mishler, D., Ghosh, I., Hamilton, A.D., and Regan, L. (2005). Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J Am Chem Soc* *127*, 146-157.
- Mailu, B.M., Li, L., Arthur, J., Nelson, T.M., Ramasamy, G., Fritz-Wolf, K., Becker, K., and Gardner, M.J. (2015). Plasmodium Apicoplast Gln-tRNA^{Gln} Biosynthesis Utilizes a Unique GatAB Amidotransferase Essential for Erythrocytic Stage Parasites. *The Journal of biological chemistry* *290*, 29629-29641.
- Mailu, B.M., Ramasamay, G., Mudeppa, D.G., Li, L., Lindner, S.E., Peterson, M.J., DeRocher, A.E., Kappe, S.H., Rathod, P.K., and Gardner, M.J. (2013). A nondiscriminating glutamyl-tRNA synthetase in the plasmodium apicoplast: the first enzyme in an indirect aminoacylation pathway. *The Journal of biological chemistry* *288*, 32539-32552.
- Majumdar, R., Bandyopadhyay, A., and Maitra, U. (2003). Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex. *The Journal of biological chemistry* *278*, 6580-6587.
- Maraia, R.J., and Lamichhane, T.N. (2011). 3' processing of eukaryotic precursor tRNAs. *Wiley interdisciplinary reviews RNA* *2*, 362-375.
- Marquet, R., Isel, C., Ehresmann, C., and Ehresmann, B. (1995). tRNAs as primer of reverse transcriptases. *Biochimie* *77*, 113-124.
- Martens-Uzunova, E.S., Jalava, S.E., Dits, N.F., van Leenders, G.J., Moller, S., Trapman, J., Bangma, C.H., Litman, T., Visakorpi, T., and Jenster, G. (2012). Diagnostic and prognostic signatures from the small non-coding RNA transcriptome in prostate cancer. *Oncogene* *31*, 978-991.
- Martin, J., Mahlke, K., and Pfanner, N. (1991). Role of an energized inner membrane in mitochondrial protein import. Delta psi drives the movement of presequences. *The Journal of biological chemistry* *266*, 18051-18057.
- Mascarenhas, A.P., and Martinis, S.A. (2008). Functional segregation of a predicted "hinge" site within the beta-strand linkers of Escherichia coli leucyl-tRNA synthetase. *Biochemistry* *47*, 4808-4816.
- McCarthy, J.E. (1998). Posttranscriptional control of gene expression in yeast. *Microbiology and molecular biology reviews : MMBR* *62*, 1492-1553.
- Meinecke, M., Wagner, R., Kovermann, P., Guiard, B., Mick, D.U., Hutu, D.P., Voos, W., Truscott, K.N., Chacinska, A., Pfanner, N., *et al.* (2006). Tim50 maintains the permeability barrier of the mitochondrial inner membrane. *Science* *312*, 1523-1526.
- Meisinger, C., Pfanner, N., and Truscott, K.N. (2006). Isolation of yeast mitochondria. *Methods Mol Biol* *313*, 33-39.

- Mereschkowsky, C. (1905). Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol Centralbl*, 25, 593–604.
- Mick, D.U., Wagner, K., van der Laan, M., Frazier, A.E., Perschil, I., Pawlas, M., Meyer, H.E., Warscheid, B., and Rehling, P. (2007). Shy1 couples Cox1 translational regulation to cytochrome c oxidase assembly. *The EMBO journal* 26, 4347-4358.
- Mirabella, R., Franken, C., van der Krogt, G.N., Bisseling, T., and Geurts, R. (2004). Use of the fluorescent timer DsRED-E5 as reporter to monitor dynamics of gene activity in plants. *Plant Physiol* 135, 1879-1887.
- Mirande, M., Kellermann, O., and Waller, J.P. (1982). Macromolecular complexes from sheep and rabbit containing seven aminoacyl-tRNA synthetases. II. Structural characterization of the polypeptide components and immunological identification of the methionyl-tRNA synthetase subunit. *The Journal of biological chemistry* 257, 11049-11055.
- Mirande, M., Le Corre, D., Louvard, D., Reggio, H., Pailliez, J.P., and Waller, J.P. (1985). Association of an aminoacyl-tRNA synthetase complex and of phenylalanyl-tRNA synthetase with the cytoskeletal framework fraction from mammalian cells. *Experimental cell research* 156, 91-102.
- Mireau, H., Lancelin, D., and Small, I.D. (1996). The same Arabidopsis gene encodes both cytosolic and mitochondrial alanyl-tRNA synthetases. *Plant Cell* 8, 1027-1039.
- Miyatsuka, T., Li, Z., and German, M.S. (2009). Chronology of islet differentiation revealed by temporal cell labeling. *Diabetes* 58, 1863-1868.
- Moine, H., Romby, P., Springer, M., Grunberg-Manago, M., Ebel, J.P., Ehresmann, C., and Ehresmann, B. (1988). Messenger RNA structure and gene regulation at the translational level in *Escherichia coli*: the case of threonine:tRNA^{Thr} ligase. *Proceedings of the National Academy of Sciences of the United States of America* 85, 7892-7896.
- Moradpour, D., Evans, M.J., Gosert, R., Yuan, Z., Blum, H.E., Goff, S.P., Lindenbach, B.D., and Rice, C.M. (2004). Insertion of green fluorescent protein into nonstructural protein 5A allows direct visualization of functional hepatitis C virus replication complexes. *J Virol* 78, 7400-7409.
- Morin, J.G., and Hastings, J.W. (1971). Energy transfer in a bioluminescent system. *J Cell Physiol* 77, 313-318.
- Morise, H., Shimomura, O., Johnson, F.H., and Winant, J. (1974). Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* 13, 2656-2662.
- Mossmann, D., Meisinger, C., and Vogtle, F.N. (2012). Processing of mitochondrial presequences. *Biochimica et biophysica acta* 1819, 1098-1106.
- Mossmann, D., Vogtle, F.N., Taskin, A.A., Teixeira, P.F., Ring, J., Burkhart, J.M., Burger, N., Pinho, C.M., Tadic, J., Loreth, D., *et al.* (2014). Amyloid-beta Peptide Induces Mitochondrial Dysfunction by Inhibition of Preprotein Maturation. *Cell metabolism* 20, 662-669.
- Mukai, T., Kobayashi, T., Hino, N., Yanagisawa, T., Sakamoto, K., and Yokoyama, S. (2008). Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochemical and biophysical research communications* 371, 818-822.
- Mukhopadhyay, R., Jia, J., Arif, A., Ray, P.S., and Fox, P.L. (2009). The GAIT system: a gatekeeper of inflammatory gene expression. *Trends in biochemical sciences* 34, 324-331.
- Mukhopadhyay, R., Ray, P.S., Arif, A., Brady, A.K., Kinter, M., and Fox, P.L. (2008). DAPK-ZIPK-L13a axis constitutes a negative-feedback module regulating inflammatory gene expression. *Molecular cell* 32, 371-382.

- Nagai, T., Ibata, K., Park, E.S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nature biotechnology* *20*, 87-90.
- Nagao, A., Suzuki, T., Katoh, T., Sakaguchi, Y., and Suzuki, T. (2009). Biogenesis of glutaminyl-mt tRNAGln in human mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* *106*, 16209-16214.
- Nakagawa, S., Niimura, Y., Gojobori, T., Tanaka, H., and Miura, K. (2008). Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic acids research* *36*, 861-871.
- Namgoong, S., Sheppard, K., Sherrer, R.L., and Soll, D. (2007). Co-evolution of the archaeal tRNA-dependent amidotransferase GatCAB with tRNA(Asn). *FEBS letters* *581*, 309-314.
- Natsoulis, G., Hilger, F., and Fink, G.R. (1986). The HTS1 gene encodes both the cytoplasmic and mitochondrial histidine tRNA synthetases of *S. cerevisiae*. *Cell* *46*, 235-243.
- Navarre, W.W., and Schneewind, O. (1999). Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiology and molecular biology reviews* : *MMBR* *63*, 174-229.
- Negrutskii, B.S., Stapulionis, R., and Deutscher, M.P. (1994). Supramolecular organization of the mammalian translation system. *Proc Natl Acad Sci U S A* *91*, 964-968.
- Netzer, N., Goodenbour, J.M., David, A., Dittmar, K.A., Jones, R.B., Schneider, J.R., Boone, D., Eves, E.M., Rosner, M.R., Gibbs, J.S., *et al.* (2009). Innate immune and chemically triggered oxidative stress modifies translational fidelity. *Nature* *462*, 522-526.
- Niwa, H., Inouye, S., Hirano, T., Matsuno, T., Kojima, S., Kubota, M., Ohashi, M., and Tsuji, F.I. (1996). Chemical nature of the light emitter of the Aequorea green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* *93*, 13617-13622.
- Nixon, A.E., and Benkovic, S.J. (2000). Improvement in the efficiency of formyl transfer of a GAR transformylase hybrid enzyme. *Protein Eng* *13*, 323-327.
- Norcum, M.T. (1989). Isolation and electron microscopic characterization of the high molecular mass aminoacyl-tRNA synthetase complex from murine erythroleukemia cells. *The Journal of biological chemistry* *264*, 15043-15051.
- O'Connor, J.P., and Peebles, C.L. (1991). In vivo pre-tRNA processing in *Saccharomyces cerevisiae*. *Molecular and cellular biology* *11*, 425-439.
- 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.
- Ohira, T., and Suzuki, T. (2011). Retrograde nuclear import of tRNA precursors is required for modified base biogenesis in yeast. *Proceedings of the National Academy of Sciences of the United States of America* *108*, 10502-10507.
- Ohlmeier, S., Kastaniotis, A.J., Hiltunen, J.K., and Bergmann, U. (2004). The yeast mitochondrial proteome, a study of fermentative and respiratory growth. *The Journal of biological chemistry* *279*, 3956-3979.
- Ormo, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., and Remington, S.J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. *Science* *273*, 1392-1395.
- Orrenius, S., Zhivotovsky, B., and Nicotera, P. (2003). Regulation of cell death: the calcium-apoptosis link. *Nature reviews Molecular cell biology* *4*, 552-565.

Otani, A., Slike, B.M., Dorrell, M.I., Hood, J., Kinder, K., Ewalt, K.L., Cheresch, D., Schimmel, P., and Friedlander, M. (2002). A fragment of human TrpRS as a potent antagonist of ocular angiogenesis. *Proceedings of the National Academy of Sciences of the United States of America* *99*, 178-183.

Ozanick, S.G., Wang, X., Costanzo, M., Brost, R.L., Boone, C., and Anderson, J.T. (2009). Rex1p deficiency leads to accumulation of precursor initiator tRNAMet and polyadenylation of substrate RNAs in *Saccharomyces cerevisiae*. *Nucleic acids research* *37*, 298-308.

Ozawa, T., Natori, Y., Sato, M., and Umezawa, Y. (2007). Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat Methods* *4*, 413-419.

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

Paley, E.L., Baranov, V.N., Alexandrova, N.M., and Kisselev, L.L. (1991). Tryptophanyl-tRNA synthetase in cell lines resistant to tryptophan analogs. *Experimental cell research* *195*, 66-78.

Park, B.J., Kang, J.W., Lee, S.W., Choi, S.J., Shin, Y.K., Ahn, Y.H., Choi, Y.H., Choi, D., Lee, K.S., and Kim, S. (2005). The haploinsufficient tumor suppressor p18 upregulates p53 via interactions with ATM/ATR. *Cell* *120*, 209-221.

Park, B.J., Oh, Y.S., Park, S.Y., Choi, S.J., Rudolph, C., Schlegelberger, B., and Kim, S. (2006). AIMP3 haploinsufficiency disrupts oncogene-induced p53 activation and genomic stability. *Cancer Res* *66*, 6913-6918.

Park, H., Park, S.G., Kim, J., Ko, Y.G., and Kim, S. (2002). Signaling pathways for TNF production induced by human aminoacyl-tRNA synthetase-associating factor, p43. *Cytokine* *20*, 148-153.

Park, K., Yi, S.Y., Lee, C.S., Kim, K.E., Pai, H.S., Seol, D.W., Chung, B.H., and Kim, M. (2007). A split enhanced green fluorescent protein-based reporter in yeast two-hybrid system. *Protein J* *26*, 107-116.

Patterson, G.H., Knobel, S.M., Sharif, W.D., Kain, S.R., and Piston, D.W. (1997). Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys J* *73*, 2782-2790.

Paukstelis, P.J., and Lambowitz, A.M. (2008). Identification and evolution of fungal mitochondrial tyrosyl-tRNA synthetases with group I intron splicing activity. *Proceedings of the National Academy of Sciences of the United States of America* *105*, 6010-6015.

Paushkin, S.V., Patel, M., Furia, B.S., Peltz, S.W., and Trotta, C.R. (2004). Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3' end formation. *Cell* *117*, 311-321.

Pedelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nature biotechnology* *24*, 79-88.

Peebles, C.L., Ogden, R.C., Knapp, G., and Abelson, J. (1979). Splicing of yeast tRNA precursors: a two-stage reaction. *Cell* *18*, 27-35.

Pelc, S.R. (1965). Correlation between coding-triplets and amino-acids. *Nature* *207*, 597-599.

Pelet, S. (2017). Nuclear relocation of Kss1 contributes to the specificity of the mating response. *Sci Rep* *7*, 43636.

Perkins, E., Sun, D., Nguyen, A., Tulac, S., Francesco, M., Tavana, H., Nguyen, H., Tugendreich, S., Barthmaier, P., Couto, J., *et al.* (2001). Novel inhibitors of poly(ADP-ribose) polymerase/PARP1 and PARP2 identified using a cell-based screen in yeast. *Cancer Res* *61*, 4175-4183.

Perona, J.J., and Gruic-Sovulj, I. (2014). Synthetic and editing mechanisms of aminoacyl-tRNA synthetases. *Topics in current chemistry* *344*, 1-41.

- Perozzo, M.A., Ward, K.B., Thompson, R.B., and Ward, W.W. (1988). X-ray diffraction and time-resolved fluorescence analyses of Aequorea green fluorescent protein crystals. *The Journal of biological chemistry* *263*, 7713-7716.
- Petschnigg, J., Groisman, B., Kotlyar, M., Taipale, M., Zheng, Y., Kurat, C.F., Sayad, A., Sierra, J.R., Mattiazzi Usaj, M., Snider, J., *et al.* (2014). The mammalian-membrane two-hybrid assay (MaMTH) for probing membrane-protein interactions in human cells. *Nat Methods* *11*, 585-592.
- Phair, R.D., and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. *Nature* *404*, 604-609.
- Pham, J.S., Sakaguchi, R., Yeoh, L.M., De Silva, N.S., McFadden, G.I., Hou, Y.M., and Ralph, S.A. (2014). A dual-targeted aminoacyl-tRNA synthetase in *Plasmodium falciparum* charges cytosolic and apicoplast tRNACys. *Biochem J* *458*, 513-523.
- Phizicky, E.M., and Hopper, A.K. (2010). tRNA biology charges to the front. *Genes & development* *24*, 1832-1860.
- Pierrel, F., Bestwick, M.L., Cobine, P.A., Khalimonchuk, O., Cricco, J.A., and Winge, D.R. (2007). Coa1 links the Mss51 post-translational function to Cox1 cofactor insertion in cytochrome c oxidase assembly. *The EMBO journal* *26*, 4335-4346.
- Pinaud, F., and Dahan, M. (2011). Targeting and imaging single biomolecules in living cells by complementation-activated light microscopy with split-fluorescent proteins. *Proceedings of the National Academy of Sciences of the United States of America* *108*, E201-210.
- Pinton, P., Giorgi, C., Siviero, R., Zecchini, E., and Rizzuto, R. (2008). Calcium and apoptosis: ER-mitochondria Ca²⁺ transfer in the control of apoptosis. *Oncogene* *27*, 6407-6418.
- Pisarev, A.V., Hellen, C.U., and Pestova, T.V. (2007). Recycling of eukaryotic posttermination ribosomal complexes. *Cell* *131*, 286-299.
- Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., Skabkina, O.V., Rakotondrafara, A.M., Hentze, M.W., Hellen, C.U., and Pestova, T.V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Molecular cell* *37*, 196-210.
- Piston, D.W., and Kremers, G.J. (2007). Fluorescent protein FRET: the good, the bad and the ugly. *Trends in biochemical sciences* *32*, 407-414.
- Polycarpo, C., Ambrogelly, A., Berube, A., Winbush, S.M., McCloskey, J.A., Crain, P.F., Wood, J.L., and Soll, D. (2004). An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. *Proceedings of the National Academy of Sciences of the United States of America* *101*, 12450-12454.
- Popenko, V.I., Cherny, N.E., Beresten, S.F., Ivanova, J.L., Filonenko, V.V., and Kisselev, L.L. (1993). Immunoelectron microscopic location of tryptophanyl-tRNA synthetase in mammalian, prokaryotic and archaeobacterial cells. *Eur J Cell Biol* *62*, 248-258.
- Praetorius-Ibba, M., Hausmann, C.D., Paras, M., Rogers, T.E., and Ibba, M. (2007). Functional association between three archaeal aminoacyl-tRNA synthetases. *The Journal of biological chemistry* *282*, 3680-3687.
- Prasher, D., McCann, R.O., and Cormier, M.J. (1985). Cloning and expression of the cDNA coding for aequorin, a bioluminescent calcium-binding protein. *Biochemical and biophysical research communications* *126*, 1259-1268.
- Prendergast, F.G., and Mann, K.G. (1978). Chemical and physical properties of aequorin and the green fluorescent protein isolated from *Aequorea forskalea*. *Biochemistry* *17*, 3448-3453.

Preston, M.A., and Phizicky, E.M. (2010). The requirement for the highly conserved G-1 residue of *Saccharomyces cerevisiae* tRNA^{His} can be circumvented by overexpression of tRNA^{His} and its synthetase. *Rna* 16, 1068-1077.

Pujol, C., Bailly, M., Kern, D., Marechal-Drouard, L., Becker, H., and Duchene, A.M. (2008). Dual-targeted tRNA-dependent amidotransferase ensures both mitochondrial and chloroplastic Gln-tRNA^{Gln} synthesis in plants. *Proceedings of the National Academy of Sciences of the United States of America* 105, 6481-6485.

Putney, S.D., and Schimmel, P. (1981). An aminoacyl tRNA synthetase binds to a specific DNA sequence and regulates its gene transcription. *Nature* 291, 632-635.

Quevillon, S., Agou, F., Robinson, J.C., and Mirande, M. (1997). The p43 component of the mammalian multi-synthetase complex is likely to be the precursor of the endothelial monocyte-activating polypeptide II cytokine. *The Journal of biological chemistry* 272, 32573-32579.

Quevillon, S., and Mirande, M. (1996). The p18 component of the multisynthetase complex shares a protein motif with the beta and gamma subunits of eukaryotic elongation factor 1. *FEBS letters* 395, 63-67.

Quevillon, S., Robinson, J.C., Berthonneau, E., Siatecka, M., and Mirande, M. (1999). Macromolecular assemblage of aminoacyl-tRNA synthetases: identification of protein-protein interactions and characterization of a core protein. *Journal of molecular biology* 285, 183-195.

Rabl, J., Leibundgut, M., Ataide, S.F., Haag, A., and Ban, N. (2011). Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* 331, 730-736.

Rackham, O., and Brown, C.M. (2004). Visualization of RNA-protein interactions in living cells: FMRP and IMP1 interact on mRNAs. *The EMBO journal* 23, 3346-3355.

Raff, R.A., and Mahler, H.R. (1972). The non symbiotic origin of mitochondria. *Science* 177, 575-582.

Rahman, B., Kawano, S., Yunoki-Esaki, K., Anzai, T., and Endo, T. (2014). NMR analyses on the interactions of the yeast Tim50 C-terminal region with the presequence and Tim50 core domain. *FEBS letters* 588, 678-684.

Raina, M., Elgamal, S., Santangelo, T.J., and Ibba, M. (2012). Association of a multi-synthetase complex with translating ribosomes in the archaeon *Thermococcus kodakarensis*. *FEBS letters* 586, 2232-2238.

Rajkowitsch, L., Vilela, C., Berthelot, K., Ramirez, C.V., and McCarthy, J.E. (2004). Reinitiation and recycling are distinct processes occurring downstream of translation termination in yeast. *Journal of molecular biology* 335, 71-85.

Ray, P.D., Huang, B.W., and Tsuji, Y. (2012). Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cellular signalling* 24, 981-990.

Reece-Hoyes, J.S., Diallo, A., Lajoie, B., Kent, A., Shrestha, S., Kadreppa, S., Pesyna, C., Dekker, J., Myers, C.L., and Walhout, A.J. (2011). Enhanced yeast one-hybrid assays for high-throughput gene-centered regulatory network mapping. *Nat Methods* 8, 1059-1064.

Rehling, P., Model, K., Brandner, K., Kovermann, P., Sickmann, A., Meyer, H.E., Kuhlbrandt, W., Wagner, R., Truscott, K.N., and Pfanner, N. (2003). Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. *Science* 299, 1747-1751.

Reichert, A.S., and Neupert, W. (2004). Mitochondriomics or what makes us breathe. *Trends in genetics : TIG* 20, 555-562.

Reid, B.G., and Flynn, G.C. (1997). Chromophore formation in green fluorescent protein. *Biochemistry* 36, 6786-6791.

- Reinders, J., and Sickmann, A. (2007). Proteomics of yeast mitochondria. *Methods Mol Biol* 372, 543-557.
- Renault, L., Kerjan, P., Pasqualato, S., Menetrey, J., Robinson, J.C., Kawaguchi, S., Vassilyev, D.G., Yokoyama, S., Mirande, M., and Cherfils, J. (2001). Structure of the EMAPII domain of human aminoacyl-tRNA synthetase complex reveals evolutionary dimer mimicry. *The EMBO journal* 20, 570-578.
- Renvoise, M., Bonhomme, L., Davanture, M., Valot, B., Zivy, M., and Lemaire, C. (2014). Quantitative variations of the mitochondrial proteome and phosphoproteome during fermentative and respiratory growth in *Saccharomyces cerevisiae*. *Journal of proteomics* 106, 140-150.
- Rettig, J., Wang, Y., Schneider, A., and Ochsenreiter, T. (2012). Dual targeting of isoleucyl-tRNA synthetase in *Trypanosoma brucei* is mediated through alternative trans-splicing. *Nucleic acids research* 40, 1299-1306.
- Rho, S.B., Kim, M.J., Lee, J.S., Seol, W., Motegi, H., Kim, S., and Shiba, K. (1999). Genetic dissection of protein-protein interactions in multi-tRNA synthetase complex. *Proceedings of the National Academy of Sciences of the United States of America* 96, 4488-4493.
- Rho, S.B., Lee, K.H., Kim, J.W., Shiba, K., Jo, Y.J., and Kim, S. (1996). Interaction between human tRNA synthetases involves repeated sequence elements. *Proceedings of the National Academy of Sciences of the United States of America* 93, 10128-10133.
- Rho, S.B., Lincecum, T.L., Jr., and Martinis, S.A. (2002). An inserted region of leucyl-tRNA synthetase plays a critical role in group I intron splicing. *The EMBO journal* 21, 6874-6881.
- Ribas de Pouplana, L., and Schimmel, P. (2001). Two classes of tRNA synthetases suggested by sterically compatible dockings on tRNA acceptor stem. *Cell* 104, 191-193.
- Richards, F.M., and Vithayathil, P.J. (1959). The preparation of subtilisin-modified ribonuclease and the separation of the peptide and protein components. *The Journal of biological chemistry* 234, 1459-1465.
- Rinehart, J., Horn, E.K., Wei, D., Soll, D., and Schneider, A. (2004). Non-canonical eukaryotic glutamyl- and glutamyl-tRNA synthetases form mitochondrial aminoacyl-tRNA in *Trypanosoma brucei*. *The Journal of biological chemistry* 279, 1161-1166.
- Rinehart, J., Krett, B., Rubio, M.A., Alfonzo, J.D., and Soll, D. (2005). *Saccharomyces cerevisiae* imports the cytosolic pathway for Gln-tRNA synthesis into the mitochondrion. *Genes & development* 19, 583-592.
- Rizzo, M.A., Springer, G.H., Granada, B., and Piston, D.W. (2004). An improved cyan fluorescent protein variant useful for FRET. *Nature biotechnology* 22, 445-449.
- Rocak, S., Landeka, I., and Weygand-Durasevic, I. (2002). Identifying Pex21p as a protein that specifically interacts with yeast seryl-tRNA synthetase. *FEMS Microbiol Lett* 214, 101-106.
- Rocheleau, J.V., Edidin, M., and Piston, D.W. (2003). Intrasequence GFP in class I MHC molecules, a rigid probe for fluorescence anisotropy measurements of the membrane environment. *Biophys J* 84, 4078-4086.
- Rogers, M.J., and Soll, D. (1988). Discrimination between glutamyl-tRNA synthetase and seryl-tRNA synthetase involves nucleotides in the acceptor helix of tRNA. *Proceedings of the National Academy of Sciences of the United States of America* 85, 6627-6631.
- Roise, D., Horvath, S.J., Tomich, J.M., Richards, J.H., and Schatz, G. (1986). A chemically synthesized pre-sequence of an imported mitochondrial protein can form an amphiphilic helix and perturb natural and artificial phospholipid bilayers. *The EMBO journal* 5, 1327-1334.

Romby, P., Caillet, J., Ebel, C., Sacerdot, C., Graffe, M., Eyermann, F., Brunel, C., Moine, H., Ehresmann, C., Ehresmann, B., *et al.* (1996). The expression of E.coli threonyl-tRNA synthetase is regulated at the translational level by symmetrical operator-repressor interactions. *The EMBO journal* *15*, 5976-5987.

Rook, M.S., Lu, M., and Kosik, K.S. (2000). CaMKIIalpha 3' untranslated region-directed mRNA translocation in living neurons: visualization by GFP linkage. *The Journal of neuroscience : the official journal of the Society for Neuroscience* *20*, 6385-6393.

Rould, M.A., Perona, J.J., Soll, D., and Steitz, T.A. (1989). Structure of E. coli glutamyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 Å resolution. *Science* *246*, 1135-1142.

Roy, H., Becker, H.D., Reinbolt, J., and Kern, D. (2003). When contemporary aminoacyl-tRNA synthetases invent their cognate amino acid metabolism. *Proceedings of the National Academy of Sciences of the United States of America* *100*, 9837-9842.

Roy, H., and Ibba, M. (2008). Monitoring Lys-tRNA(Lys) phosphatidylglycerol transferase activity. *Methods* *44*, 164-169.

Saad, N.Y., Schiel, B., Braye, M., Heap, J.T., Minton, N.P., Durre, P., and Becker, H.D. (2012). Riboswitch (T-box)-mediated control of tRNA-dependent amidation in *Clostridium acetobutylicum* rationalizes gene and pathway redundancy for asparagine and asparaginyl-tRNA synthesis. *The Journal of biological chemistry* *287*, 20382-20394.

Sachs, M.S., and Geballe, A.P. (2006). Downstream control of upstream open reading frames. *Genes & development* *20*, 915-921.

Sagan, L. (1967). On the origin of mitosing cells. *Journal of theoretical biology* *14*, 255-274.

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-gamma and p53 signaling. *Nat Chem Biol* *8*, 547-554.

Sakaue-Sawano, A., Ohtawa, K., Hama, H., Kawano, M., Ogawa, M., and Miyawaki, A. (2008). Tracing the silhouette of individual cells in S/G2/M phases with fluorescence. *Chemistry & biology* *15*, 1243-1248.

Sampath, P., Mazumder, B., Seshadri, V., and Fox, P.L. (2003). Transcript-selective translational silencing by gamma interferon is directed by a novel structural element in the ceruloplasmin mRNA 3' untranslated region. *Molecular and cellular biology* *23*, 1509-1519.

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.

Santos, M.A., Gomes, A.C., Santos, M.C., Carreto, L.C., and Moura, G.R. (2011). The genetic code of the fungal CTG clade. *C R Biol* *334*, 607-611.

Santos, M.A., Moura, G., Massey, S.E., and Tuite, M.F. (2004). Driving change: the evolution of alternative genetic codes. *Trends in genetics : TIG* *20*, 95-102.

Sarkar, J., Poruri, K., Boniecki, M.T., McTavish, K.K., and Martinis, S.A. (2012). Yeast mitochondrial leucyl-tRNA synthetase CP1 domain has functionally diverged to accommodate RNA splicing at expense of hydrolytic editing. *The Journal of biological chemistry* *287*, 14772-14781.

Sauerwald, A., Zhu, W., Major, T.A., Roy, H., Palioura, S., Jahn, D., Whitman, W.B., Yates, J.R., 3rd, Ibba, M., and Soll, D. (2005). RNA-dependent cysteine biosynthesis in archaea. *Science* *307*, 1969-1972.

Schimmel, P., and Ribas de Pouplana, L. (1995). Transfer RNA: from minihelix to genetic code. *Cell* *81*, 983-986.

- Schimmel, P., and Wang, C.C. (1999). Getting tRNA synthetases into the nucleus. *Trends in biochemical sciences* 24, 127-128.
- Schwartz, R.M., and Dayhoff, M.O. (1978). Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science* 199, 395-403.
- Schwille, P., Meyer-Almes, F.J., and Rigler, R. (1997). Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. *Biophys J* 72, 1878-1886.
- Segev, N., and Hay, N. (2012). Hijacking leucyl-tRNA synthetase for amino acid-dependent regulation of TORC1. *Molecular cell* 46, 4-6.
- Sekigawa, M., Kunoh, T., Wada, S., Mukai, Y., Ohshima, K., Ohta, S., Goshima, N., Sasaki, R., and Mizukami, T. (2010). Comprehensive screening of human genes with inhibitory effects on yeast growth and validation of a yeast cell-based system for screening chemicals. *J Biomol Screen* 15, 368-378.
- Sekine, S., Nureki, O., Shimada, A., Vassilyev, D.G., and Yokoyama, S. (2001). Structural basis for anticodon recognition by discriminating glutamyl-tRNA synthetase. *Nat Struct Biol* 8, 203-206.
- Sengupta, S., and Higgs, P.G. (2015). Pathways of Genetic Code Evolution in Ancient and Modern Organisms. *Journal of molecular evolution* 80, 229-243.
- 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.
- Shalak, V., Kaminska, M., Mitnacht-Kraus, R., Vandenabeele, P., Clauss, M., and Mirande, M. (2001). The EMAPII cytokine is released from the mammalian multisynthetase complex after cleavage of its p43/proEMAPII component. *The Journal of biological chemistry* 276, 23769-23776.
- Sheppard, K., Akochy, P.M., Salazar, J.C., and Soll, D. (2007). The *Helicobacter pylori* amidotransferase GatCAB is equally efficient in glutamine-dependent transamidation of Asp-tRNA^{Asn} and Glu-tRNA^{Gln}. *The Journal of biological chemistry* 282, 11866-11873.
- Shimomura, O. (1979). Structure of the chromophore of Aequorea green fluorescent protein. *FEBS Lett* 104, 220-222.
- Shimomura, O., Johnson, F.H., and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusa, Aequorea. *J Cell Comp Physiol* 59, 223-239.
- Shoemaker, C.J., Eyler, D.E., and Green, R. (2010). Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science* 330, 369-372.
- Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proceedings of the National Academy of Sciences of the United States of America* 108, E1392-1398.
- Sickmann, A., Reinders, J., Wagner, Y., Joppich, C., Zahedi, R., Meyer, H.E., Schonfisch, B., Perschil, I., Chacinska, A., Guiard, B., *et al.* (2003). The proteome of *Saccharomyces cerevisiae* mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* 100, 13207-13212.
- Siegel, R.M., Chan, F.K., Zacharias, D.A., Swofford, R., Holmes, K.L., Tsien, R.Y., and Lenardo, M.J. (2000). Measurement of molecular interactions in living cells by fluorescence resonance energy transfer between variants of the green fluorescent protein. *Sci STKE* 2000, pl1.
- Simader, H., Hothorn, M., Kohler, C., Basquin, J., Simos, G., and Suck, D. (2006). Structural basis of yeast aminoacyl-tRNA synthetase complex formation revealed by crystal structures of two binary sub-complexes. *Nucleic acids research* 34, 3968-3979.

- Simos, G., Sauer, A., Fasiolo, F., and Hurt, E.C. (1998). A conserved domain within Arc1p delivers tRNA to aminoacyl-tRNA synthetases. *Molecular cell* *1*, 235-242.
- Simos, G., Segref, A., Fasiolo, F., Hellmuth, K., Shevchenko, A., Mann, M., and Hurt, E.C. (1996). The yeast protein Arc1p binds to tRNA and functions as a cofactor for the methionyl- and glutamyl-tRNA synthetases. *Embo Journal* *15*, 5437-5448.
- Sissler, M., Giege, R., and Florentz, C. (1998). The RNA sequence context defines the mechanistic routes by which yeast arginyl-tRNA synthetase charges tRNA. *Rna* *4*, 647-657.
- Sjostrand, F.S. (1953). Electron microscopy of mitochondria and cytoplasmic double membranes. *Nature* *171*, 30-32.
- Smirnova, E.V., Lakunina, V.A., Tarassov, I., Krasheninnikov, I.A., and Kamenski, P.A. (2012). Noncanonical functions of aminoacyl-tRNA synthetases. *Biochemistry (Mosc)* *77*, 15-25.
- Solimena, M., and Gerdes, H.H. (2003). Secretory granules: and the last shall be first. *Trends in cell biology* *13*, 399-402.
- Soll, D. (1988). Genetic code: enter a new amino acid. *Nature* *331*, 662-663.
- Song, H., Mugnier, P., Das, A.K., Webb, H.M., Evans, D.R., Tuite, M.F., Hemmings, B.A., and Barford, D. (2000). The crystal structure of human eukaryotic release factor eRF1--mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell* *100*, 311-321.
- Sonneborn, T.M. (1965). Nucleotide Sequence of a Gene: First Complete Specification. *Science* *148*, 1410.
- Souciet, G., Menand, B., Ovesna, J., Cosset, A., Dietrich, A., and Wintz, H. (1999). Characterization of two bifunctional *Arabidopsis thaliana* genes coding for mitochondrial and cytosolic forms of valyl-tRNA synthetase and threonyl-tRNA synthetase by alternative use of two in-frame AUGs. *Eur J Biochem* *266*, 848-854.
- Sprinzi, M., and Cramer, F. (1975). Site of aminoacylation of tRNAs from *Escherichia coli* with respect to the 2'- or 3'-hydroxyl group of the terminal adenosine. *Proceedings of the National Academy of Sciences of the United States of America* *72*, 3049-3053.
- Srinivasan, G., James, C.M., and Krzycki, J.A. (2002). Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* *296*, 1459-1462.
- Stansfield, I., Jones, K.M., Kushnirov, V.V., Dagkesamanskaya, A.R., Poznyakovski, A.I., Paushkin, S.V., Nierras, C.R., Cox, B.S., Ter-Avanesyan, M.D., and Tuite, M.F. (1995). The products of the SUP45 (eRF1) and SUP35 genes interact to mediate translation termination in *Saccharomyces cerevisiae*. *The EMBO journal* *14*, 4365-4373.
- Stathopoulos, C., Jacquin-Becker, C., Becker, H.D., Li, T., Ambrogelly, A., Longman, R., and Soll, D. (2001). *Methanococcus jannaschii* prolyl-cysteinyl-tRNA synthetase possesses overlapping amino acid binding sites. *Biochemistry* *40*, 46-52.
- Stroud, D.A., Oeljeklaus, S., Wiese, S., Bohnert, M., Lewandrowski, U., Sickmann, A., Guiard, B., van der Laan, M., Warscheid, B., and Wiedemann, N. (2011). Composition and topology of the endoplasmic reticulum-mitochondria encounter structure. *Journal of molecular biology* *413*, 743-750.
- Subach, F.V., Subach, O.M., Gundorov, I.S., Morozova, K.S., Piatkevich, K.D., Cuervo, A.M., and Verkhusha, V.V. (2009). Monomeric fluorescent timers that change color from blue to red report on cellular trafficking. *Nat Chem Biol* *5*, 118-126.
- Sugiyama, M., Akase, S.P., Nakanishi, R., Horie, H., Kaneko, Y., and Harashima, S. (2014). Nuclear localization of Haa1, which is linked to its phosphorylation status, mediates lactic acid tolerance in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* *80*, 3488-3495.

- Sung, M.K., and Huh, W.K. (2007). Bimolecular fluorescence complementation analysis system for in vivo detection of protein-protein interaction in *Saccharomyces cerevisiae*. *Yeast* 24, 767-775.
- Sung, M.K., Lim, G., Yi, D.G., Chang, Y.J., Yang, E.B., Lee, K., and Huh, W.K. (2013). Genome-wide bimolecular fluorescence complementation analysis of SUMO interactome in yeast. *Genome research* 23, 736-746.
- Tardiff, D.F., Jui, N.T., Khurana, V., Tambe, M.A., Thompson, M.L., Chung, C.Y., Kamadurai, H.B., Kim, H.T., Lancaster, A.K., Caldwell, K.A., *et al.* (2013). Yeast reveal a "druggable" Rsp5/Nedd4 network that ameliorates alpha-synuclein toxicity in neurons. *Science* 342, 979-983.
- Tatsuta, T., Augustin, S., Nolden, M., Friedrichs, B., and Langer, T. (2007). m-AAA protease-driven membrane dislocation allows intramembrane cleavage by rhomboid in mitochondria. *The EMBO journal* 26, 325-335.
- Taylor, A.B., Smith, B.S., Kitada, S., Kojima, K., Miyaura, H., Otwinowski, Z., Ito, A., and Deisenhofer, J. (2001). Crystal structures of mitochondrial processing peptidase reveal the mode for specific cleavage of import signal sequences. *Structure* 9, 615-625.
- Theobald-Dietrich, A., Frugier, M., Giege, R., and Rudinger-Thirion, J. (2004). Atypical archaeal tRNA pyrrolysine transcript behaves towards EF-Tu as a typical elongator tRNA. *Nucleic acids research* 32, 1091-1096.
- Tolerico, L.H., Benko, A.L., Aris, J.P., Stanford, D.R., Martin, N.C., and Hopper, A.K. (1999). *Saccharomyces cerevisiae* Mod5p-II contains sequences antagonistic for nuclear and cytosolic locations. *Genetics* 151, 57-75.
- 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. *The Journal of biological chemistry* 275, 35063-35069.
- Treusch, S., Hamamichi, S., Goodman, J.L., Matlack, K.E., Chung, C.Y., Baru, V., Shulman, J.M., Parrado, A., Bevis, B.J., Valastyan, J.S., *et al.* (2011). Functional links between Abeta toxicity, endocytic trafficking, and Alzheimer's disease risk factors in yeast. *Science* 334, 1241-1245.
- Triana-Alonso, F.J., Chakraborty, K., and Nierhaus, K.H. (1995). The elongation factor 3 unique in higher fungi and essential for protein biosynthesis is an E site factor. *The Journal of biological chemistry* 270, 20473-20478.
- Trotta, C.R., Miao, F., Arn, E.A., Stevens, S.W., Ho, C.K., Rauhut, R., and Abelson, J.N. (1997). The yeast tRNA splicing endonuclease: a tetrameric enzyme with two active site subunits homologous to the archaeal tRNA endonucleases. *Cell* 89, 849-858.
- Truscott, K.N., Kovermann, P., Geissler, A., Merlin, A., Meijer, M., Driessen, A.J., Rassow, J., Pfanner, N., and Wagner, R. (2001). A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. *Nat Struct Biol* 8, 1074-1082.
- Tsien, R., and Prasher, D. (1997). In *GFP: Green Fluorescent Protein Strategies and Applications*. New York: Wiley & Sons.
- Tumbula, D.L., Becker, H.D., Chang, W.Z., and Soll, D. (2000). Domain-specific recruitment of amide amino acids for protein synthesis. *Nature* 407, 106-110.
- Turakhiya, U., von der Malsburg, K., Gold, V.A.M., Guiard, B., Chacinska, A., van der Laan, M., and Ieva, R. (2016). Protein Import by the Mitochondrial Presequence Translocase in the Absence of a Membrane Potential. *Journal of molecular biology* 428, 1041-1052.
- Turner, R.J., Lovato, M., and Schimmel, P. (2000). One of two genes encoding glycyl-tRNA synthetase in *Saccharomyces cerevisiae* provides mitochondrial and cytoplasmic functions. *The Journal of biological chemistry* 275, 27681-27688.

- Ullmann, A., Jacob, F., and Monod, J. (1967). Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the beta-galactosidase structural gene of *Escherichia coli*. *Journal of molecular biology* 24, 339-343.
- Ussery, M.A., Tanaka, W.K., and Hardesty, B. (1977). Subcellular distribution of aminoacyl-tRNA synthetases in various eukaryotic cells. *Eur J Biochem* 72, 491-500.
- Valencia-Burton, M., McCullough, R.M., Cantor, C.R., and Broude, N.E. (2007). RNA visualization in live bacterial cells using fluorescent protein complementation. *Nat Methods* 4, 421-427.
- van der Blik, A.M., and Koehler, C.M. (2003). A mitochondrial rhomboid protease. *Dev Cell* 4, 769-770.
- Van Engelenburg, S.B., and Palmer, A.E. (2010). Imaging type-III secretion reveals dynamics and spatial segregation of *Salmonella* effectors. *Nat Methods* 7, 325-330.
- van Hoof, A., Lennertz, P., and Parker, R. (2000). Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Molecular and cellular biology* 20, 441-452.
- van Rooyen, J.M., Murat, J.-B., Hammoudi, P.-M., Kieffer-Jaquinod, S., Coute, Y., Sharma, A., Pelloux, H., Belrhali, H., and Hakimi, M.-A. (2014). Assembly of the Novel Five-Component Apicomplexan Multi-Aminoacyl-tRNA Synthetase Complex Is Driven by the Hybrid Scaffold Protein Tg-p43. *PLoS one* 9, e89487.
- Vennegoor, C.J., Stols, A.L., and Bloemendal, H. (1972). More evidence for a particle character of aminoacyl-transfer RNA synthetases isolated from rat liver. *Journal of molecular biology* 65, 375-378.
- Verkhusha, V.V., Chudakov, D.M., Gurskaya, N.G., Lukyanov, S., and Lukyanov, K.A. (2004). Common pathway for the red chromophore formation in fluorescent proteins and chromoproteins. *Chemistry & biology* 11, 845-854.
- Vogt, W. (1995). Oxidation of methionyl residues in proteins: tools, targets, and reversal. *Free radical biology & medicine* 18, 93-105.
- Vogtle, 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.
- Vyas, K., Chaudhuri, S., Leaman, D.W., Komar, A.A., Musiyenko, A., Barik, S., and Mazumder, B. (2009). Genome-wide polysome profiling reveals an inflammation-responsive posttranscriptional operon in gamma interferon-activated monocytes. *Molecular and cellular biology* 29, 458-470.
- Wachter, R.M., Elsliger, M.A., Kallio, K., Hanson, G.T., and Remington, S.J. (1998). Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. *Structure* 6, 1267-1277.
- 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., Cheresch, D.A., and Schimmel, P. (2002). A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 99, 173-177.
- Waldo, G.S., Standish, B.M., Berendzen, J., and Terwilliger, T.C. (1999). Rapid protein-folding assay using green fluorescent protein. *Nature biotechnology* 17, 691-695.
- Wallin, I. (1927). *Symbiogenesis and the origin of species* London, UK: Bailliere, Tindall and Co.
- Wang, S., and Hazelrigg, T. (1994). Implications for bcd mRNA localization from spatial distribution of exu protein in *Drosophila oogenesis*. *Nature* 369, 400-403.

- Ward, W.W., and Bokman, S.H. (1982). Reversible denaturation of Aequorea green-fluorescent protein: physical separation and characterization of the renatured protein. *Biochemistry* *21*, 4535-4540.
- Watford, M. (1991). The urea cycle: a two-compartment system. *Essays Biochem* *26*, 49-58.
- Wehrman, T., Kleaveland, B., Her, J.H., Balint, R.F., and Blau, H.M. (2002). Protein-protein interactions monitored in mammalian cells via complementation of beta -lactamase enzyme fragments. *Proceedings of the National Academy of Sciences of the United States of America* *99*, 3469-3474.
- Wei, N., Shi, Y., Truong, L.N., Fisch, K.M., Xu, T., Gardiner, E., Fu, G., Hsu, Y.S., Kishi, S., Su, A.I., *et al.* (2014). Oxidative Stress Diverts tRNA Synthetase to Nucleus for Protection against DNA Damage. *Molecular cell*.
- Weil, T.T., Forrest, K.M., and Gavis, E.R. (2006). Localization of bicoid mRNA in late oocytes is maintained by continual active transport. *Dev Cell* *11*, 251-262.
- Wiedemann, N., van der Laan, M., Hutu, D.P., Rehling, P., and Pfanner, N. (2007). Sorting switch of mitochondrial presequence translocase involves coupling of motor module to respiratory chain. *The Journal of cell biology* *179*, 1115-1122.
- Wilcox, M., and Nirenberg, M. (1968). Transfer RNA as a cofactor coupling amino acid synthesis with that of protein. *Proceedings of the National Academy of Sciences of the United States of America* *61*, 229-236.
- Wilson, D.N., and Doudna Cate, J.H. (2012). The structure and function of the eukaryotic ribosome. *Cold Spring Harb Perspect Biol* *4*.
- Wilusz, J.E., Whipple, J.M., Phizicky, E.M., and Sharp, P.A. (2011). tRNAs marked with CCACCA are targeted for degradation. *Science* *334*, 817-821.
- Woese, C.R. (1965). On the evolution of the genetic code. *Proceedings of the National Academy of Sciences of the United States of America* *54*, 1546-1552.
- Wong, F.C., Beuning, P.J., Silvers, C., and Musier-Forsyth, K. (2003). An isolated class II aminoacyl-tRNA synthetase insertion domain is functional in amino acid editing. *The Journal of biological chemistry* *278*, 52857-52864.
- Wong, J.T. (1975). A co-evolution theory of the genetic code. *Proceedings of the National Academy of Sciences of the United States of America* *72*, 1909-1912.
- Yadavalli, S.S., and Ibba, M. (2012). Quality control in aminoacyl-tRNA synthesis its role in translational fidelity. *Adv Protein Chem Struct Biol* *86*, 1-43.
- Yang, D., Oyaizu, Y., Oyaizu, H., Olsen, G.J., and Woese, C.R. (1985). Mitochondrial origins. *Proceedings of the National Academy of Sciences of the United States of America* *82*, 4443-4447.
- Yang, F., Moss, L.G., and Phillips, G.N., Jr. (1996). The molecular structure of green fluorescent protein. *Nature biotechnology* *14*, 1246-1251.
- 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.
- Yao, P., and Fox, P.L. (2012). A truncated tRNA synthetase directs a "translational trickle" of gene expression. *Cell Cycle* *11*, 1868-1869.
- Yao, P., Potdar, A.A., Arif, A., Ray, P.S., Mukhopadhyay, R., Willard, B., Xu, Y., Yan, J., Saidel, G.M., and Fox, P.L. (2012). Coding region polyadenylation generates a truncated tRNA synthetase that counters translation repression. *Cell* *149*, 88-100.

- Yeung, M.L., Houzet, L., Yedavalli, V.S., and Jeang, K.T. (2009). A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. *The Journal of biological chemistry* 284, 19463-19473.
- Yoo, C.J., and Wolin, S.L. (1997). The yeast La protein is required for the 3' endonucleolytic cleavage that matures tRNA precursors. *Cell* 89, 393-402.
- Yoshihisa, T., Yunoki-Esaki, K., Ohshima, C., Tanaka, N., and Endo, T. (2003). Possibility of cytoplasmic pre-tRNA splicing: the yeast tRNA splicing endonuclease mainly localizes on the mitochondria. *Molecular biology of the cell* 14, 3266-3279.
- Young, J.C., Hoogenraad, N.J., and Hartl, F.U. (2003). Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. *Cell* 112, 41-50.
- Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H., and Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 Å resolution. *Science* 292, 883-896.
- Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomov, S., Kisselev, L., and Philippe, M. (1995). Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. *The EMBO journal* 14, 4065-4072.
- Zhu, X., Liu, Y., Yin, Y., Shao, A., Zhang, B., Kim, S., and Zhou, J. (2009). MSC p43 required for axonal development in motor neurons. *Proceedings of the National Academy of Sciences of the United States of America* 106, 15944-15949.
- Zimyanin, V.L., Belaya, K., Pecreaux, J., Gilchrist, M.J., Clark, A., Davis, I., and St Johnston, D. (2008). In vivo imaging of oskar mRNA transport reveals the mechanism of posterior localization. *Cell* 134, 843-853.
- Zywicky, M., Bakowska-Zywicka, K., and Polacek, N. (2012). Revealing stable processing products from ribosome-associated small RNAs by deep-sequencing data analysis. *Nucleic acids research* 40, 4013-4024.

Ingénierie d'un outil basé sur une GFP fragmentée pour l'étude des protéines multi-localisées chez les eucaryotes



Les aminoacyl-ARNt synthétases catalysent la formation des aminoacyl-ARNt, utilisés lors de la synthèse protéique et peuvent également former des complexes multi-synthétasiques (MSC). Chez *S. cerevisiae*, le complexe AME associe les glutamyl- et méthionyl-ARNt synthétases à la protéine d'ancrage Arc1 et joue un rôle primordial dans la coordination de l'expression des génomes nucléaire et mitochondrial. Tous les composants de ce MSC sont multi-localisés et assurent des fonctions essentielles dans d'autres compartiments. Pour étudier ces localisations multiples, nous avons élaboré un outil, basé sur la Split-GFP, qui nous permet de visualiser spécifiquement la fraction organellaire d'une protéine multi-localisée. Pour cela, la GFP a été séparée en deux fragments : i) β 1-10, restreint à un compartiment subcellulaire et ii) β 11, fusionné aux protéines d'intérêts. Cet outil nous a permis d'étudier diverses relocalisations, ainsi que de délimiter des signaux d'import.

Mots-clés : Aminoacyl-ARNt synthétases, *S. cerevisiae*, Split-GFP, localisation subcellulaire, mitochondries

Aminoacyl-tRNA synthetases catalyze aminoacyl-tRNA formation, required for protein synthesis but can also associate into multi-synthetase complexes (MSC). In *S. cerevisiae*, the AME complex contains glutamyl- and methionyl-tRNA synthetases bound to the anchor protein Arc1 and is responsible for the coordination of nuclear and mitochondrial genome expression. The three MSC partners are multi-localized and present simultaneously in several compartments. The detection of the organellar pools of these multi-localized proteins *in vivo* is difficult, since they are mainly cytosolic. Therefore, we engineered a split-GFP based localization tool that allows us to specifically visualize organellar fractions of multi-localized proteins. To do so, GFP was split into two parts: β 1-10, restricted to a subcellular compartment and β 11, fused to the protein of interest. This tool allowed us to study relocalization of cytosolic proteins and characterize targeting signals.

Key-words: Aminoacyl-tRNA synthetases, *S. cerevisiae*, Split-GFP, subcellular localization, mitochondria