HAL
open science

Etude d'un nouveau type de RNase P spécifique des eucaryotes chez Arabidopsis thaliana

Bernard Gutmann

To cite this version:

Bernard Gutmann. Etude d'un nouveau type de RNase P spécifique des eucaryotes chez Arabidopsis thaliana. Génomique, Transcriptomique et Protéomique [q-bio.GN]. Université de Strasbourg, 2012. Français. NNT : 2012STRAJ133 . tel-01159764

HAL Id: tel-01159764

https://theses.hal.science/tel-01159764

Submitted on 3 Jun 2015

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

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

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE 414

UPR 2357

THÈSE présentée par :

Bernard GUTMANN

soutenue le : **14 septembre 2012**

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

Discipline/ Spécialité : Aspects moléculaires et cellulaires de la biologie

Etude d'un nouveau type de RNase P spécifique des eucaryotes chez *Arabidopsis thaliana*

THÈSE dirigée par :

Dr. GIEGÉ Philippe	Docteur, IBMP-CNRS (Université de Strasbourg)

RAPPORTEURS :

Dr. LURIN Claire	Docteur, URGV-INRA-CNRS (Université d'Évry Val d'Essonne)
Prof. SCHROEDER Renée	Professeur, Max F Perutz Laboratories (Universität Wien)

AUTRES MEMBRES DU JURY :

Prof. Otten Léon	Professeur, IBMP-CNRS (Université de Strasbourg)

REMERCIEMENTS

Je tiens tout d'abord à remercier les Docteurs Claire Lurin et Ivan Tarassov, ainsi que les Professeurs Renée Schroeder et Léon Otten pour avoir accepté de juger mon travail de thèse.

Je remercie mon directeur de thèse Philippe Giegé, de m'avoir accueilli dans son laboratoire, ainsi que pour sa confiance et ses encouragements.

J'adresse tous mes remerciements à Anthony Gobert pour son soutien, ses conseils et sa patience qui m'ont été d'une aide inestimable.

Je souhaite également remercier tous mes collègues et amis de l'Institut de Biologie Moléculaire des Plantes avec qui j'ai partagé des moments agréables, avec une mention spéciale pour les membres du "labo 306", sans qui la vie de tous les jours aurait sans doute été bien terne et certains samedis soirs trop reposants.

Enfin je remercie mon entourage, famille et amis, pour leur soutien au quotidien.

RESUME

La RNase P est impliquée dans la maturation des ARNt en libérant l'extrémité 5' leader de leurs précurseurs. Jusqu'à très récemment il était admis que cette enzyme soit universellement conservée en tant que complexe ribonucléoprotéique dont l'ARN, un ribozyme, possède l'activité catalytique. L'activité RNase P existe chez les plantes dans les trois compartiments possédant une expression génétique. Cependant les gènes codant les composants de la RNase P ribonucléoprotéique ne sont pas présents dans le génome des plantes.
La rupture avec le modèle établi est venue avec l'identification d'une RNase P uniquement protéique dans les mitochondries humaines et chez les plantes. Ces protéines ont été nommées au laboratoire PRORP (pour « Proteinaceous RNase P »). Elles sont caractérisées par la présence de motifs PPR qui pourraient être impliqués dans la liaison à l'ARN et d'un domaine métallonucléase NYN sûrement impliqué dans l'activité catalytique de la protéine. Chez *Arabidopsis*, trois paralogues ont été localisés soit dans les organelles (PRORP1), soit dans le noyau (PRORP2 et 3).

Des tests d'activité *in vitro* montrent que chaque protéine PRORP possède seule une activité RNase P. L'étude de lignées ADN-T a montré que la fonction des protéines PRORP est essentielle et la fonction des protéines nucléaires PRORP2 et 3 est redondante.
L'hypothèse est que les protéines PRORP soient entièrement responsables de l'activité RNase P chez les plantes, i.e. qu'elles aient complètement remplacés les ribonucléoprotéines pour l'activité RNase P. Cependant, la RNase MRP, une autre ribonucléoprotéine impliquée dans la maturation des ARNr, est présente dans le noyau des cellules de plante. Chez les animaux et les levures, cette enzyme contient des sous-unités également associées à la RNase P ribonucléoprotéique. Ainsi nous ne pouvions exclure que la RNase MRP ait acquis au cours de l'évolution l'activité RNase P chez les plantes. Pour répondre à ces questions nous avons construit des lignées de dérégulation de PRORP1, PRORP2/3 et de sous unités de la RNase MRP. Nous avons trouvé que les protéines PRORP sont bien impliquées dans l'activité RNase P alors que la RNase MRP n'est pas impliquée dans la maturation des ARNt.

Ces résultats suggèrent que les protéines PRORP seraient bien les seules enzymes impliquées dans la maturation de l'extrémité 5' des ARNt chez *Arabidopsis*. De plus l'analyse des lignées de dérégulation montre que les protéines PRORP possèdent une diversité de substrat. Ainsi, en plus de la maturation des ARNt, PRORP1 est impliquée dans la maturation d' ARNm mitochondriaux en réalisant une coupure endonucléolytique au niveau d'une structure ARNt-like, tandis que PRORP2 et 3 interviennent indirectement dans la maturation de snoARN en clivant les transcrits nucléaires dicistroniques ARNt-snoARN.

ABRÉVIATIONS

Ω	ohm
ABAD	amyloid-beta peptide-binding alcohol dehydrogenase
AcSG	acétosyringone
ADN	acide désoxyribonucléique
ADNc	ADN complémentaire
ADN-T	ADN du plasmide Ti
AGO	Argonaute
amiARN	miARN arificiel
ARN	acide ribonucléique
ARNm	ARN messager
ARNr	ARN ribosomique
ARNt	ARN de transfert
Asn	Asparagine
Asp	Aspartate
A. thaliana	*Arabidopsis thaliana*
ATP	adénosine triphosphate
atp 6	ATP synthase subunit 6
BET	bromure d'éthidium
BSA	Bovine Serum Albumin
B. subtilis	*Bacillus subtilis*
C	cytidine
CaCl2	chlorure de calcium
ccmC	cytochrome c maturation C
Ci	Curie
CMS	stérilité mâle cytoplasmique
Col-0	Columbia-0
cox1	cytochrome c oxidase subunit I
CRR	chlororespiratory reduction (CRR4, CRR21)
cRT-PCR	circular reverse transcription PCR
CsCl	chlorure de césium
Ct	crossing threshold
Da, kDa	Dalton, kilodalton
DAPI	4', 6' diamidino-2- phénylindole
DCL	Dicer-like
DNase	désoxyribonucléase
dNTP	2'-désoxyribonucléoside-5'-triphosphate
DTT	dithiothreitol
DYW	motif aspartate tyrosine tryptophane
E	efficacité (qRT-PCR)
E. coli	*Escherichia coli*
eIF4E	eukaryotic translation initiation factor 4E
EDTA	éthylène diamine tétra-acétate de sodium
E-YFP	enhanced yellow fluorescent protein
µF	micro-farad
g, mg	gramme, milligramme
µg, ng	microgramme, nanogramme
GlmS	glucosamine-6P synthase
Gln	Glutamine
Gly	Glycine
h	heure
HCl	acide chlorhydrique
His	histidine
HYL1	Hyponastic leaves1
Ile	Isoleucine
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITS	internal transcribed spacer

kb, Mb	kilobase, megabase
KCl	chlorure de potassium
KH_2PO_4	phosphate de potassium monobasique
KO	knock out
L, ml, µl	litre, millilitre, microlitre
LB	Luria Broth
Lys	Lysine
M, mM, µM	molaire, millimolaire, micromolaire
mA	milliampère
MES	2-(N-Morpholino)ethanesulfonic acid
Met-e	Méthionine élongateur
$MgCl_2$	chlorure de magnésium
miARN	microARN
min	minute
MOPS	3-(*N*-morpholino)propanesulfonic acid
MRPP	Mitochondrial RNase P Protein
MS	Murashige et Skoog
Mth	*Methanothermobacter thermoautotrophicus*
NaCl	chlorure de sodium
nad1, ND1	NADH dehydrogenase subunit 1 (sous-unité du complexe I de la chaine respiratoire)
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
Na_2HPO_4	sodium phosphate dibasique
ndhD	sous-unité D de la NAD(P)H déshydrogénase
NEP	nuclear-encoded polymérase
NH_4Ac	acétate d'ammonium
Ni-NTA	nickel nitrilotriacetic acid
No-0	Nössen-0
NOR	nucleolus organizer région
nt	nucléotide
NYN	N4BP1, YacP-like Nuclease
ORF	open reading frame
OTP43	organelle transcript processing defect 43
PBS	phosphate buffered saline
PCR	polymerase chain réaction
PDS	phytoène désaturase
PEG-CMS	polyéthylène glycol chlorométhyl styrène
PEP	platid-encoded RNA polymerase
Phe	Phénylalanine
PLS	sous-famille des PPR possédant les motifs P-S-L
PNK	polynucleotidyl kinase
pol	polymérase
Pop	processing of precursor RNA (gène codant des sous-unités protéiques de la RNase P et la RNase MRP chez les levures et les animaux et de la RNase MRP chez les plantes)
PPR	pentatricopeptide repeat
PPR E	sous-groupe des protéines PPR PLS possédant des motifs E et E+
PPR-like L	Long motif PPR de 35 à 36 acides aminés
PPR-like S	Short motif PPR de 31 acides aminés
PPR P	PPR-like motif PPR canonique de 35 acides aminés
pre-tmRNA	transfer-messenger RNA precursor
Pro	Proline
PRORP	proteinaceous RNase P
PVC	*Planctomycetes*, *Verrucomicrobia*, *Chlamydiae*
PVP40	polyvinylpyrrolidone 40
qRT-PCR	quantitative real-time PCR
région C	C pour catalytic ; région responsable de l'activité catalytique
région S	S pour specificity ; région portant le site de liaison au substrat
Rf	Restaurateur de fertilité

RIP-Chip	RNA ImmunoPrecipitation, suivie d'hybridation sur puce à ADN
RISC	RNA-induced silencing complex
RNase	ribonucléase
RNase P	P pour précurseur
RNase MRP	RNase mitochondrial RNA processing
RNP	ribonucléoprotéine
rNTP	ribonucléotide triphosphate
RPF	RNA processing factor
Rpm1r	mitochondrial RNase P RNA
Rpm2p	mitochondrial RNase P protein
RpoTm	ARN polymérase mitochondriale de type phagique
RpoTmp	ARN polymérase mitochondriale et plastidiale de type phagique
RpoTp	ARN polymérase plastidiale de type phagique
RPP	RNase P protein
RPR	RNase P RNA
Rpr2p	RNase P ribonucleoprotein-2 protein (chez *Saccharomyces cerevisiae*)
RT-PCR	reverse transcriptase polymerase chain reaction
S	Svedberg
S. cerevisiae	Saccharomyces cerevisiae
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	seconde
Ser	Sérine
siARN	small interfering RNA
site A	A pour aminoacyl, site présent dans la grande sous-unité ribosomique
site E	E pour exit, site présent dans la grande sous-unité ribosomique
site P	P pour peptidyl, site présent dans la grande sous-unité ribosomique
Snm1p	suppressor of nuclear mitochondrial endoribonuclease 1 protein
snoARN	small nucleolar RNA
snoR43	snoARN R43
snRNP	small nuclear ribonucleoprotein
SSC	saline-sodium citrate
TAE	Tris/Acétate/EDTA
TBE	Tris/Borate/EDTA
TCA	acide trichloroacétique
TEMED	tétraméthyléthylène diamine
TPR	tetratricopeptide repat
Trm	tRNA methyltransferase (Trm10p)
TRV	Tobacco Rattle Virus
Tyr	Tyrosine
U	unité enzymatique
U	uridine
UCE	upstream control element
UV	ultraviolet
V, kV	volt, kilovolt
Val	Valine
VIGS	Virus Induced Gene Silencing
WMD	Web MicroRNA Designer
WS-4	Wassilewskja-4

SOMMAIRE

1	INTRODUCTION	1
1.1	LES COMPARTIMENTS PRESENTANT UNE EXPRESSION GENETIQUE CHEZ LES PLANTES	1
1.1.1	Origine et évolution	1
1.1.2	Les génomes nucléaires et organellaires	2
1.1.3	Fonctions du noyau et des organites	4
1.2	L'EXPRESSION DES GENOMES CHEZ LA PLANTE ET SA REGULATION	5
1.2.1	L'expression du génome nucléaire de plante	5
1.2.1.1	La transcription et la régulation de l'expression des gènes	5
1.2.1.2	Modifications post-transcriptionnelles	6
1.2.1.2.1	L'ajout d'une coiffe en 5'	6
1.2.1.2.2	L'ajout d'une queue poly A	7
1.2.1.2.3	Epissage	7
1.2.1.2.4	Coupure endonucléolytique	8
1.2.2	L'expression des génomes organellaires de plante	8
1.2.2.1	La transcription	8
1.2.2.2	Modifications post-transcriptionnelles	9
1.2.2.2.1	L'édition des ARN	10
1.2.2.2.2	L'épissage des ARN	10
1.2.2.2.3	Les extrémités des transcrits	11
1.2.2.2.4	Les coupures endonucléolytiques	12
1.3	LES PROTEINES A "PENTATRICOPEPTIDE REPEAT"	12
1.3.1	Les différentes classes de protéines PPR	13
1.3.2	Distribution des protéines PPR dans le vivant	13
1.3.3	L'implication des protéines PPR dans les mécanismes post-transcriptionnels	14
1.3.3.1	L'édition de l'ARN	14
1.3.3.2	L'épissage	15
1.3.3.3	Les coupures endonucléolytiques	16
1.4	LES RIBOZYMES ET RIBONUCLEOPROTEINES (RNP)	16
1.4.1	Les virus et viroïdes	17
1.4.2	Bactéries	17
1.4.3	Eucaryotes	18
1.5	RNASE P	19
1.5.1	Fonction de la RNase P	19
1.5.1.1	Mécanisme cinétique dans la RNP	19
1.5.1.2	Diversité de substrat pour la RNase P	20
1.5.2	Présence dans les différents phyla	21
1.5.2.1	Procaryotes	21
1.5.2.1.1	Bactéries	21
1.5.2.1.2	*Archaea*	22
1.5.2.2	Eucaryotes	24
1.5.2.2.1	RNP nucléaire	24
1.5.2.2.1.1	Levure	24
1.5.2.2.1.2	Animaux	25
1.5.2.2.2	RNP mitochondriale	26

1.5.2.2.3 RNase P protéique 26
1.5.2.2.3.1 Animaux 26
1.5.2.2.3.2 Plantes 27

2 OBJECTIFS DE MA THESE 29

3 MATERIELS ET METHODES 30
3.1 MATERIELS 30
3.1.1 Souches bactériennes 30
3.1.2 Plantes 30
3.1.3 Plasmides 31
3.1.4 Amorces nucléotidiques 32
3.1.5 Plasmides disponibles 37

3.2 METHODES 37
3.2.1 ADN 37
3.2.1.1 Amplification par PCR 37
3.2.1.2 Électrophorèse sur gel d'agarose 38
3.2.1.3 Digestion et ligation de fragments d'ADN 38
3.2.1.4 Construction amiARN 39
3.2.1.5 Transformation de bactéries par électroporation 39
3.2.1.6 Transformation des bactéries par choc thermique 40
3.2.1.7 Préparation d'ADN plasmidique par lyse alcaline 40
3.2.1.8 Séquençage d'ADN 40
3.2.1.9 Extraction de l'ADN total de feuille d'*Arabidopsis* 40
3.2.1.10 Mutagenèse dirigée 41
3.2.2 ARN 42
3.2.2.1 Extraction d'ARN total de plante 42
3.2.2.2 Transcription *in vitro* 43
3.2.2.3 Marquage en 5' de l'ARN par la polynucléotidyle kinase 43
3.2.2.4 Electrophorèse sur gel de polyacrylamide 44
3.2.2.5 Electrophorèse sur gel d'agarose 44
3.2.2.6 Analyse de type "northern blot" 45
3.2.2.7 Purification des transcrits extraits sur gel de polyacrylamide 46
3.2.2.8 RT-PCR pour l'obtention d'ADNc total 46
3.2.2.9 CRT-PCR 46
3.2.2.10 QRT-PCR 47
3.2.3 Protéines 48
3.2.3.1 Expression dans E. coli et purification 48
3.2.3.2 Dosage des protéines par la méthode de Bradford 49
3.2.3.3 Electrophorèse sur gel de polyacrylamide en conditions dénaturantes (SDS-PAGE) 50
3.2.3.4 Coloration du gel de polyacrylamide 50
3.2.3.5 Transfert des protéines sur membrane 51
3.2.3.6 Immunodétection de type Western 51
3.2.3.7 Test d'activité 52
3.2.4 Plantes 52
3.2.4.1 Expression transitoire
dans des protoplastes issus de cellules mésophylle d'*A. thaliana* 52
3.2.4.2 Virus Induced Gene Silencing (VIGS) 54
3.2.4.3 Transformation de plante par floraldip 54
3.2.4.4 Stérilisation des graines 55

3.2.5 Microscopie 55
3.2.5.1 Observation au microscope confocal des protoplastes transformés afin de déterminer la localisation des protéines de fusion PRORP-EYFP 55
3.2.5.2 Observation au stéréomicroscope 56

4 RESULTATS 57
4.1 LOCALISATION DES PROTEINES PRORP ET FONCTION DE PRORP1 57
4.1.1 Préambule 57
4.1.2 Publication 1 : A single *Arabidopsis* organellar protein has RNase P activity 57

4.2 FONCTION *IN VITRO* DES PROTEINES PRORP NUCLEAIRES ET *IN PLANTA* DES TROIS PROTEINES PRORP 59
4.2.1 Préambule 59
4.2.2 Publication 2 : PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis* 59
4.3 AUTRES OUTILS DEVELOPPES LORS DES TRAVAUX DE THESE 61
4.3.1 Outils pour l'identification des substrats PRORP 61
4.3.1.1 VIGS 61
4.3.1.2 Lignées de surexpression 62
4.3.1.3 Lignées miARN 62
4.3.2 Outil pour l'identification du mode d'action des protéines PRORP 62

5 DISCUSSION 64
5.1 LA VIE SANS RNASE P RIBONUCLEOPROTEIQUE 64
5.1.1 Les protéines PRORP ont une activité RNase P 64
5.1.2 Les protéines PRORP sont essentielles 64
5.1.3 Les protéines PRORP sont seules responsables de l'activité RNase P 65

5.2 PRORP REFLETE-T-ELLE UNE DIVERSITE DE MODE D'ACTION PARMI LES PROTEINES PPR ? 66
5.2.1 Une nouvelle catégorie de protéines PPR 66
5.2.2 Mode de reconnaissance des ARN par PRORP 66
5.2.3 Substrat minimal 67

5.3 L'EVOLUTION DES RNASE P 68
5.3.1 La perte du ribozyme de la RNase P 68
5.3.2 *Arabidopsis* est-elle une exception chez les plantes terrestre ? 69
5.3.3 Diversité de l'évolution de la RNase P ? 70

5.4 INTEGRATION DE LA FONCTION DES PROTEINES PRORP AVEC D'AUTRES MECANISMES D'EXPRESSION GENETIQUE 70

ANNEXES 73

BIBLIOGRAPHIE 74

1 INTRODUCTION

Le sujet de ma thèse est l'étude de la RNase P nucléaire et organellaire chez les plantes. Dans ces organismes, l'activité RNase P est portée par un nouveau type de RNase P purement protéique, qui appartient à la famille des protéines à "pentatricopeptide repeat" (PPR).

Après quelques généralités sur les mécanismes d'expression génétique dans les différents compartiments où l'activité RNase P a lieu, je présenterai une introduction décrivant l'état des connaissances sur les protéines PPR en général ainsi que sur les RNase P et leur distribution en tant que ribonucléoprotéine ou protéine.

1.1 LES COMPARTIMENTS PRÉSENTANT UNE EXPRESSION GENETIQUE CHEZ LES PLANTES

Les plantes sont des organismes autotrophes, dont les cellules possèdent un noyau et deux organites semi-autonomes, grâce à la présence de leur propre génome, la mitochondrie et le chloroplaste.

1.1.1 ORIGINE ET ÉVOLUTION

La formation du noyau est le prélude de la formation des eucaryotes, cependant cette étape est encore mal connue et actuellement quatre théories essayent de répondre à son origine.

La première énonce une endosymbiose entre une *Archaea* méthanogènique et une myxobactérie (Lopez-Garcia and Moreira, 2006). Elle s'appuie sur une analyse des génomes eucaryotiques montrant la présence de gènes d'origine bactérienne et archéal dans le génome nucléaire.

Cependant cette hypothèse est en contradiction avec les observations faites sur *Gemmata obscuriglobus* (Fuerst, 2005). Dans ces bactéries, le cytoplasme est divisé en plusieurs compartiments possédant une ou plusieurs membranes. Le nucleoïde est entouré de deux membranes et possède une structure proche du noyau des eucaryotes (Fuerst, 2005).

Une autre théorie, qui reste très controversée, propose la présence persistante de virus dans un mycoplasme méthanogénique qui « phagocyterait » des bactéries (Bell, 2001). Cette triple association aurait évolué pour former le noyau, les bactéries « phagocytées »

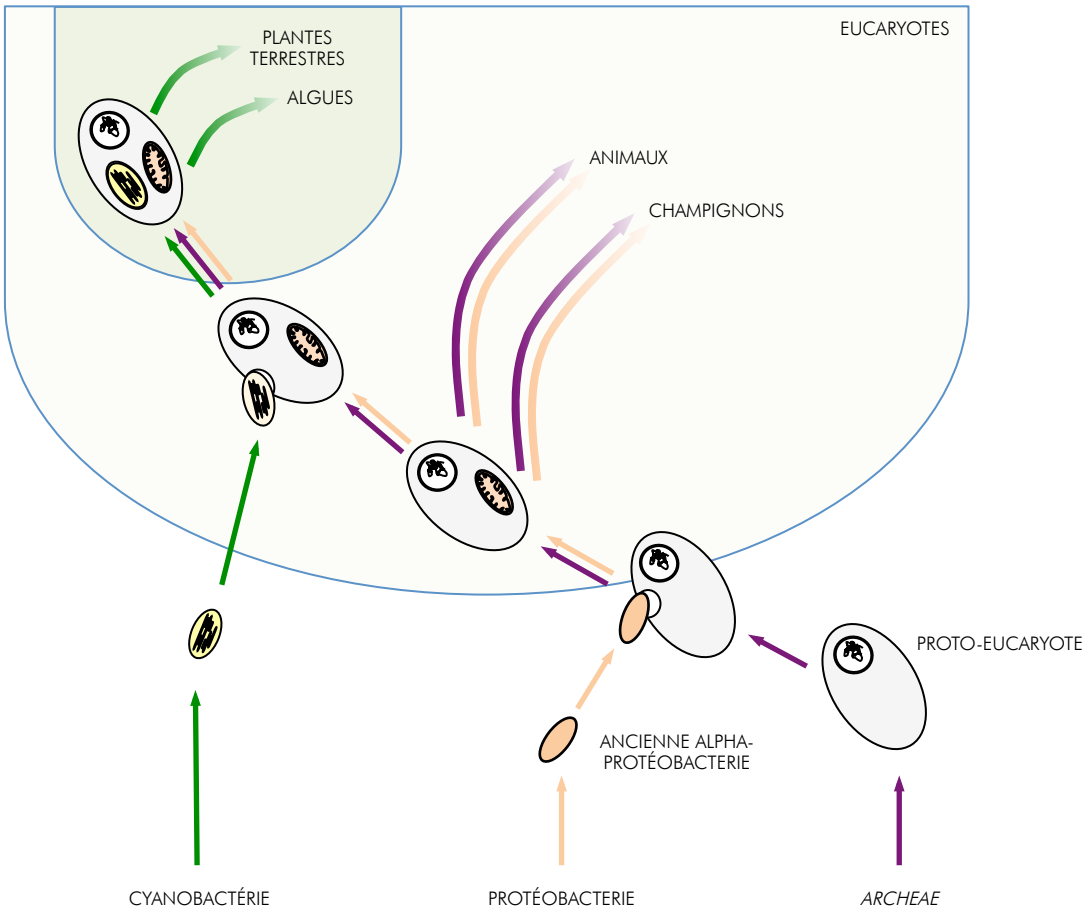


Figure 1 : LES ÉVÉNEMENTS D'ENDOSYMBIOSES MENANT AUX PLANTES SUPÉRIEURES (Timmis et al., 2004)

La formation du noyau, qui ferait intervenir une archeae, est une étape encore mal connue.

Deux étapes d'endosymbiose sont à l'origine de la formation des organelles. Lors de la première endosymbiose, un eucaryote unicellulaire (proto-eucaryote) a incorporé une α-protéobactérie. La symbiose entre ces deux organismes a conduit à la formation d'une proto-mitochondrie ayant évoluée en une mitochondrie. La seconde endosymbiose, concernant un eucaryote issu de la première endosymbiose et une cyanobactérie, a formé un plaste. Cet événement a conduit à la séparation entre d'une part les plantes terrestres et les algues, possédant les deux organelles, et d'autre part les animaux et les champignons ayant uniquement les mitochondries.

fournissant l'ATP, les lipides, les acides aminés et les nucléotides et les virus des fonctions spécifiques aux cellules eucaryotiques comme le capping des ARNm ou la séparation entre transcription et traduction.
Enfin une dernière hypothèse (Forterre, 2011) reprend les trois théories et considère la formation d'un symbionte entre une thaumarchaeon et une bactérie du super-embranchement PVC (*Planctomycetes*, *Verrucomicrobia* et *Chlamydiae*) suivie d'une invasion virale.

La mitochondrie et le chloroplaste ont évolué par des étapes d'endosymbiose (figure 1). Il y a environ 1.5 milliards d'années une cellule eucaryote primitive a incorporé une α-protéobactérie (Gray, 1999). Cette endosymbiose primaire, à l'origine de la mitochondrie, a engendré le premier eucaryote unicellulaire hétérotrophe. Une endosymbiose secondaire entre un eucaryote et une cyanobactérie a eu lieu il y a 1.6 à 1 milliards d'années (Raven and Allen, 2003). Cet événement, conduisant à la formation des plastes, est à l'origine du règne végétal.

Au cours de l'évolution un transfert massif de gènes a eu lieu entre ces deux symbiontes et le noyau de la cellule hôte (Timmis et al., 2004), permettant le contrôle des organelles par leur hôte et la protection de leur génome contre les radicaux libres (Allen and Raven, 1996). Ainsi 18 % des gènes nucléaires d'*A. thaliana* semblent être issus d'une cyanobactérie donc du proto-plaste (Martin et al., 2002). Les protéines, issues de ces gènes nucléaires et essentielles aux fonctions organellaires, sont synthétisées dans le cytosol et importées dans les mitochondries ou les chloroplastes par deux systèmes d'importation indépendants. Cependant, dans certains cas, des protéines synthétisées par un même gène et possédant des séquences de double adressage peuvent être importées aux deux organites, comme par exemple l'aminoacyl ARNt synthétase (Duchene et al., 2005).
On peut noter qu'un transfert inverse de gène a aussi eu lieu au cours de l'évolution dans la mitochondrie. En effet 4 % du génome mitochondrial serait représenté par des séquences résiduelles de rétrotransposons d'origine nucléaire (Unseld et al., 1997).

1.1.2 LES GÉNOMES NUCLÉAIRES ET ORGANELLAIRES

La taille des génomes nucléaires végétaux sont très variables et peuvent aller jusqu'à une taille de 16000 Mb pour *Triticum aestivum*, alors que le génome d'*A. thaliana* ne

représente que 125 Mb à 157 Mb (Arabidopsis Genome Initiative, 2000; Bennett et al., 2003). Ce dernier est constitué de cinq chromosomes, possédant 24 grands segments dupliqués de 100 kb ou plus représentant 58 % du génome (Arabidopsis Genome Initiative, 2000). Ils seraient issus d'événements de duplication totale ou partielle du génome au cours de l'évolution (Simillion et al., 2002). Le génome nucléaire renferme un peu plus de 25000 gènes avec une densité variant entre 1 à 38 gènes par 100 kb (Arabidopsis Genome Initiative, 2000). En plus des ARNm à l'origine de protéines cytosoliques, membranaires et organellaires, le génome d'*Arabidopsis* code pour de nombreux ARN non codant. Par exemple, 599 ARNt nucléaires et 24 ARNt dérivant des organites (23 mitochondriaux et 1 plastidial) (Michaud et al., 2011) sont synthétisés à partir du génome nucléaire. C'est le chromosome 1 qui possède une grande partie de ces gènes qui sont divisés en deux groupes, l'un est constitué d'une répétition de 27 $ARNt^{Pro}$ et le second de 27 répétitions en tandem des gènes $ARNt^{Tyr}$ $ARNt^{Tyr}$ $ARNt^{Ser}$ (Theologis et al., 2000). Les ARNt importés dans la mitochondrie sont codés en grande partie par le chromosome 2 qui possède 270 kb du génome mitochondrial, soit 75 % (Lin et al., 1999; Michaud et al., 2011). Chez les plantes les ARNt peuvent être transcrits sous la forme d'un précurseur dicistronique ARNt-snoARN, comme par exemple chez le riz certains $ARNt^{Met\text{-}e}$ ou chez *Arabidopsis* certains $ARNt^{Gly}$ (Michaud et al., 2011). Les snoARN, présents sur ces transcrits dicistroniques, font uniquement partie de la famille des snoARN à boite C/D (Kruszka et al., 2003). Les gènes codant pour les ARN ribosomaux (ARNr) sont présents dans les organisateurs du nucléole (NORs). Ils sont codés sous la forme d'un précurseur, l'ARNr 35S, qui va subir différents événements de maturation afin de donner les ARNr 25S, 18S et 5.8S (Copenhaver and Pikaard, 1996). Les ARNr 5S sont codés par des séquences répétées présententes dans les régions centromériques des chromosomes 3, 4 et 5 (Arabidopsis Genome Initiative, 2000; Vaillant et al., 2007)}. Le génome nucléaire code également pour des molécules d'ARN simple brin, de 20 à 25 nucléotides, appelées microARN (miARN). Ils interviennent dans le développement (Zhu et al., 2011) et différents mécanismes biologiques (Chen et al., 2011) ou de stress (Iyer et al., 2012). Pour cela il régule l'expression de gènes cibles en recrutant la machinerie du silencing renfermant en particulier les protéines Dicer-like (DCL) et Argonautes (AGO) (Fang and Spector, 2007).

Chez les plantes, le génome mitochondrial a une taille comprise entre 180 et 2400 kb. Chez *Arabidopsis*, ce génome circulaire de 367 kb code pour 57 gènes, soit en moyenne un gène tous les 8 kb (Unseld et al., 1997). Il est donc environ deux fois moins dense que le génome nucléaire. Ainsi, 62 % du génome mitochondrial est composé de régions à

fonction inconnue et pourrait servir de région tampon permettant au génome de subir des mutations en particulier par recombinaison. L'ADN mitochondrial possède 4 % de séquences répétées non recombinantes présentant 90 % d'homologie et deux séquences répétées et recombinantes de 6.5 kb et de 4.2 kb conduisant à la formation de deux molécules circulaires de 233 kb et 134 kb. Trois ARNr (18S, 26S et 5S) et 22 ARNt mitochondriaux participent à la synthèse de 38 protéines. Cependant le nombre d'ARNt mitochondriaux n'est pas suffisant pour la lecture de tous les codons. Pour suppléer à ce défaut, les ARNt codant pour six acides aminés manquant sont codés par le génome nucléaire et importés par les mitochondries (Salinas et al., 2008).

Le génome chloroplastique d'*Arabidopsis thaliana* est plus petit que le mitochondrial (Sato et al., 1999). Il est circulaire et possède une taille de 155 kb. Il est séparé en deux parties (84 kb et 18 kb) par la présence d'une paire de séquences répétées et inversées de 27 kb. Ce génome possède 132 gènes dont certains sont dupliqués et code pour 79 protéines, les ARNr (16S, 23S, 4.5S, 5S), présents dans les zones répétées et inversées, et pour 30 ARNt représentant 20 acides aminés. Les gènes sont organisés sous la forme de 13 clusters, correspondant à deux ou plusieurs gènes adjacents.

1.1.3 FONCTIONS DU NOYAU ET DES ORGANITES

L'enveloppe nucléaire régule la transcription en activant certains facteurs de transcription en réponse à des stress (Seo et al., 2008). De plus, elle sépare la transcription et la traduction. Avant d'être exportés dans le cytoplasme, les ARN vont subir différentes modifications post-transcriptionnelles, telles que l'ajout de coiffe et de queue polyA, l'épissage et le clivage des ARN. Un ARN exporté dans le cytoplasme et ne possédant pas ces modifications, ne sera pas pris en charge par la machinerie traductionnelle et sera dégradé (Fasken and Corbett, 2005).

Les deux symbiotes vont fournir des fonctions supplémentaires à la cellule. Grâce à la mitochondrie les cellules vont pouvoir synthétiser leur propre énergie. La mitochondrie renferme le cycle de Krebs permettant la synthèse du NADH, nécessaire à la phosphorylation oxydative conduisant à la synthèse d'ATP. Cela permettra la création d'eucaryotes multicellulaires, dans lesquels une seconde fonction est devenue importante pour la différenciation ou la défense de la cellule lors d'une infection, i.e. la mort cellulaire programmée (Scott and Logan, 2008). On la retrouve aussi impliquée dans les

voies de signalisation grâce à la régulation de la concentration en calcium (Logan and Knight, 2003).

Les chloroplastes vont transformer les organismes hétérotrophes en autotrophes, grâce à la photosynthèse qui est divisée en deux étapes. La première phase fait intervenir les photosystèmes permettant la synthèse de molécule de NADPH et d'ATP. Ces molécules vont être prises en charge par le cycle de Calvin pour fixer le carbone du CO_2 afin de former des molécules organiques.

1.2 L'EXPRESSION DES GENOMES CHEZ LA PLANTE ET SA REGULATION

1.2.1 L'EXPRESSION DU GÉNOME NUCLÉAIRE DE PLANTE

1.2.1.1 LA TRANSCRIPTION ET LA REGULATION DE L'EXPRESSION DES GENES

Dans le génome nucléaire de plante, il existe une transcription monocistronique ou polycistronique. Elle nécessite cinq ARN polymérase (pol), qui interviennent dans la transcription par exemple des ARNm (pol II), ARNr (pol I et III), ARNt (pol III), miARN (pol II) et siARN (pol IV et pol V). Ce mécanisme nécessite la présence de séquence consensus en amont du site d'initiation de la transcription. Ainsi on retrouve la boite TATA dans la majorité des gènes codant pour des protéines. Elle est située à environ 30 paires de bases du site d'initiation et est le lieu de fixation de l'ARN pol II. Généralement deux boites supplémentaires (une boîte CAT et un domaine riche en GC) sont présentes entre les positions -40 et -200 afin de réguler la transcription. En amont de cette région d'autres éléments, appelés activateurs ou bloqueurs, peuvent moduler la transcription, en recrutant des protéines, appelées facteurs de transcription, augmentant ou diminuant la transcription (Deppmann et al., 2006).

Au contraire l'activité de l'ARN polymérase I nécessite la présence de deux éléments, un core dont sa présence est indispensable pour avoir une transcription *in vitro* et une upstream control element (UCE) (Moss et al., 2007). En ce qui concerne la transcription des ARNt et de l'ARNr 5S, l'ARN pol III reconnaît des promoteurs intragéniques (Bogenhagen et al., 1980; Sakonju et al., 1980; Galli et al., 1981; Sharp et al., 1981).

En plus des facteurs de transcription, des modifications épigénétiques peuvent intervenir afin de réguler l'expression d'un gène. Elles peuvent se faire à deux niveaux,

soit sur les protéines associées à l'ADN, les histones, soit directement sur la molécule d'ADN.

Les histones subissent des modifications post-traductionnelles (acétylations, méthylations ou réactions inverses) qui permettent de condenser ou décondenser la chromatine (Berr et al., 2011). Ainsi ils vont soit relâcher la molécule d'ADN et donc libérer l'accès aux facteurs de transcription ou au contraire restreindre l'accès en condensant l'ADN. L'ADN subit également des méthylations qui vont diminuer l'expression des gènes touchés. Chez les plantes cette méthylation est dû à des siARN dérivés de miARN (Chellappan et al., 2010; Wu et al., 2010).

Les miARN régulent également l'expression des gènes en recrutant la machinerie de silencing qui va soit dégrader le transcrit cible (Ronemus et al., 2006), soit inhiber la traduction de ce transcrit (Brodersen et al., 2008).

1.2.1.2 MODIFICATIONS POST-TRANSCRIPTIONNELLES

Il existe un grand nombre de mécanismes post-transcriptionnels essentiels à l'expression génétique. Parmi eux, on peut citer des modifications impliquées dans les processus de maturation et de stabilité des ARN comme les modifications des extrémités 5' et 3' (clivage endonucléolytique, coiffe en 5', queue polyA en 3') ou encore l'épissage.

1.2.1.2.1 L'AJOUT D'UNE COIFFE EN 5'

L'ajout d'une coiffe en 5' spécifique au noyau est nécessaire à l'export des ARNm en se liant au cap-binding complex (Nojima et al., 2007), à leur stabilité en les protégeant contre l'action des 5'-exonuclease (Shimotohno et al., 1977) et à leur traduction en interagissant avec le facteur d'initiation eIF4E qui recrute indirectement le ribosome (Mitchell et al., 2010).

Cette modification de l'extrémité 5' de l'ARNm néosynthétisé fait intervenir trois étapes successives. Une ARN triphosphatase supprime un phosphate à l'extrémité 5', permettant l'ajout par la guanylyltranférase d'un résidu guanosine. Ce résidu sera ensuite méthylé en position 7 par une méthyltransferase (Banerjee, 1980) afin d'obtenir une extrémité 5' 7mGpppN.

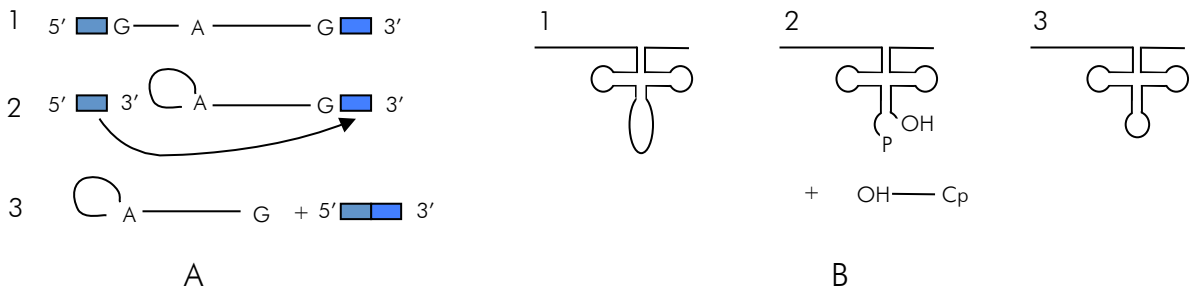


Figure 2 : MÉCANISMES D'ÉPISSAGE DES ARNm ET DES ARNt NUCLÉAIRES DE PLANTES (Cech, 1990)

L'épissage des ARNm fait intervenir un résidu adénylique (A). Une transestérification entre cette adénosine et un groupement guanylique présent sur le site de branchement de l'intron libérera le premier exon (A2). L'extrémité 3′ de l'intron est ensuite libérée, puis les deux exons sont assemblés par hydrolyse d'une molécule d'ATP (A3). Toutes les étapes sont réalisées par des complexes ribonucleoprotéiques (snRNP) associés à des facteurs protéiques.

L'épissage des ARNt (B) fait intervenir une endonucléase qui va libérer l'intron (B2). Les deux exons sont associés par une ligase puis le 2′P présent à la jonction va être supprimé par une phosphatase (B3).

1.2.1.2.2 L'AJOUT D'UNE QUEUE POLY A

Chez les eucaryotes, comme la coiffe, la queue polyA est nécessaire au maintien de l'intégrité de l'ARNm (Moore and Proudfoot, 2009). La polyadénylation consiste en l'addition d'une queue polyA en l'extrémité 3' du transcrit néo-synthétisé. Cette modification est réalisée par la poly(A) polymérase canonique qui ajoute des adénosines monophosphates en utilisant comme substrat l'ATP (Edmonds and Abrams, 1960; Chen et al., 2010). Des protéines se liant à la queue polyA, les PABP (poly A binding protein), interviennent dans l'export, la stabilité et la traduction des ARNm polyadénylés (Kuhn and Wahle, 2004) (Huang and Carmichael, 1996).

Cependant il existe un phénomène inverse dans le noyau des eucaryotes. En effet la polyadénylation de certains ARN par une polyA polymérase non canonique va les orienter vers la dégradation via l'exosome (Schmid and Jensen, 2008).

1.2.1.2.3 EPISSAGE

Dans le noyau des cellules végétales, l'épissage des ARNm concerne les introns splicéosome-dépendant (figure 2A). Ce mécanisme fait intervenir le splicéosome qui est constitué de cinq petits complexes nucléoprotéiques appelés snRNP (small nuclear ribonucleoprotein), associés à plusieurs facteurs protéiques. Il se déroule en trois étapes. Une transestérification libère l'extrémité 5' de l'intron qui se lie à une adénosine présente sur le site de branchement de l'intron créant une forme intermédiaire de lasso. L'extrémité 3' est ensuite clivé et les exons libérés sont ligués par hydrolyse de l'ATP (Lorkovic et al., 2000).

Les précurseurs des $ARNt^{Met\text{-}e}$ et $ARNt^{Tyr}$ nucléaires possèdent un intron dans la boucle anti-codon (figure 2B). Leur épissage requiert la présence de trois enzymes : une endonucléase pour le clivage de l'intron au niveau des deux sites d'épissage, une ARN ligase pour reformer les deux extrémités de l'ARNt et une 2'-phosphotransferase pour supprimer le 2' phosphate présent à la jonction d'épissage (Abelson et al., 1998).

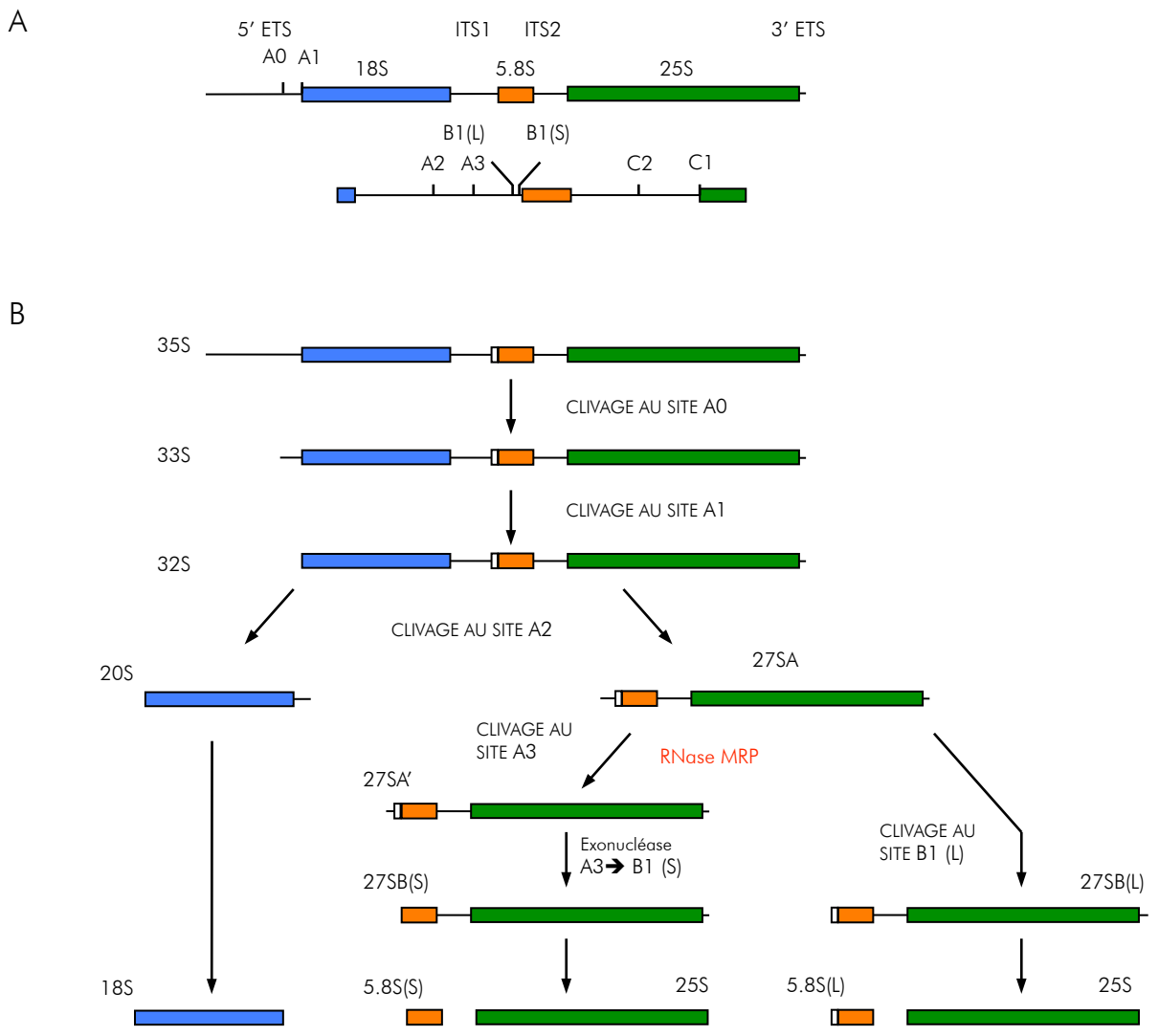


Figure 3 : ETAPES DE MATURATION DES ARNr NUCLÉAIRES CHEZ *S. CEREVISIAE* (Lygerou et al. 1994)

L'ARN polymérase I synthétise le pre-ARNr 35S (A). (B) Il va subir trois étapes de maturation successives correspondant à des clivages endonucléolytiques en A0 dans le 5′ ETS, en A1 en 5′ du 18S et en A2 divisant le preARNr en deux intermédiaires de maturation (ARNr 20S et 27SA). L'ARNr 20S subit une coupure endonucléotidique à l'origine de l'ARNr 18S. L'ARNr 27SA peut être maturé par deux voies alternatives. Le 27SA peut subir une coupure endonucléolytique au site A3 par la RNase MRP, avant d'être digéré par une exonucléase jusqu'au site B1(S). Cet intermédiaire 27B(S) est ensuite maturé afin de donner l'ARNr 5.8S(S) et 25S. L'autre voie alternative consiste à un clivage endonucléolytique au site B1(L), puis à une maturation finale afin d'obtenir l'ARNr 5.8S(L) et 25S.

1.2.1.2.4 COUPURE ENDONUCLEOLYTIQUE

Chez les eucaryotes, de nombreuses endonucléases interviennent dans la maturation des ARN. On citera par exemple la RNase MRP (mitochondrial RNA processing) qui est présente dans le noyau et la mitochondrie chez les mammifères et la levure (Kiss et al., 1992; Reilly and Schmitt, 1995), alors qu'elle est détectée uniquement dans le noyau chez les plantes (Kiss et al., 1992). Elle intervient dans la maturation des ARNr cytosolique qui sont transcrits sous la forme d'un précurseur polycistronique 35S qui renferme les ARNr 18S, 5.8S et 25S (figure 3). Les pré-ARNr 18S et 5.8S ainsi que 5.8S et 25S sont séparés par l'ITS1 (internal transcribed spacer) et l'ITS2. L'ITS1 renferme le site de A3 qui subit un clivage endonucléolytique par la RNase MRP chez la levure (Allmang et al., 1996). Chez les plantes, le site de clivage de la RNase MRP n'a pas été identifié. Cependant elle cosédimenterait, dans des extraits de feuilles de tabac avec les précurseurs ribosomaux (Kiss et al., 1992).

Des coupures endonucléolytiques sont également nécessaires pour la maturation des miARN. Ces molécules sont synthétisées sous la forme d'un transcrit primaire (primiARN) en tige boucle qui sera protégé par des protéines de coiffe (Kim et al., 2008). Le primiARN sera clivé par un complexe constitué de deux protéines : DCL1 un homologue de la protéine Dicer humaine qui a une activité endonucléase (Kurihara and Watanabe, 2004) et HYL1 nécessaire à une activité efficace de DCL1 (Kurihara et al., 2006; Szarzynska et al., 2009). DCL1 va réaliser deux clivages endonucléolytique de l'ARN double brin, un premier en amont du pri-miARN et un second au niveau de la boucle afin d'obtenir un miARN mature (Kurihara and Watanabe, 2004).

Après son export vers le cytosol, le miARN va être incorporé dans le complexe RISC (RNA-induced silencing complex) et va servir de patron lors de la reconnaissance des ARNm cibles. Ces derniers seront alors clivés par la protéine AGO1 (Vaucheret et al., 2004).

1.2.2 L'EXPRESSION DES GÉNOMES ORGANELLAIRES DE PLANTE

1.2.2.1 LA TRANSCRIPTION

Dans les mitochondries, la transcription est réalisée par des ARN polymérases de type phagique synthétisées par le noyau : les NEP (nuclear-encoded polymerase) (figure 4).

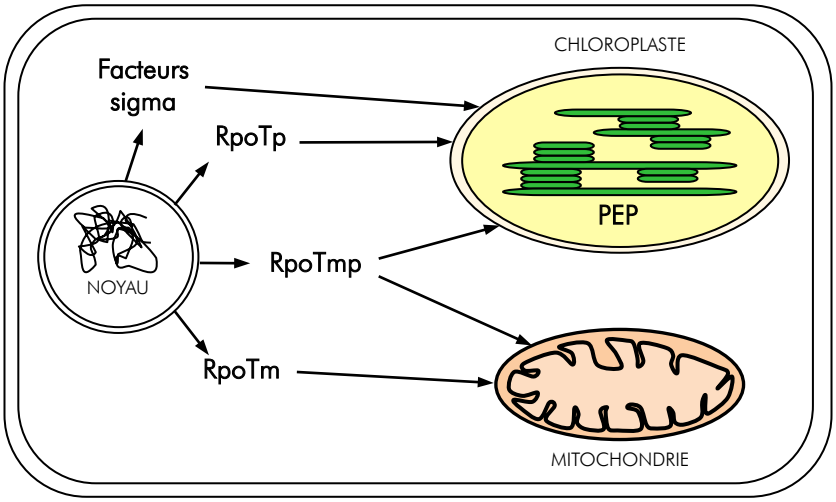


Figure 4 : MACHINERIE TRANSCRIPTIONNELLE DANS LES ORGANELLES D'*A. THALIANA* (Liere et al., 2011)

Le génome nucléaire code pour trois ARN polymerases de type phagique, qui sont importées soit dans le chloroplaste (RpoTp), soit dans la mitochondrie (RpoTm), soit doublement localisées (RpoTmp). Le chloroplaste code pour une ARN polymerase de type eubactérienne. Une transcription spécifique par la PEP nécessite la présence de facteurs sigma, codés par le génome nucléaire.

Trois gènes nucléaires codent pour ces ARN polymérases chez *Arabidopsis*, qui seront adressées à la mitochondrie RpoTm, au chloroplaste RpoTp ou doublement adressées RpoTmp (Chang et al., 1999; Hedtke et al., 2000). Alors que RpoTm intervient dans la transcription de la plupart des gènes mitochondriaux (Kuhn et al., 2007), la fonction de RpoTmp reste incomprise malgré le fait qu'il a été montré récemment qu'elle intervient dans la transcription de certains gènes mitochondriaux (Kuhn et al., 2009). Pour reconnaître la séquence promotrice ces polymérases nécessitent la présence de facteurs supplémentaires dans les mitochondries de levures et de vertébrés (Tracy and Stern, 1995) (Shutt et al., 2010). Ces promoteurs sont distribués de manière homogène, peut-être pour éviter une diminution de la transcription en raison du grand nombre de recombinaisons. Ainsi l'ensemble du génome est transcrit et la régulation se fait majoritairement au niveau post-transcriptionnel (Holec et al., 2006).

Dans le chloroplaste, la situation est différente (figure 4). En plus des NEP importées à partir du cytosol, le génome chloroplastique code pour une ARN polymérase de type eubactérienne, la PEP (plastid-encoded polymerase) (Sato et al., 1999). La reconnaissance des promoteurs par la PEP se fait de manière spécifique par l'intervention de facteurs sigma codés par le noyau (Nagashima et al., 2004b; Nagashima et al., 2004a; Ishizaki et al., 2005). Ces promoteurs de type bactérien possèdent deux séquences conservées TATAAT et TTGACA respectivement en position -10 et -35 (Weihe, 2004). L'expression des facteurs sigma, régulant la transcription, est différente en fonction du cycle circadien (Ichikawa et al., 2008) du développement (Kasai et al., 2004) ou du tissu (Isono et al., 1997a; Isono et al., 1997b). La régulation de l'expression des gènes se fait aussi au niveau traductionnel, la lumière étant un des signaux majeurs de contrôle de l'expression (Marin-Navarro et al., 2007).

1.2.2.2 MODIFICATIONS POST-TRANSCRIPTIONNELLES

Comme les ARN nucléaires, les ARN organellaires de plantes subissent un grand nombre de modifications post-transcriptionnnelles, dont la plupart leur sont spécifiques, dans leur rôle ou leur mécanisme. On se concentrera sur l'édition, l'épissage, l'ajout de queue polyA et certaines coupures endonucléolytiques.

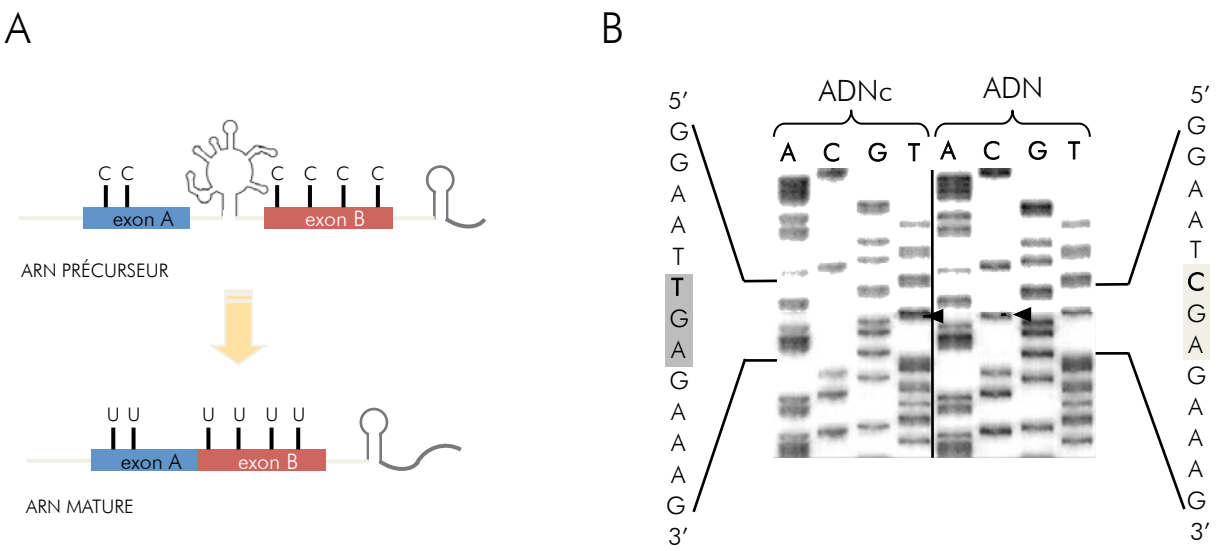


Figure 5 : L'ÉDITION DE L'ARN DANS LES ORGANELLES DE PLANTE (Gualberto et al., 1989)

A. L'édition de l'ARN consiste en des modifications de cytidines en uridines à des positions précises des transcrits.

B. Mise en évidence d'un site d'édition de l'ARN. La séquence nucléotidique d'un gène (ADN) est comparée à la séquence nucléotidique de l'ADN complémentaire obtenue à partir du transcrit de ce gène (ADNc). Les triangles noirs montrent la différence de séquence (C dans l'ADN et T dans l'ADNc). Ceci indique la modification d'un codon CGA en codon UGA au niveau de l'ARN.

1.2.2.2.1 L'EDITION DES ARN

Le terme d'édition a été employé pour la première fois chez le trypanosome pour désigner des insertions et plus rarement des délétions d'uridine dans des ARNm mitochondriaux (Benne, 1994). Cette appellation a ensuite été reprise dans le cas des plantes terrestres pour décrire le mécanisme de déamination de cytidines (C) en uridines (U) (figure 5) et plus rarement une conversion inverse d'U en C (Gualberto et al., 1989), (Pring et al., 1993), (Covello and Gray, 1989), (Gualberto et al., 1990). Les sites d'édition sont beaucoup plus fréquents dans le transcriptome mitochondrial (400 à 500 sites) (Unseld et al., 1997; Giege and Brennicke, 1999) que dans le transcriptome chloroplastique (34 sites) (Chateigner-Boutin and Small, 2007).

Ce mécanisme peut être considéré comme correcteur de mutations. En effet celui-ci peut restaurer un codon d'initiation ou de terminaison (Chaudhuri et al., 1995), est nécessaire à l'épissage de certains introns (Castandet et al., 2010) ou à la maturation en 5' et 3' des ARNt (Marechal-Drouard et al., 1996), mais surtout permet de recouvrer des codons codant pour des acides aminés conservés dans les protéines respectives (Hiesel et al., 1989). Ainsi, il permet la traduction correcte des ARNm et est nécessaire pour obtenir des protéines fonctionnelles.

Le mécanisme en lui-même reste obscur. Un complexe est recruté et activé grâce à la présence de séquence cis intervenant dans la trans-activation d'éléments en trans. Jusqu'à récemment, seule une hélicase à ARN avait était proposée comme élément de ce mécanisme (Takenaka and Brennicke, 2003). Cependant un grand nombre de travaux montre maintenant l'implication des protéines PPR dans ce mécanisme (voir ci-dessous).

1.2.2.2.2 L'EPISSAGE DES ARN

Dans les organelles de plantes, l'épissage concerne des introns de groupe I et II (figure 6). Les introns du groupe I sont présents principalement dans les chloroplastes. En effet dans le génome mitochondrial tous les introns sont de groupe II à l'exception d'un seul intron de groupe I présent dans le gène *cox1* de certaines plantes (Cho et al., 1998). Cependant il n'y a pas de corrélation phylogénétique. Cet intron serait en effet issu d'un transfert génétique horizontal. Les introns de type II sont présents en très grand nombre dans les organelles des plantes terrestres. Ils sont

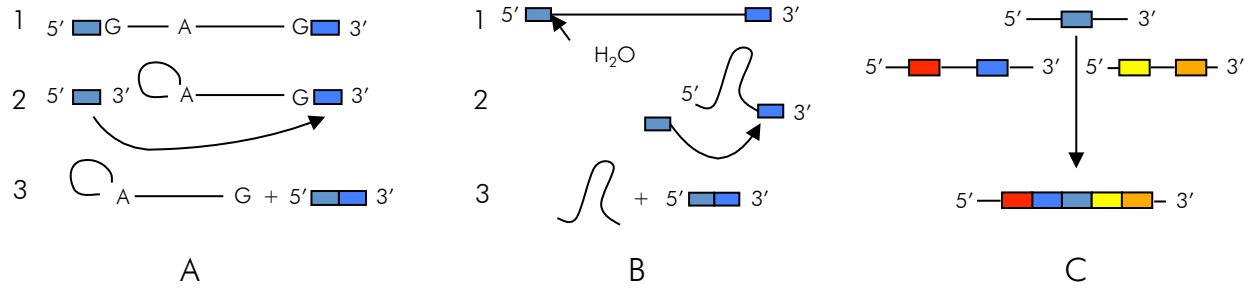


Figure 6 : MÉCANISMES D'ÉPISSAGE DANS LES ORGANELLES DE PLANTES (Cech, 1990)

L'épissage en cis consiste en l'épissage d'un intron séparant deux exons sur une même molécule d'ARN. Cet épissage fait intervenir un résidu adénylique (A1) dans le cas des introns de type II ou guanylique pour les introns de type I. Une transestérification entre cette adénosine et un groupement guanylique présent en amont de l'intron libérera le premier exon (A2). Une attaque nucléophile suivra permettant la fusion des deux exons (A3). En absence du résidu adénylique (B1), une molécule d'H_2O peut jouer le même rôle que le groupement adénylique libérant le premier exon (B2). L'étape suivante sera comme dans le premier cas (A) une attaque nucléophile entre les deux exons (B3).

L'épissage en trans permet de réunir des exons présents sur des molécules d'ARN différentes. L'ARNm mature est obtenu à partir de trois précurseurs, via deux épissages en cis et deux en trans (C).

constitués de six domaines leur conférant une structure secondaire conservée, facilitant l'épissage. La différence entre ces deux groupes réside principalement dans le mécanisme d'épissage, constitué de deux étapes (Cech, 1990). Dans les introns de groupe II, la première étape fait intervenir un site de branchement interne à l'intron possédant un résidu adénylique.

La transestérification, entre le 2'H de l'adénosine et le 5'P d'un résidu guanylique situé au niveau de l'extrémité 5' de l'intron, formera une liaison phosphodiester. La deuxième étape consiste en une attaque nucléophile entre les deux exons, libérant le 3'OH de l'intron. En absence de site de branchement, l'hydrolyse par une molécule d'H_2O libère le premier exon (figure 6B). Dans les introns du groupe I, l'adénosine sera remplacée par une guanosine ou un nucléotide guanylique (figure 6A) (Cech, 1990).

Chez certaines bactéries et champignons, des introns autocatalytiques codent pour des maturases nécessaires à leur épissage. De telles maturases sont aussi trouvées dans les introns de groupe II chez les plantes. Cependant, des études sur *Chlamydomonas reinhardtii* et *Arabidopsis* montrent que les protéines nécessaires à l'épissage de ces introns dans les organelles, comme par exemple les protéines PPR (voir-ci dessous), sont codées par le génome nucléaire (Herrin and Nickelsen, 2004).

1.2.2.2.3 LES EXTREMITES DES TRANSCRITS

Contrairement au noyau, dans les organelles aucune coiffe n'est ajoutée afin de préserver les ARN contre la dégradation. À l'inverse les ARN codant et non codant peuvent être polyadénylés, cependant les acteurs ne sont pas encore identifiés. Dans les organelles, la queue polyA est un signal qui oriente les ARN vers la dégradation par la polynucléotide phosphorylase (Walter et al., 2002; Holec et al., 2006; Schuster and Stern, 2009).

Dans le chloroplaste, les transcrits néosynthétisés, n'étant pas protégés en 5' et 3', subissent une dégradation par des endo- et exonucléases (Drager et al., 1999). Afin de protéger ces deux extrémités et d'augmenter la stabilité des transcrits, des protéines s'associent aux deux extrémités du transcrit (Pfalz et al., 2009).

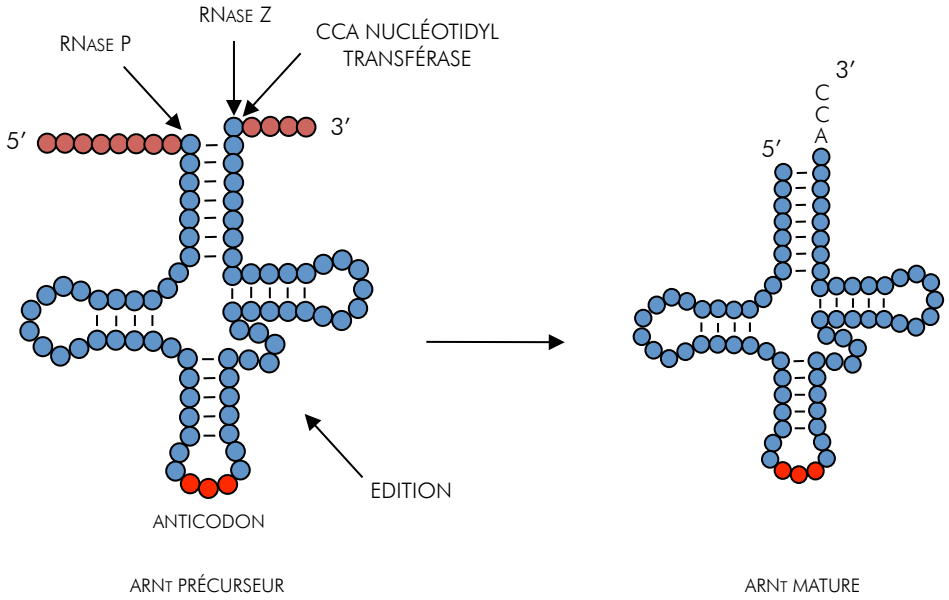


Figure 7 : LA MATURATION DE L'ARNT (Hopper and Phizicky, 2003)

Les transcrits primaires possèdent des extrémités 5′ et 3′ devant être maturées. La RNase P va cliver en 5′ la séquence nucléotidique tandis que la RNase Z la clive en 3′. Un groupement CCA sera ensuite ajouté grâce une activité nucléotidyl transférase. Dans certains cas des cytidines seront éditées en uridines.

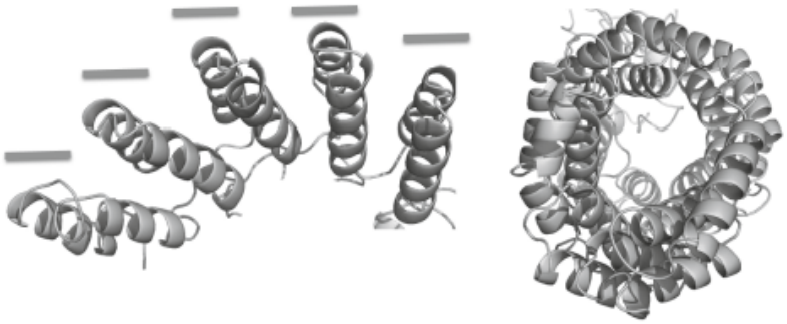


Figure 8 : PRÉDICTION DE STRUCTURE ET D'ORGANISATION DES MOTIFS PPR (Gutmann et al., 2012b)

Les 35 acides aminés constituant les motifs PPR présentent une séquence primaire très dégénérée. Toutefois, des prédictions de structure proposent qu'ils puissent s'organiser de manière à former deux hélices α. Les bars en gris indiquent chaque motif PPR.

Une succession de motifs PPR conduirait à une super-hélice formant un tunnel chargé positivement qui serait une plate-forme d'interaction pour les ARN.

1.2.2.2.4 LES COUPURES ENDONUCLEOLYTIQUES

Dans les chloroplastes deux endonucléases, RNase E et la RNase J, interviennent dans la maturation des extrémités 5' des ARN (Schein et al., 2008; Sharwood et al., 2011). La RNase J a une activité exo et endonuclease, et intervient dans la maturation des ARNm polycistroniques (Sharwood et al., 2011). Une dérégulation de l'expression de la RNase J conduit à une suraccumulation d'ARN antisens, qui a un effet délétère sur l'expression des gènes chloroplastiques.

Dans les organelles ainsi que dans le noyau, les précurseurs d'ARNt vont subir une activité de type RNase Z préalable à l'ajout du CCA (figure 7) (Kunzmann et al., 1998). Cette activité endonucléase est précédée par une autre en 5' de type RNase P. Il a été montré *in vivo* que des structures ressemblant à des ARNt (t-elements) dans les précurseurs d'ARNm mitochondriaux étaient clivées par une activité de type RNase Z ou RNase P (Forner et al., 2007). Au niveau nucléaire, les précurseurs dicistronique $ARNt^{Gly}$-snoR43 peuvent être clivés *in vitro* par la RNase Z. Cependant, cela nécessite la maturation au préalable de l'$ARNt^{Gly}$ en 5' (Kruszka et al., 2003).

1.3 LES PROTEINES A "PENTATRICOPEPTIDE REPEAT"

Bien que les mécanismes intervenant dans la maturation des transcrits issus du génome des organites de plantes soient largement connus dans leur principe, les acteurs eux, demeuraient dans l'ensemble inconnus. En 2000, le séquençage du génome d'*A. thaliana* a permis de caractériser une nouvelle classe de protéine : les protéines à pentatricopeptide repeat (PPR) (Aubourg et al., 2000; Small and Peeters, 2000), dont un nombre croissant d'étude montre leur implication dans les modifications post-transcrptionnelles.

Les protéines PPR comprennent une succession de 2 à 26 motifs PPR répétés en tandem, avec 12 motifs en moyenne. Le motif PPR classique est constitué de 35 acides aminés ayant une séquence primaire très dégénérée. Elle ne comprend que très peu de résidus conservés, qui sont surtout des tyrosines. Ce motif ressemble aux motifs TPR (tetratricopeptide repat) (Small and Peeters, 2000) qui sont impliqués dans les interactions protéine-protéine (Blatch and Lassle, 1999). La plupart des protéines PPR possède une séquence d'adressage vers les organites (Small and Peeters, 2000). Le motif PPR pourrait former deux hélices alpha anti-parallèle (figure 8) (Small and Peeters, 2000). La succession des motifs PPR formerait une super-hélice. Les tyrosines conservées pourraient être impliquées dans un empaquetage

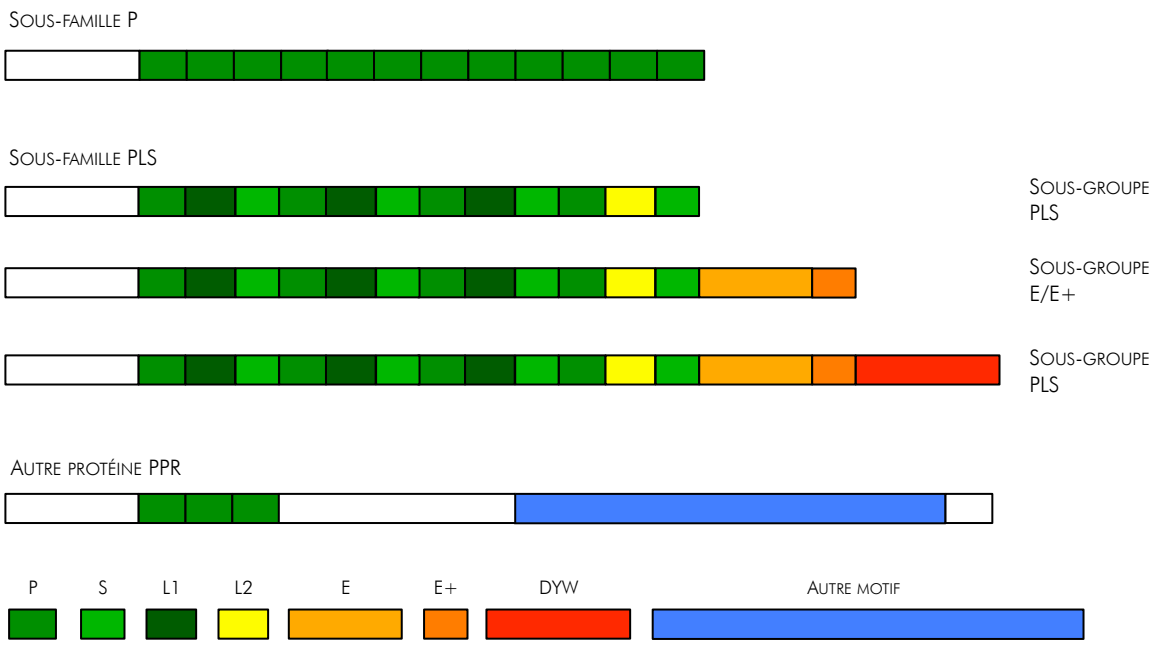


Figure 9 : LES DIFFÉRENTES CLASSES DE PROTÉINES PPR (Lurin et al., 2004).

Les protéines PPR sont formées de motifs PPR possédant une séquence de 35 acides aminés dégénérée. Ces motifs sont appelés PPR P. Il existe des variantes de tels motifs : les motifs PPR L (35 à 36 acides aminés) et les PPR S (31 acides aminés). Ces motifs permettent de classer les protéines PPR en deux sous-familles : P, possédant uniquement des motifs PPR P, et PLS possédant les trois motifs. Cette seconde sous-famille peut être divisée en sous-groupes. Ils dépendront de la présence en 3' de motifs supplémentaires : E, E+ et DYW. Ainsi on aura les sous-groupes PLS, ne possédant pas de motif additionnel, le sous-groupe E avec un motif E, le sous-groupe E+ avec les motifs E et E+, et enfin le sous-groupe DYW avec les motifs E, E+ et DYW. Le nombre des motifs PPR est variable et les motifs additionnels sont tronqués pour certaines protéines.

D'autres protéines, possédant un faible nombre de motifs PPR en N-terminal et un domaine catalytique en C-terminal, ne peuvent être classées dans les deux premières catégories.

inter hélice. L'intérieur de la super-hélice pourrait former un tunnel chargé positivement, qui pourrait être une plateforme d'interaction avec l'ARN (Small and Peeters, 2000).

1.3.1 LES DIFFÉRENTES CLASSES DE PROTÉINES PPR

Le motif PPR canonique, constitué de 35 acide aminés et nommé PPR P, possède deux variantes : PPR-like S (Short) et PPR-like L (Long) constitué respectivement de 31 et de 35 à 36 acides aminés (Lurin et al., 2004). Les protéines PPR peuvent être divisées en deux sous-familles chez *Arabidopsis*. La sous-famille P possédant uniquement le motif P répété en tandem et la sous-famille PLS ayant le long de ses protéines une répétition du triple motif P-L-S (figure 9).

La sous-famille PLS peut être subdivisée en sous-groupes grâce à la présence de trois motifs additionnels : E (91 acides aminés en moyenne), E+ (33) et DYW (106). Le motif DYW est caractérisé par la présence d'un motif $HxEx_nCxxC$ (Salone et al., 2007). Ainsi il existe quatre sous-groupes : PLS (sans motif supplémentaire), E (avec le motif E), E+ (avec les motifs E-E+) et DYW (avec les motifs E-E+-DYW).

Récemment, des protéines à motif PPR ont été identifiées mais ne pouvant être classées dans les groupes définis ci-dessus (Lurin et al., 2004). Elles possèdent en N-terminal des motifs PPR et en C-terminal un domaine catalytique, ne correspondant pas à celui des protéines du sous-groupe DYW. On retrouve par exemple dans cette sous-famille l'ARN polymérase mitochondriale de levure et humaine (Lipinski et al., 2011; Rackham and Filipovska, 2011) ainsi que les protéines PRORP (voir ci-dessous) (Gobert et al., 2010).

1.3.2 DISTRIBUTION DES PROTÉINES PPR DANS LE VIVANT

Chez les procaryotes, les protéines PPR sont absentes à l'exception de *Ralstonia solanacearum* (Salanoubat et al., 2002), un parasite de plante. Cette présence est certainement dû à un transfert génétique horizontal (Lurin et al., 2004). Au contraire, elles sont présentes dans tous les eucaryotes mais en nombre très variable (figure 10).

Elles sont présentes en petit nombre chez les animaux, les champignons et les protistes, mais présentes en très grand nombre chez les plantes terrestres. En effet elles sont par exemple au nombre de 7 chez l'humain (Rackham and Filipovska, 2011), environ 30 chez *Trypanosoma brucei* (Mingler et al., 2006; Pusnik et al., 2007) et 15 chez *Saccharomyces cerevisiae* (Lipinski et al., 2011), ainsi que 455 chez *Arabidopsis thaliana* et 470 chez

Figure 10 : RÉPARTITION DES PROTÉINES PPR DANS LE MONDE VIVANT (Lurin, et al., 2004).

Alors que les gènes codant pour les protéines PPR sont absents chez les procaryotes (il existe cependant quelques exceptions dû à un transfert génétique horizontal) et en faible quantité chez les animaux, les champignons et les algues (une dizaine), ils sont présents par centaines chez les plantes terrestres.

Organisme	Nombre de gènes PPR
Ralstonia solanacearum	1
Homo sapiens	7
Chlamydomonas reinhardtii	10
Saccharomyces cerevisae	15
Trypanosoma brucei	30
Arabidopsis thaliana	455
Oryza sativa	470

Tableau 1 : COMPARAISON DU NOMBRE DE GÈNES CODANT POUR DES PROTÉINES PPR DANS LE MONDE VIVANT

Oryza sativa (O'Toole et al., 2008) (tableau 1). On notera aussi que la sous-famille PLS est uniquement trouvée chez les plantes terrestres.

Les gènes codant pour les protéines PPR sont répartis de manière homogène sur les chromosomes chez *Arabidopsis*. Leur expression est faible : les protéines PPR P sont peu exprimées et les PPR PLS le sont encore moins (Lurin et al., 2004). 80% des ORFs possèdent un seul exon et 13% un intron. Elles possèdent au niveau N terminal, dans 70% des cas, un peptide signal dont on prédit un adressage soit vers les mitochondries (75%) soit vers les chloroplastes (25%).

1.3.3 L'IMPLICATION DES PROTÉINES PPR DANS LES MÉCANISMES POST-TRANSCRIPTIONNELS

Dans les plantes la diversité des protéines PPR se traduit également par une diversité de fonction. En effet ces protéines sont impliquées dans la transcription, la stabilité des ARN et leur dégradation, l'édition, l'épissage, les coupures endonucléolytiques et la traduction. On détaillera en particulier les protéines PPR impliquées dans les modifications post-transcriptionnelles, i.e. l'édition, l'épissage et les coupures endonucléolytiques.

1.3.3.1 L'EDITION DE L'ARN

Les protéines impliquées dans l'édition font partie de la sous-famille PLS. Cependant il existe une exception pour la protéine PPR596 (At1g80270), impliquée dans l'édition partielle de rps3 dans les mitochondries d'*A. thaliana* (Doniwa et al., 2010).

Ces protéines PPR possèdent un domaine N-terminal composé de motifs PPR qui seraient impliqués dans la reconnaissance à l'ARN. On peut citer pour exemple les protéines CRR4 (At2g45350) (Kotera et al., 2005; Okuda et al., 2006) et CRR21 (At5g55740) (Okuda et al., 2007), faisant partie du sous-groupe E+. Elles ont été identifiées par un crible génétique visant à identifier des mutants d'activité du complexe NA(P)DH déshydrogénase du chloroplaste. CRR4 et CRR21 sont impliquées dans la maturation de *ndhD* en éditant respectivement le premier et le second site d'édition de *ndhD*. Elles se lient avec 25 nucléotides en amont et 10 en aval du premier site d'édition de *ndhD*, permettant une reconnaissance spécifique du site d'édition.

Ces protéines PPR possèdent également un domaine C-terminal composé soit du motif E-E+ seul (sous-groupe E+) soit des motifs E-E+ et DYW (sous-groupe DYW). Les protéines PPR-DYW pourraient porter l'activité de déamination des C en U. En effet le DYW possède un motif $HxEx_nCxxC$, présentant une identité de séquence avec le site actif des cytidines désaminases (Salone et al., 2007). Au contraire les protéines PPR-E agiraient dans ce mécanisme post-transcriptionnel en recrutant le complexe d'édition, par l'intermédiaire de leur domaine C-terminal. En effet ce domaine, présentant une forte similitude, peut être interverti (Okuda et al., 2007) et les protéines chimériques qui en résultent gardent leur spécificité d'activité.

1.3.3.2 L'EPISSAGE

La mutation d'un gène codant pour une protéine PPR peut aussi affecter l'épissage de l'ARN. Le gène mitochondrial *nad1* code pour la sous-unité ND1 du complexe I de la chaîne respiratoire. Le transcrit est constitué de 5 exons et le premier intron est supprimé par un épissage en trans. Une mutation du gène *OTP43* (At1g74900) codant pour une protéine PPR mitochondriale, se traduit par une impossibilité de germination des graines en raison d'une absence d'épissage du premier intron du transcrit *nad1* (de Longevialle et al., 2007).

La reconnaissance protéine / intron peut se faire directement avec les introns intervenant dans l'épissage en trans. Cela a été montré pour la protéine PPR4 (Schmitz-Linneweber et al., 2006) qui est localisée dans le chloroplaste et constituée de 16 motifs PPR. La mutation du gène de PPR4 se traduit par l'absence d'épissage des transcrits du gène rps12. Afin de confirmer la cible, une étude par RIP-chip a été réalisée. Dans cette technique, les ARN qui coimmunoprécipitent avec la protéine cible sont identifiés par hybridation sur des microarrays (Keene et al., 2006). La protéine a été détectée comme se liant à deux séquences de *rps12*, correspondant au même intron des deux ARN épissés en trans. *rps12* code pour la protéine ribosomale S12 présente dans la sous-unité ribosomale 30S. Chez *E. coli*, S12 est essentielle à la viabilité cellulaire, car elle interagit avec la tige-boucle de l'ARNr 16S. L'absence de la protéine S12 pourrait donc provoquer une déficience en ARNr 16S dans les chloroplastes d'*A. thaliana*, celui-ci ne pouvant plus être stabilisé. Ainsi, cela expliquerait comment un défaut d'épissage des transcrits de *rps12* provoque un disfonctionnement dans la traduction chloroplastique (Schmitz-Linneweber et al., 2006).

La majorité des protéines PPR impliquées dans l'épissage font partie de la sous-famille P. Dans le cas des protéines du groupe PLS, les domaines E et DYW ne sont pas nécessaires à l'épissage. Cela a été montré pour la protéine PPR-DYW PpPPR_43 codée par le génome de *Physcomitrella patens* (Ichinose et al., 2012). L'épissage du troisième intron de *cox1* a pu être restauré après complémentation du mutant KO avec une construction codant pour une protéine tronquée au niveau des motifs E et DYW.

1.3.3.3 LES COUPURES ENDONUCLEOLYTIQUES

Chez les plantes la modification des extrémités 5' et 3' des ARNm mitochondriaux est nécessaire à l'obtention d'ARNm mature (Forner et al., 2007). Chez *Arabidopsis* trois protéines PPR-P (RPF1 (At1g12700), RPF2 (At1g62670) et RPF3 (At1g62930)) ont été identifiées comme impliquées de manière spécifique dans la maturation des extrémités 5' de certains ARNm mitochondriaux (Jonietz et al., 2010; Holzle et al., 2011; Jonietz et al., 2011). Les gènes de ces protéines sont regroupés sur un locus codant pour des restaurateurs de fertilité (Rf) dans la stérilité mâle cytoplasmique (CMS). Ces protéines interviennent en général dans la stabilisation des ARN, leur maturation et leur traduction, afin de synthétiser des protéines fonctionnelles nécessaires à l'obtention d'un pollen fertile.

Chez le riz, le locus Rf code pour la protéine RF1a qui va restaurer la fertilité mâle cytoplasmique en bloquant la synthèse de ORF79 (Wang et al., 2006). Le gène *orf79* est présent dans un opéron contenant une répétition du gène *atp6* (B-*atp6*). RF1a agit par des coupures endonucléolytiques sur le transcrit dicistronique B-*atp6*-*orf79* en trois régions différentes possédant chacune plusieurs sites de clivage.

Ainsi il apparaît que les protéines PPR sont bien des acteurs majeurs de l'expression génétique dans les organelles. Toutefois, dans certains cas, leur fonction affecte aussi l'expression génétique dans le noyau (Cooper et al., 2006; Hammani et al., 2011).

1.4 LES RIBOZYMES ET RIBONUCLEOPROTEINES (RNP)

L'hypothèse du monde à ARN repose sur la découverte de molécules d'ARN capables de s'autorépliquer et d'avoir une activité catalytique. La complexité du métabolisme cellulaire aurait poussé à la création des protéines et de l'ADN pour le stockage de l'information génétique, c'est le monde actuel dans lequel il existe des reliques de ce monde à ARN : les

ribozymes. Pour leur activité, ces ribozymes nécessitent souvent la présence d'ions divalent qui sont généralement des ions Mg^{2+} (Steitz and Steitz, 1993; Cochrane et al., 2007; Lee et al., 2008).

1.4.1 LES VIRUS ET VIROÏDES

Le ribozyme à tête de marteau a été caractérisé chez des virus et viroïdes qui affectent aussi bien les procaryotes que les eucaryotes (Perreault et al., 2011). Ces virus se répliquent selon la technique du cercle roulant synthétisant des intermédiaires multimériques qui seront clivés en monomères avant d'être circularisés (Symons, 1997). Afin d'obtenir des monomères de même longueur, le clivage autocatalytique de l'ARN monobrin est réalisé par une structure interne au virus le ribozyme à tête de marteau (Doudna, 1998). Ce ribozyme est constitué d'environ 50 nt avec un core central, très conservé, d'où partent trois petites hélices (Scott et al., 1995; Doudna, 1998).

D'autres ribozymes possèdent le même mode d'action que le ribozyme à tête de marteau, comme par exemple le ribozyme en épingle à cheveux, spécifique aux virus satellites à ARN de plante. Le site actif du ribozyme est constitué de deux hélices irrégulières, dont l'une possède le site de clivage (Rupert and Ferre-D'Amare, 2001). Contrairement aux autres ribozymes, son activité ne nécessite pas la présence d'un métal, laissant penser qu'il possède un mode d'action différent des autres ribozymes présents dans les virus (Hampel and Cowan, 1997; Nesbitt et al., 1997).

1.4.2 BACTÉRIES

Chez les bactéries Gram+, le transcrit de la glucosamine-6-phosophate synthase (*glmS*) possède dans la région 5' non-transcrite un ribozyme qui joue le rôle de riboswitch. La glucosamine-6-phosphate, qui est le produit de la GlmS, se lie au ribozyme (Cochrane et al., 2007). Cette reconnaissance conduit à un auto-clivage du ribozyme (Winkler et al., 2004), orientant le transcrit vers une dégradation intra-cellulaire (Collins et al., 2007).

1.4.3 EUCARYOTES

L'une des ribonucléoprotéines communes aux procaryotes et aux eucaryotes est le ribosome. Il intervient dans la traduction en catalysant la formation d'une protéine à partir d'un ARNm. Pour cela il utilise comme matrice l'ARNm, et il ajoute un acide aminé délivré par un ARNt à la protéine en cours de formation par une réaction de type peptidyl-transferase.

Le ribosome est constitué de deux sous-unités : une petite et une grande, dont la composition varie entre les procaryotes et les eucaryotes. Les mitochondries et les chloroplastes possèdent des ribosomes très proches des ribosomes bactériens. La grande sous-unité possède trois sites de fixation à l'ARNt : le site A (aminoacyl) qui est le site de reconnaissance codon et anti-codon, le site P (peptidyl) qui lie l'ARNt possédant le peptide en cours de formation et le site E (exit) qui est le site de sortie de l'ARNt.

Chez les eucaryotes, la petite sous-unité 40S est constituée d'un ARNr (18S) et de 33 protéines (Ben-Shem et al., 2011), tandis que chez les procaryotes la petite sous-unité 30S est constituée de 21 protéines et d'un ARNr (16S). Les sites A, P et E sont présents à l'interface de la sous-unité. C'est le cas aussi pour l'ARNm au niveau des sites A et P avant de traverser la sous-unité (Yusupova et al., 2001). Chez les bactéries, l'ARNr 16S intervient dans la reconnaissance codon-anti-codon grâce à une interaction avec l'anti-codon et le site A (Yoshizawa et al., 1999; Ogle et al., 2001). De plus un changement de conformation de cet ARN est nécessaire à une initiation de la transcription fidèle (Qin et al., 2012).

Chez les eucaryotes, la grande sous-unité 60S du ribosome est constituée de trois ARNr (5.8S, 5S et 28S / 25S) et de 46-47 protéines. Chez les procaryotes la sous-unité 50S possède 33 protéines et deux ARNr (5S et 23S). On retrouve les sites A, P et E ainsi que le site "peptidyl-transferase" qui est le lieu de catalyse de la liaison peptidique. La chaine néo-synthétisée progressera ensuite à travers un tunnel. Il est intéressant de noter que chez la bactérie, le site de transfert de l'acide aminé à la molécule en cours de synthèse est composé uniquement des ARNr (Rodnina et al., 2007).

Contrairement au virus la grande majorité des ribozymes présents chez les eucaryotes nécessite la présence de protéines pour avoir une activité *in vivo*, comme par exemple dans le cas des introns de groupe I et II (voir précédemment). Cependant il existe une exception chez *Neurospora*. Le plasmide Varkud, un type de rétroélément (Kuiper and Lambowitz, 1988), code pour l'ARN satellite Varkud. Cet ARN est capable de s'auto-cliver par une réaction semblable à celle des ribozymes de virus. Toutefois, la région

autocatalytique possède une séquence et une structure secondaire différentes des ribozymes nucléolytiques (Beattie et al., 1995). L'interaction entre le ribozyme et le substrat se fera essentiellement par la structure tertiaire, comme dans le cas de la RNase P (Lilley, 2004).

1.5 RNASE P

La ribonucléase P (RNase P) intervient dans la maturation des ARNt en libérant l'extrémité 5' des précurseurs d'ARNt, par une coupure endonucléolytique (Altman et al., 1993). Elle est impliquée également dans la maturation d'autres substrats, qui varient suivant les espèces (voir ci-dessous).

La RNase P est universellement conservée à l'exception de certaines *Archaea* où elle est absente (Randau et al., 2008). Classiquement, la RNase P est une RNP qui possède un ARN, portant l'activité catalytique (Guerrier-Takada et al., 1983), en association avec une ou plusieurs protéines (Lai et al., 2010a). Chez les procaryotes la RNase P est une RNP (Kouzuma et al., 2003; Kazantsev and Pace, 2006; Lai et al., 2012b). Chez les eucaryotes, l'activité RNase P est réalisée soit par une RNP, comme dans le noyau et les mitochondries de la levure (Dang and Martin, 1993; Kachouri et al., 2005) ou dans le noyau des animaux (Jarrous and Altman, 2001), soit par des enzymes uniquement protéiques (voir-ci-dessous).

1.5.1 FONCTION DE LA RNASE P

1.5.1.1 MECANISME CINETIQUE DANS LA RNP

Le clivage par la RNase P conduit à la formation d'un fragment d'ARN possédant en 3' un groupement phosphate, cependant il existe quelques exceptions chez *Halobacterium volcanii* ou chez les vertébrés supérieurs (Gupta, 1984; Lee et al., 1987) où l'ARN est transcrit sans séquence 5' leader et possède une extrémité 5' triphosphate.

Cette activité nécessite la présence de Mg^{2+} qui se fixera sur un boucle interne de la région core du ribozyme (Kufel and Kirsebom, 1998). Il peut être remplacé par le Mn^{2+} ou le Ca^{2+}, mais cela conduit à une diminution de l'activité et de la spécificité (Smith et al., 1992; Brannvall and Kirsebom, 1999).

1.5.1.2 DIVERSITE DE SUBSTRAT POUR LA RNASE P

La RNase P semble être essentielle à toutes les cellules, puisqu'elle est présente dans *Mycoplasma genitalium*, qui possède le génome bactérien le plus compact (Fraser et al., 1995). Sa fonction serait d'autant plus importante qu'elle intervient dans la maturation d'un grand nombre de substrats différents chez les bactéries.

Ainsi, en plus de la maturation des précurseurs d'ARNt, la RNase P bactérienne intervient dans la maturation de l'ARNr 4.5 S (Peck-Miller and Altman, 1991), d'ARNm polycistronique comme l'opéron histidine (Alifano et al., 1994; Li and Altman, 2003), de certains riboswitch dont celui du coenzyme B12 (Altman et al., 2005), de l'ARN antisens C4 des bactériophages P1 et P7 au niveau d'une structure en ARNt-like présentant une boucle D et T courte (Hartmann et al., 1995) et du 10Sa, un petit ARN stable présentant uniquement une boucle T (Komine et al., 1994). Cette diversité de substrat chez les bactéries a été confirmée *in vivo* en réprimant l'expression du gène *rpnB* codant pour le ribozyme de la RNase P de *Synechocystis* 6803 (Tous et al., 2001).

Le ribozyme de la RNase P d'*E. coli* est également capable de cliver des ARN eucaryotiques, comme le 2S ARNr de la *Drosophila melanogaster* (Hori et al., 2000b, a), le pre-tmRNA de la cyanelle maturé ou non en 3' (Gimple and Schon, 2001) et un riboswitch arginine putatif chez *Aspergillus nidulans* (Altman et al., 2005).

Chez les eucaryotes, les seules études ont été menées jusqu'à présent sur la RNase P de *S. cerevisiae*, qui a été caractérisée comme clivant le précurseur des ARNr de façon discrète au niveau de plusieurs sites différents (Chamberlain et al., 1996). Cependant une approche globale visant à déterminer le spectre de substrat de la RNase P de levure a permis d'isoler de nouveaux candidats comme par exemple certains snoARN à boite C/D (Coughlin et al., 2008). Après une purification par affinité de la RNase P, une analyse par genome-wide microarray a permis d'isoler des substrats potentiels. Ces cibles ont été analysées par northern blot en comparant l'ARN d'une levure sauvage avec celui de mutant thermosensible pour deux sous-unités protéiques de la RNase P (Coughlin et al., 2008).

Cette diversité de substrat est concordante avec l'analyse du substrat minimal à partir d'ARNt. Ainsi la RNase P bactérienne reconnait le bras accepteur et la tige-boucle T des ARNt (McClain et al., 1987), alors que la nucléaire humaine requiert, en plus, 7 nucléotides s'hybridant sur le bras accepteur (Yuan and Altman, 1995).

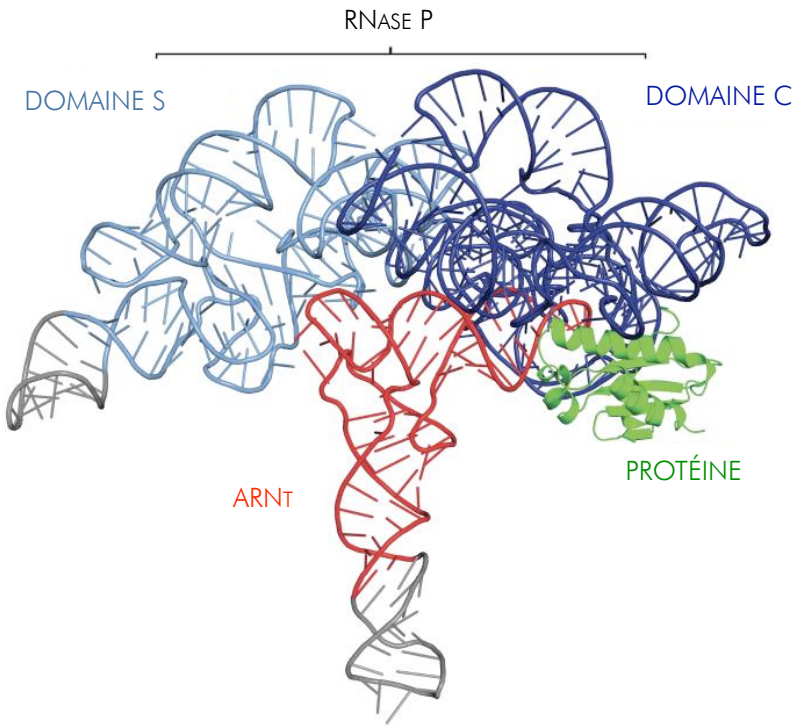


Figure 11 : STRUCTURE CRISTALLOGRAPHIQUE DE LA RNASE P RNP DE *T. MARITIMA* ASSOCIÉE À UN ARNT (Reiter et al., 2010)

La RNase P bactérienne est constituée d'un complexe ribonucléoprotéique dont l'ARN est un ribozyme. Ce dernier est constitué d'un domaine S (bleu clair) impliqué dans la reconnaissance du substrat et d'un domaine C (bleu foncé) impliqué dans l'activité catalytique de l'enzyme. L'ARN est associé à une petite sous-unité protéique (vert) assistant, *in vivo*, le clivage de l'ARNt précurseur (rouge) en se fixant sur l'extrémité 5' leader.

En gris sont représentés les nucléotides additionnels nécessaires à la cristallisation du complexe.

1.5.2 PRÉSENCE DANS LES DIFFÉRENTS PHYLA

1.5.2.1 PROCARYOTES

1.5.2.1.1 BACTERIES

La RNase P a longtemps été considérée comme une ribonucléoprotéine universelle qui serait l'un des derniers reliquats du monde à ARN et qui aurait déjà été présente dans le dernier ancêtre commun (Joyce, 1989). L'holoenzyme de la RNase P est constituée d'un ARN et d'une petite sous-unité protéique (figure 11).

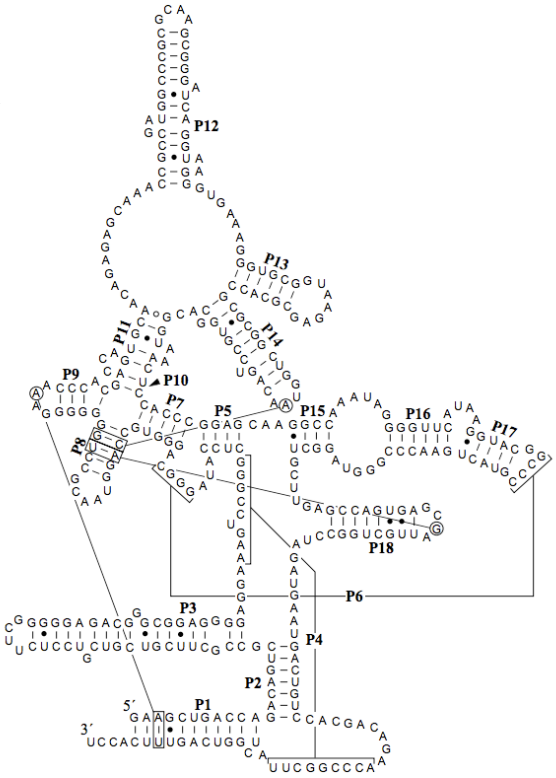
La composante ARN (RPR) est codée par le gène *rpnB* et a une longueur de 338 à 444 nucléotides (Brown and Pace, 1992). Elle peut être divisée en deux groupes en fonction de leur structure secondaire : le type A (Ancestral) ayant pour représentant *E. coli* et le type B (*Bacillus subtilus*) (figure 12). Une troisième forme, type C, retrouvée chez *Thermomicrobium roseum*, une bactérie verte non-sulfureuse, serait issue d'une évolution convergente avec le type B (Haas and Brown, 1998).

Les variations entre les différentes espèces résident au niveau de quatre domaines structuraux périphériques par rapport à une structure de base appelée core (Brown et al., 1991). Cette structure, composée de 263 nt (Waugh et al., 1989), possède deux régions : la région S (Specificity) porte le site de liaison reconnaissant la tige-boucle T (Hardt et al., 1995) et le CCA terminal qui est transcrit chez les bactéries (Guerrier-Takada et al., 1984; Harris et al., 2001) et la région C (catalytic) responsable de l'activité (Harris and Pace, 1995).

Les RPR bactériennes divergent aussi par leur biosynthèse. Ainsi chez *E. coli*, le ribozyme est synthétisé sous la forme d'un précurseur dont son extrémité 3' est maturée par un clivage endonucléolytique par la RNase E (Lundberg and Altman, 1995) avant de subir une digestion exonucléase par la RNase T (Li et al., 1998). Au contraire chez *B. subtilis*, les extrémités 5' et 3' du précurseur sont maturées de façon autolytique (Loria and Pan, 2000).

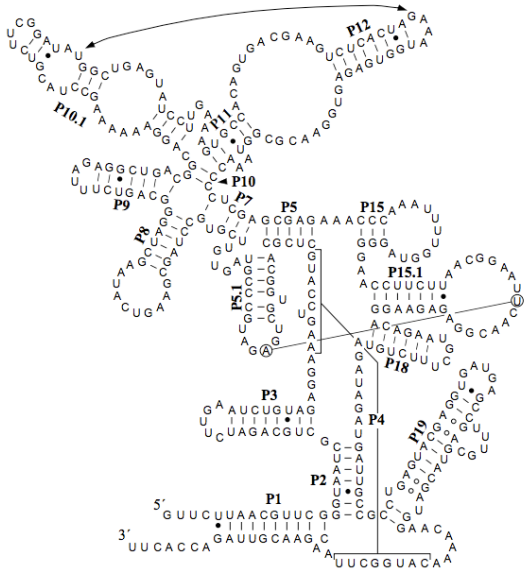
In vitro, l'ARN de la RNase P d'*E. coli* et de *B. subtilis* peut cliver, seul, les ARNt précurseurs en présence d'une forte concentration saline (Guerrier-Takada et al., 1983). Cependant, *in vivo* ou en présence d'une faible concentration en sel, le ribozyme ne peut fonctionner sans la présence d'une petite sous-unité protéique (RPP) (Schedl and Primakoff, 1973; Guerrier-Takada et al., 1983). Cette protéine d'environ 120 acides aminés est codée par le gène *rnpA*. Sa séquence est très peu conservée entre les différentes espèces, cependant certains acides aminés sont

A



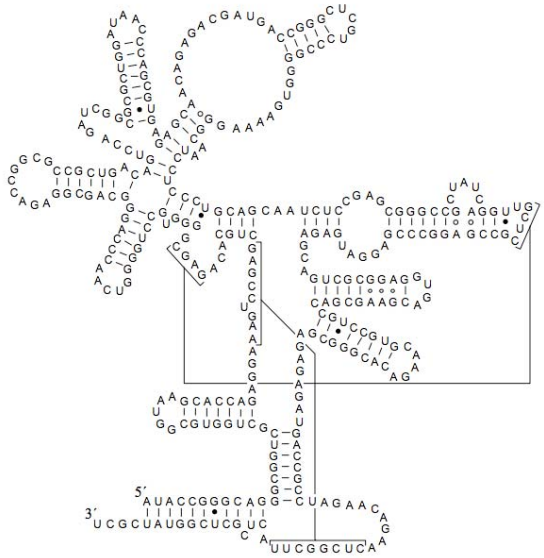
Type A : *Escherichia coli*

B



Type B : *Bacillus subtilis*

C



Type C : *Thermomicrobium roseum*

Figure 12 : REPRÉSENTATION DE LA STRUCTURE SECONDAIRE DES ARN DE LA RNASE P BACTÉRIENNE (Brown, 1999)

Chez les bactéries, les ARN de la RNase P peuvent être regroupés en trois types différents en fonction de leur structure secondaire : A, B et C. Le type A est représenté par l'ARN de la RNase P d'*E. coli*, le type B par l'ARN de la RNase P de *B. subtilis* et le type C par l'ARN de la RNase P de *T. roseum*.

conservés formant le motif $KX_{4-5}AX_2RNX_2(K/R)RX_2(R/K)$ (Kirsebom and Vioque, 1995; Gopalan et al., 1997). Malgré leur faible homologie de séquence, les sous-unités protéiques de *E. coli* et *B. subtilis* peuvent être échangées pour la maturation des précurseurs d'ARNt (Guerrier-Takada et al., 1983). Par contre la maturation d'autres substrats, comme l'ARNr 4.5S d'*E. coli*, ne pourra se faire que si le ribozyme de la RNase P d'*E. coli* est en présence de sa propre sous-unité protéique (Guerrier-Takada et al., 1983).

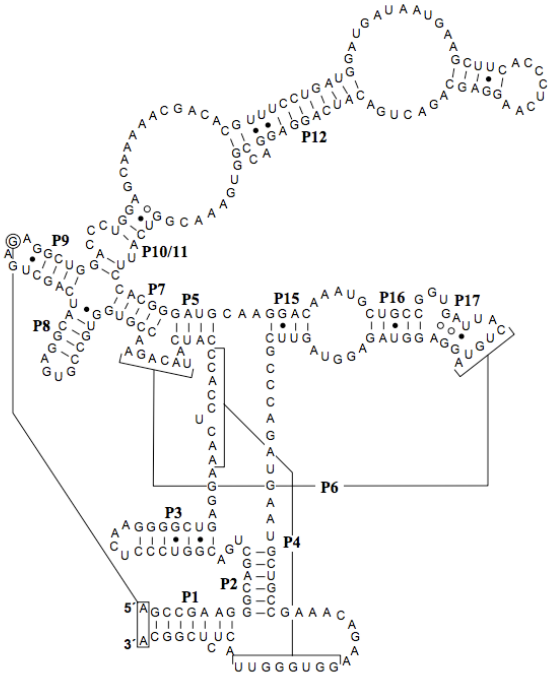
RPP est essentielle pour la stabilité et l'interaction avec le substrat. En effet, la présence de la sous-unité protéique augmente par un facteur de 10^4 l'activité du ribozyme de *B. subtilis* (Kurz et al., 1998) en augmentant l'interaction avec la séquence 5' leader (Crary et al., 1998) et en réduisant celle avec les ARNt matures en 5' (Tallsjo and Kirsebom, 1993). Une mutation de certains résidus hydrophobes ou basiques, va diminuer ou rendre inactif l'action du ribozyme et va diminuer son spectre d'action (Gopalan et al., 1997). Ceci peut s'expliquer par l'analyse de la structure tridimensionnelle de la RNase P ; la sous-unité protéique présente un sillon central qui interagit directement avec l'ARNt précurseur, le plaçant à proximité du site actif du ribozyme (Niranjanakumari et al., 1998). Cette structure est aussi présente au niveau de la protéine ribosomale S5 et du facteur d'élongation G de la translocase ribosomale (Stams et al., 1998). La sous-unité protéique de la RNase P et la protéine ribosomale S5 pourraient donc toutes dériver d'un ancêtre commun présent dans la machinerie traductionnelle primordiale.

Contrairement à l'expression de la RPR, la régulation de la RPP pourrait se faire au niveau de la transcription. En effet, RPP est cotranscrit avec la sous-unité ribosomale L34 (Panagiotidis et al., 1992). Cependant le rapport entre les deux protéines varie de 60 à 100 fois. Ceci pourrait être dû à la présence de deux transcrits, le long transcrit contenant l'ARN codant pour la RPP n'étant pas l'espèce prédominante. Cette différence pourrait s'expliquer par la présence d'un hairpin entre les deux ORF, bloquant la transcription du gène RPP (Panagiotidis et al., 1992).

1.5.2.1.2 *ARCHAEA*

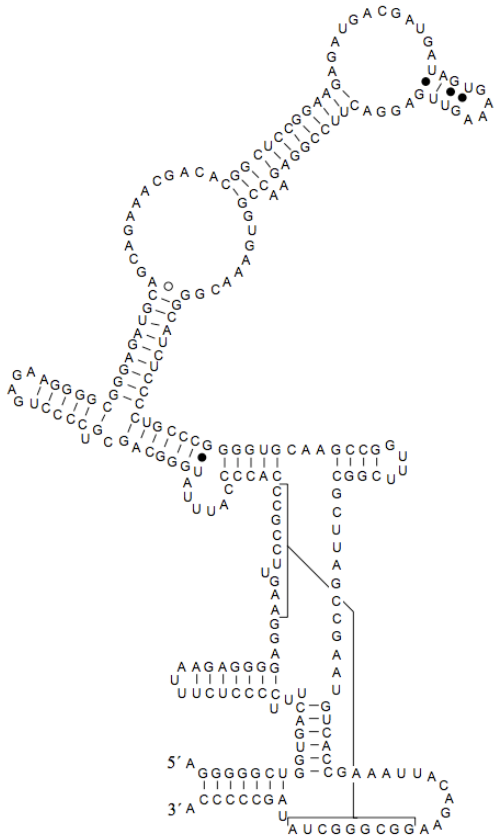
Les *Archaea* possèdent une forme intermédiaire de la RNase P, avec un ARN similaire à celui des bactéries, associé à quatre sous-unités protéiques semblables à celles des eucaryotes. La plupart des *Archaea* possèdent un ARN similaire à celui

A



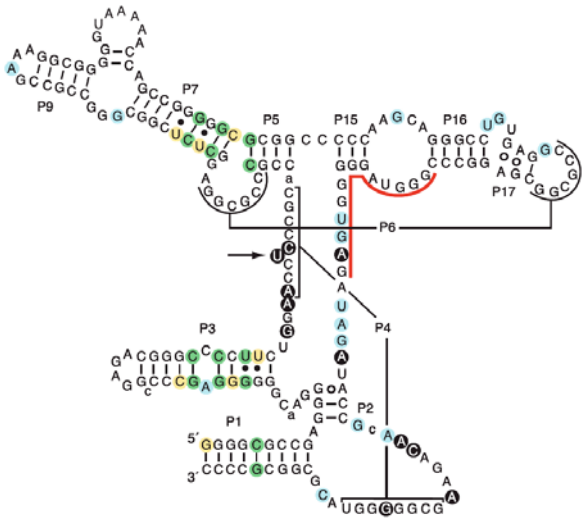
Type A : *Methanobacterium thermoautotrophicum*

B



Type M : *Methanococcus jannaschii*

C



Type T : *Pyrobaculum aerophilum*

Figure 13 : REPRÉSENTATION DE LA STRUCTURE SECONDAIRE DES ARN DE LA RNASE P PRÉSENTS CHEZ LES *ARCHAEA*

Chez les *Archaea*, les ARN de la RNase P peuvent être regroupés en trois types différents en fonction de leur structure secondaire : A, M et T. Le type A est représenté par l'ARN de la RNase P d'*M. thermoautotrophicum* (Brown, 1999), le type M par l'ARN de la RNase P de *M. jannaschii* (Brown, 1999) et le type T par l'ARN de la RNase P de *P. aerophilum* (Lai et al., 2010b).

du type A présent dans les bactéries et serait ainsi la forme ancestrale des RPR des *Archaea*. Cependant il existe des variations avec la forme A bactérienne. Ainsi la séquence L15, fortement conservée chez les bactéries et intervenant dans la reconnaissance du 3' NCCA du précurseur des ARNt et dans la liaison au métal, est très variable chez les *Archaea* (Kufel and Kirsebom, 1998; Harris et al., 2001).

Il existe deux formes supplémentaires du ribozyme, le type M (*Methanocaldococcus jannaschii*) et le type T (*Thermoproteaceae*), qui sont des formes réduites de l'ARN ancestral sur lesquelles des éléments essentiels sont manquants (figure 13). Ainsi l'ARN de type M ne possède pas la séquence L15 ainsi que celle intervenant dans la reconnaissance avec la tige-boucle T (Harris et al., 2001). Le type T quant à lui présente une forme minimale de l'ARN avec un domaine S presque totalement absent (Lai et al., 2012b).

L'absence de ces éléments essentiels pour le ribozyme expliquerait la présence d'un plus grand nombre de protéines par rapport à la RNase P de type bactérienne. Elles sont au nombre de quatre (Mth11p, Mth687p, Mth688p et Mth1618p) ou plus et possèdent une similarité de séquence avec des sous-unités protéiques de la RNase P nucléaire de *S. cerevisaea* (Pop4p, Pop5p, Rpp1p et Rpr2p) (Hall and Brown, 2002). Malgré tout, le ribozyme possède seul, une activité catalytique en présence d'une concentration très élevée en sel (Pannucci et al., 1999). Les quatre sous-unités protéiques sont présentes sous forme d'un complexe binaire (Kikovska et al., 2007). La protéine Pop4p est conservée chez l'ensemble des archaea et des eucaryotes. Son gène, comme chez les bactéries, est étroitement lié à celui des protéines ribosomales. En effet, il est localisé dans l'opéron codant pour les protéine ribosomales (Hartmann and Hartmann, 2003).

Il est intéressant de noter que chez *Nanoarchaeum equitans*, la RNase P est absente (Randau et al., 2008). Ceci est en désaccord avec la théorie soutenant que la RNase P soit conservée à travers l'ensemble du vivant. Le promoteur du gène codant pour les ARNt est situé juste en amont de la position +1 de l'ARNt. Ainsi un ARNt est synthétisé sans séquence leader et l'ARNt mature possède une terminaison 5' triphosphate.

1.5.2.2 EUCARYOTES

Chez les eucaryotes, la situation est beaucoup plus complexe avec deux compartiments supplémentaires contenant une activité RNase P (la mitochondrie et le chloroplaste) et la présence, suivant les espèces, de deux types de RNase P (RNP ou protéique).

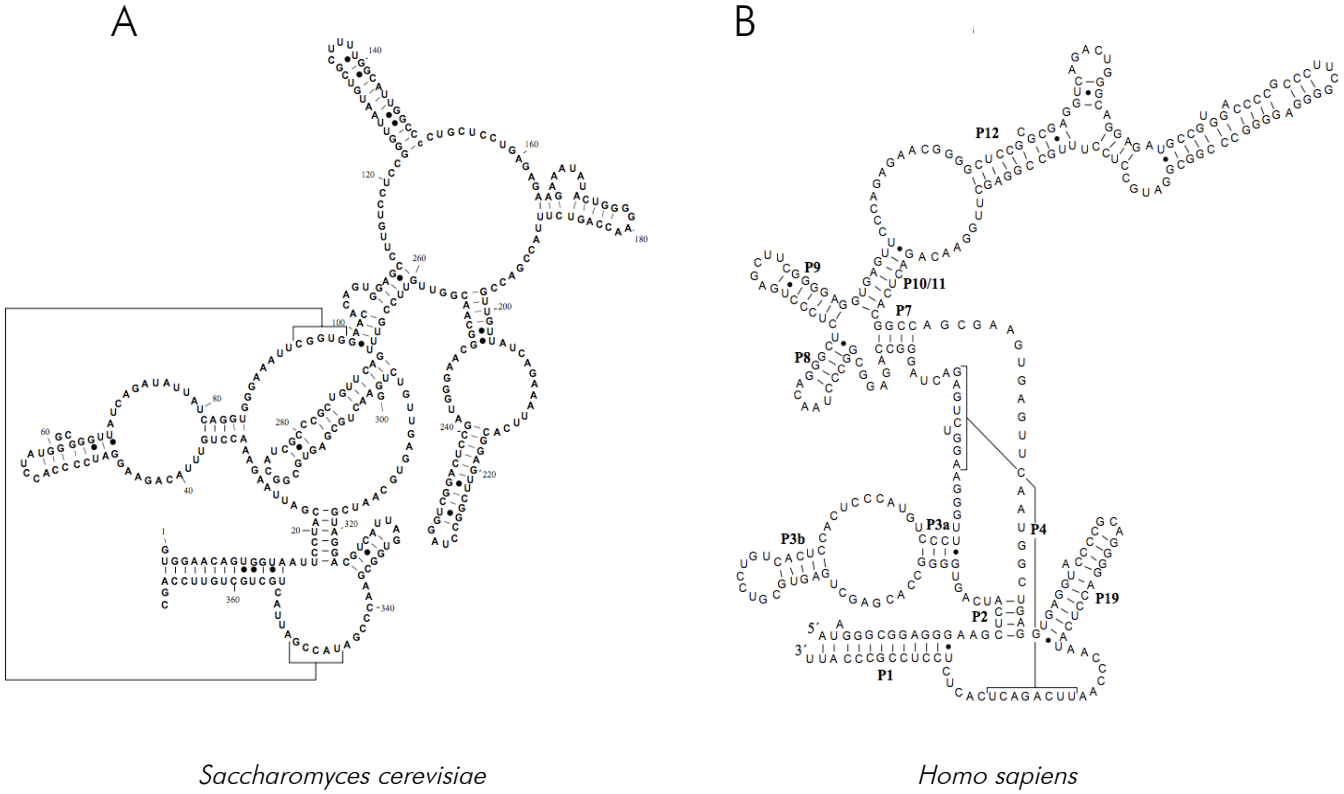
1.5.2.2.1 RNP NUCLEAIRE

Chez les levures et les animaux, la RNP nucléaire est de type bactérien. Elle est constituée d'un ARN ayant une homologie de séquence avec l'ARN de la RNase P bactérienne associé avec une dizaine de sous-unité protéique.

1.5.2.2.1.1 LEVURE

Chez la levure, la RNase P nucléaire possède neuf sous-unités protéiques (Pop1p, Rpp1p, Pop3p, Pop4p, Pop5p, Pop6p, Pop7p, Pop8p et Rpr2p) associées à l'ARN RPR1 (Lygerou et al., 1994; Chu et al., 1997; Dichtl and Tollervey, 1997; Stolc and Altman, 1997; Chamberlain et al., 1998; Stolc et al., 1998). Elle possède deux sites de reconnaissance au substrat, interagissant avec l'ARNt et le 3' trailer (Ziehler et al., 2000).
Dans le noyau des eucaryotes un autre complexe ribonucleoprotéique, la RNase MRP, a évolué en parallèle à la RNase P. Chez la levure, elle possède huit sous-unités communes, bien que le ribozyme de la RNase MRP diverge avec celui de la RNase P. Chez la RNase P une sous-unité protéique est spécifique (Rpr2p) (Chamberlain et al., 1998), tandis que chez la RNase MRP il y en a deux (Snm1p (Schmitt and Clayton, 1994) et Rmp1p (Salinas et al., 2005)).

RPR1 est synthétisée sous la forme d'un précurseur possédant un 5' leader de 84 nt et un 3' trailer de 34 nt (Lee et al., 1991), qui vont être supprimés au cours de l'assemblage de la RNase P donnant un ARN de 369 nt (figure 14). En effet des mutations dans *RPR1*, dans la région impliquée dans la liaison aux sous-unités protéiques, conduisent à un défaut de maturation du ribozyme (Ziehler et al., 2001), contrairement à des mutations dans des acides aminés conservés du domaine catalytique (Pagan-Ramos et al., 1996b, a). De plus une suppression de certaines sous-unités protéiques (à l'exception de Pop3p et Rpr2p), comme



Saccharomyces cerevisiae *Homo sapiens*

Figure 14 : REPRÉSENTATION DE LA STRUCTURE SECONDAIRE DES ARN DE LA RNASE P RNP PRÉSENTS DANS LE NOYAU DES EUCARYOTES

L'ARN de la RNase P des champignons est représenté par celui de *S. cerevisiae* et celui des animaux est représenté par celui de *H. sapiens* (Brown, 1999).

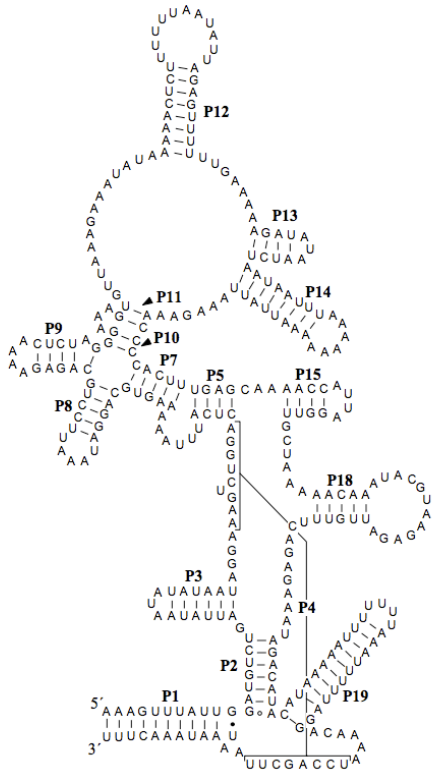
Pop1p ou Pop4p, conduit à une diminution de l'assemblage de la RNase P et à une diminution de la forme mature de *RPR1* (Chu et al., 1997; Srisawat et al., 2002).

In vitro, seulement sept des sous-unités protéiques sont nécessaires à l'activité d'une pre-RNase P (Srisawat et al., 2002). Pop1p et Pop4p sont les seules sous-unités qui interagissent de manière spécifique avec la sous-unité ARN (Houser-Scott et al., 2002). Pop3p, absente de ce core, possède une propriété de liaison à l'ARN et pourrait jouer un rôle dans la reconnaissance des différents substrats (Brusca et al., 2001). Rpr2p, également absente de ce core et présente uniquement dans la RNase P, interagit de manière forte avec elle-même et avec Pop4p (Houser-Scott et al., 2002). La même interaction est observée pour Snm1p, une protéine similaire à Rpr2p, présente uniquement dans la RNase MRP (Houser-Scott et al., 2002). Cette observation confirme une origine commune de ces deux ribonucléoprotéines.

1.5.2.2.1.2 ANIMAUX

En accord avec la sortie du monde ARN et avec les observations dans le noyau de la levure, le ribozyme de la RNase P (figure 14) aurait perdu au fur et à mesure de l'évolution son activité au profit de protéines accessoires, nécessaires à l'activité de la RNP. Cependant l'activité catalytique de l'ARN de la RNase P nucléaire humaine est conservée dans le temps (Kikovska et al., 2007). En effet *in vitro* l'ARN de la RNase P humaine (ARN H1) comme celui de *Giardia lamblia*, un eucaryote inférieur, peut cliver, seul, quatre précurseurs d'ARNt différents et une boucle en épingle à cheveux. Cependant l'affinité et le taux de clivage sont diminués. Comme dans le cas de la RNase P bactérienne, le ribozyme possède un core, mais il ne contient pas les éléments intervenant dans la catalyse et la structure tertiaire de l'ARN (Marquez et al., 2006). Cependant il est capable de se lier à l'ARN même en absence de protéines (Marquez et al., 2006).

Dans les cellules HeLa, 10 protéines, possédant des activités soit de liaison à l'ARN, soit d'ATPase hélicase, sont associées à l'ARN H1 (Jarrous and Reiner, 2007) : Rpp14, Rpp20, Rpp21, Rpp25, Rpp29, Rpp30, Rpp38, Rpp40, hPop1 et hPop5. Des homologues de la plupart des sous-unités protéiques sont



Reclinomonas americana

Saccharomyces cerevisiae

Figure 15 : REPRÉSENTATION DE LA STRUCTURE SECONDAIRE DES ARN DE LA RNASE P RNP MITOCHONDRIALE

L'ARN de la RNase P des mitochondries de *R. americana* est un exemple d'ARN très proche de celui de la RNase P de type bactérien (Brown, 1999). Au contraire l'ARN de la RNase P mitochondriale de *S. cerevisiae* a fortement divergé par rapport à celui des bactéries (Seif et al., 2003).

observables chez les autres métazoans (Hartmann and Hartmann, 2003). Cependant, contrairement à la levure, hPop1 ne se lie pas à l'ARN H1.

1.5.2.2.2 RNP MITOCHONDRIALE

Dans les mitochondries possédant une RNP, la situation est beaucoup plus complexe que dans le noyau avec deux types de RNase P ayant ou non une homologie avec la RNase P bactérienne.

Le génome mitochondrial de *Reclinomonas americana* code pour un RPR qui possède une structure proche de celle des bactéries (figure 15) (Lang et al., 1997) et la RNase P mitochondriale d'*Aspergillus nidulans* contient sept protéines associées à un ARN (Lee et al., 1996).

Chez la levure, la RNase P est constituée d'un ARN (figure 15), qui a considérablement divergé par rapport à celui des bactéries (Seif et al., 2003), associé à une sous-unité protéique, qui ne possède pas d'homologie avec celle de la RNase P bactérienne (Lai et al., 2010a). Le gène nucléaire *RPM2* code pour la sous-unité protéique de 105 kDa (Morales et al., 1992) qui intervient dans l'activité RNase P dans les mitochondries (Dang and Martin, 1993). Cette protéine serait aussi un activateur de la transcription de certaines protéines comme des composants de l'appareil d'import mitochondrial (Stribinskis et al., 2005) et serait impliquée dans le stockage et le renouvellement des ARN dans les P bodies (Stribinskis and Ramos, 2007). Rpm1r, la sous-unité ribonucléique constituant la RNase P, est codée par le génome mitochondrial (Hollingsworth and Martin, 1986). Le précurseur Rpm1r est cotranscrit avec l'ARNt$_f^{Met}$ et ARNt^{Pro} et sa maturation demande la présence de sa forme mature (Stribinskis et al., 1996) et de Rpm2p (Stribinskis et al., 2001).

1.5.2.2.3 RNASE P PROTEIQUE

1.5.2.2.3.1 ANIMAUX

Dans les mitochondries humaines la question de la nature de la RNase P est restée longtemps en suspens. Des études ont suggéré que la RNase P

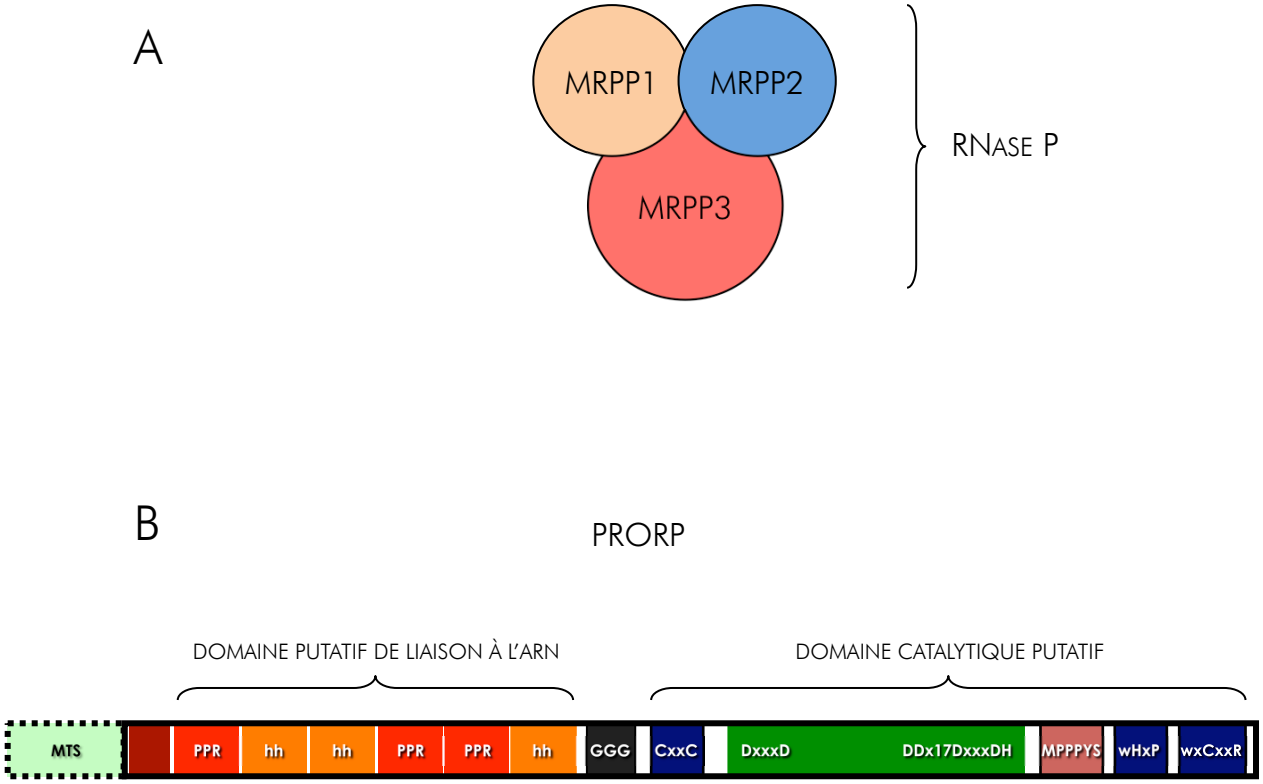


Figure 16 : LA RNASE P PROTÉIQUE

La RNase P mitochondriale humaine est constituée de trois sous-unités (A) : une ARNt méthyltransférase (MRPP1), une « short chain dehydrogenase/reductase » (MRPP2) et une protéine PPR (MRPP3) (Holzmann et al., 2008). MRPP3, contrairement à MRPP1 et MRPP2, est conservée dans de nombreux organismes. On retrouve ainsi trois orthologues chez *Arabidopsis thaliana*, nommés PRORP. (B) Ces protéines comprennent dans leur partie N terminale trois motifs PPR (rouge) ayant certainement un rôle dans la reconnaissance des ARN et 3 motifs (orange) ayant une structure secondaire en double hélice et qui pourraient être des motifs PPR non canoniques. En C terminal un motif métallonucléase (vert) pourrait être impliqué dans l'activité endonucléolytique de la RNase P et un motif CxxC dans la liaison à un cofacteur métallique. Entre les deux domaines, un motif riche en glycine (noir) pourrait conférer une flexibilité à la protéine.

mitochondriale humaine contenait un ARN identique au ribozyme de la RNase P nucléaire humaine (Doersen et al., 1985; Puranam and Attardi, 2001). La RNase P a finalement été caractérisée comme une entité strictement protéique (Holzmann et al., 2008). En effet elle est constituée de trois protéines, nommées Mitochondrial RNase P Proteins 1, 2 et 3 (MRPP1-3) (figure 16). Ces protéines, qui n'ont aucune homologie de séquence avec celles de la RNase P de type ribonculéoprotéique, sont nécessaires et suffisantes pour donner au complexe une activité de type RNase P (Holzmann et al., 2008). Comme les autres RNase P son activité nécessite du Mg^{2+} (Rossmanith and Karwan, 1998) et le clivage des précurseurs d'ARNt résulte en la formation d'un fragment 3' possédant un résidu phosphate (Holzmann et al., 2008).

MRPP1 est l'une des homologues de la protéine de levure Trm10p (Holzmann et al., 2008). Chez la levure c'est une ARNt m^1G methyltransferase responsable de la modification de l'ARNt à la position 9 (Jackman et al., 2003). Cette protéine pourrait donc jouer un rôle dans la liaison à l'ARN.

MRPP2 est une déshydrogénase 17β-hydroxystéroide de type 10 qui est essentielle pour le métabolisme de l'isoleucine et des acides gras à chaine ramifiée. Elle est aussi connue sous le nom d'ABAD pour la déshydrogénase amyloid-β liant l'alcool et est liée à la maladie d'Alzheimer (Holzmann and Rossmanith, 2009; Chen and Yan, 2010). Cependant l'activité déshydrogénase ne semble pas être nécessaire à la fonction de la RNase P. Toutefois MRPP2 possède un domaine Rossmann fold et pourrait donc être aussi impliquée dans la liaison à l'ARN (Holzmann et al., 2008).

MRPP3 possède en N-terminal des motifs PPR et pourrait donc être impliquée comme les deux autres dans la liaison à l'ARN. De plus, au niveau C-terminal, elle possède un domaine métallonucléase "NYN" (Anantharaman and Aravind, 2006) fortement conservé entre les différentes espèces et posséderait donc l'activité catalytique de l'enzyme (Holzmann et al., 2008).

1.5.2.2.3.2 PLANTES

Un homologue du ribozyme bactérien est présent dans le génome du cyanelle *Cyanophora paradoxa* (Shevelev et al., 1995) et du chloroplaste d'algues rouges comme *Cyanidioschyzon merolae* (Ohta et al., 2003) et d'algue verte comme *Nephroselmis olivacea* (Turmel et al., 1999). Chez *Daucus carrota* il semble

exister une activité RNase P dans chaque compartiment possédant une expression génétique (Franklin et al., 1995). Les études préliminaires menées sur le noyau et la mitochondrie semblent être en adéquation avec cette hypothèse, puisqu'elles définissent la RNase P comme une ribonucléoprotéine (Marchfelder, 1995; Schon, 1995).

Cependant les gènes, codant pour l'ARN et les sous-unités protéiques spécifiques de la RNase P de type bactérienne présente chez la levure et les animaux, sont absents des génomes végétaux (Hartmann and Hartmann, 2003). De plus, des études menées durant les années 1990 sur les chloroplastes d'épinard semblent montrer que la RNase P chloroplastique ne possède pas d'ARN. En effet, après séparation par un gradient de densité au CsCl, l'activité est présente dans la fraction de protéines chloroplastiques total (Wang et al., 1988). De plus après un traitement à la nucléase micrococcal, l'activité de la RNase P chloroplastique est toujours présente, alors qu'elle est inactivée pour l'enzyme d'*E. coli* (Wang et al., 1988). Enfin, la RNase P chloroplastique de l'épinard semble avoir un mécanisme d'action différent de celui de la RNase P (Thomas et al., 2000). Cependant, jusqu'à récemment, la RNase P n'avait jamais été identifiée au niveau moléculaire dans les plantes.

Un examen de banque de données suggère que MRPP3 soit la seule des trois protéines de la RNase P mitochondriale humaine à être conservée dans d'autres espèces. Trois orthologues de cette protéine sont trouvés chez *A. thaliana* (figure 16).

2 OBJECTIFS DE MA THESE

Au début de ma thèse, aucune protéine ou RNP ayant une activité RNase P n'avait été identifiée au niveau moléculaire chez les plantes.

En nous appuyant sur les travaux menés sur la RNase P mitochondriale humaine (Holzmann et al., 2008), nous avons pu identifier trois candidats potentiels pour l'activité RNase P chez *A. thaliana*, que nous avons nommé PRORP (PROteinaceous RNase P).

Tout d'abord, nous avons donc décidé de réaliser des études fonctionnelles de ces protéines *in vitro*, grâce à des tests d'activité avec des protéines PRORP recombinantes. En complément, nous avons voulu caractériser leur fonction *in vivo*. Pour cela nous nous sommes appuyés sur des lignées de mutants d'insertion ADN-T et des lignés de dérégulation par la technique du VIGS (Virus Induced Gene Silencing). Cependant la RNase MRP, une ribonucléoprotéine qui possède des sous-unités communes à RNase P chez les animaux et la levure, est présente dans le noyau. Donc on ne peut pas exclure l'hypothèse selon laquelle, en absence de RNase P ribonucléoprotéique, la RNase MRP ait acquis au cours de l'évolution l'activité RNase P chez les plantes. Cette hypothèse a été également étudiée avec des lignés de dérégulation par la technique du VIGS (Virus Induced Gene Silencing).

Par ces approches nous avons voulu voir si les plantes utilisent uniquement les protéines PRORP pour l'activité RNase P, à la fois dans le noyau et les organelles. Si c'est le cas, les plantes seraient les premiers exemples d'organismes dont l'activité RNase P serait entièrement réalisée par des protéines dans les trois compartiments possédant une expression génétique.

3 MATERIELS ET METHODES

3.1 MATERIELS

3.1.1 SOUCHES BACTÉRIENNES

Les souches BL21, DH5α et TOP10 d'*E. coli* ont été utilisées. La souche BL21 contient le gène codant pour l'ARN polymérase T7 sous le contrôle du promoteur lacUV5. Les souches DH5α et TOP10 présentent une mutation pour endA1, inactivant l'activité endonucléase intracellulaire, et pour recA éliminant les recombinaisons homologues.

Génotype des souches bactériennes :

BL21(DE3) : F– *omp*T *hsd*SB(rB–, mB–) *gal dcm* (DE3)

DH5α : F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1

TOP10 : F– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1*ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G

Les souches C58C1 et GV3101 d'Agrobacterium ont été utilisées durant cette thèse.

3.1.2 PLANTES

Les mutants d'insertion d'*A. thaliana* ont été mis en croissance à 23°C et un cycle diurne de 16 h. Ils sont issus des collections de GABI (noté GK et d'écotype Col-0), RIKEN (noté RATM, d'écotype No-0), SALK (noté SALK, d'écotype Col-0) et Versailles (noté EHQ d'écotype Ws-4).

At2g32230 : *prorp1-1* : GK_385G09 et *prorp1-2* : SALK_092567

At2g16650 : *prorp2* : EHQ_F435C08, RATM11-1141-1_G

At4g21900 : *prorp3 :* SALK_126397

At1g59760 : un mutant ADN-T *mtr4* a été fourni par Heike Lange (Lange et al., 2011)

3.1.3 PLASMIDES

pBAD27

Ce plasmide possède un ADN-T ayant une cassette de clonage juste en aval du promoteur fort 35S. Il a aussi une résistance au BASTA dans les plantes.

pER8

Ce vecteur possède un promoteur inductible à l'œstradiol qui est utilisé ici pour l'expression des miARN après transformation stable de l'ADN-T dans le génome de plante Col-0 sauvage d'*Arabidopsis thaliana*. Il possède également une résistance à l'hygromycine.

pET28

Ce plasmide est un vecteur d'expression des protéines chez *E. coli*. Un système de régulation par l'opéron lactose contrôle l'expression de la protéine d'intérêt fusionnée en N-terminal à une étiquette de 6 histidines. Dans ce système, le niveau bas d'expression peut être maintenu par un ajout de glucose dans le milieu de culture.

pGEM-Teasy

Ce plasmide possède à ces extrémités 3' un T sortant. Il est donc facile d'utilisation pour des clonages à partir de produit PCR issu d'une ADN polymérase qui ajoute un nucléotide à l'extrémité 3'. Ce vecteur est utilisé pour le clonage des produits de cRT-PCR.

pTRV1

Ce vecteur possède l'ADNc de l'ARN1 du virus TRV (Tobacco Rattle Virus), i.e. il code pour une ARN polymérase ARN dépendante, une hélicase, une protéine de mouvement et un faible suppresseur de silencing permettant au virus d'atteindre la partie méristématique. Il est utilisé dans le cadre de la technique VIGS (Virus Induced Gene Silencing) en association avec le plasmide pTRV2.

pTRV2

Ce plasmide renferme l'ADNc de l'ARN2 du TRV. Cet ARN code pour la protéine de manteau et pour deux protéines non-structurelles impliquées dans la transmission de plante à plante par le nématode. Dans le cas du vecteur pTRV2, les deux gènes codant pour les protéines non essentielles ont été remplacés par une cassette de clonage. Elle est

utilisée pour cloner une séquence de 400 à 600 nucléotides correspondant au gène cible à déréguler.

pUC19

Ce vecteur est utilisé ici pour le clonage et l'expression des transcrits précurseurs. Il est présent en grand nombre de copies et a été maintenu dans des souches de cellules TOP10.

3.1.4 AMORCES NUCLÉOTIDIQUES

Les amorces sont synthétisées par Sigma-Genosys et par Eurofins. Elles sont utilisées pour la mutagenèse des gènes codant pour les protéines PRORP, l'amplification par PCR des substrats à cloner, l'analyse par PCR et par séquençage des plasmides présents dans les clones obtenus et des mutants ADN-T de plante, la cRT-PCR, la qRT-PCR et la réalisation des sondes pour l'étude par northern blot.

MUTAGENESE DE DEUX ASPARTATES EN DEUX ALANINES (474-475 POUR PRORP1 ET 422-423 POUR PRORP2 ET 3)

prorp1_F	CCCTCCTGGCTCAAACGCTGCTTGGTATTGGCTTTACG
prorp1_R	CGTAAAGCCAATACCAAGCAGCGTTTGAGCCAGGAGGG
prorp2_F	CTCCACCAGGTTCCAATGCTGCTTGGTATTGGCTGTATGC
prorp2_R	GCATACAGCCAATACCAAGCAGCATTGGAACCTGGTGGAG
prorp3_F	CTCCACCAGGTTCCAATGCTGCTTGGTATTGGCTCTACGC
prorp3_R	GCGTAGAGCCAATACCAAGCAGCATTGGAACCTGGTGGAG

CLONAGE

TEST D'ACTIVITE IN VITRO

orf138_F	TAAAGGATCCGAATTGTAATACGACTCACTATAGCGTAGTTAATTACCTCGGACGTACCCATGG
orf138_R	TAAAGAATTCGAAAGGGTAGCCTTTTGTCACGAGCTG
nad6_F	TAAAGGATCCGAATTGTAATACGACTCACTATAGACGAAATGCTATTGATTTTAGGAGGACTATAATG
nad6_R	TAAAGAATTCCAACAGACAAGATCAGATCGTAGAATAGAGAAGG
ARNt$^{Asp(GUC)}$ n_F	TAAAGGATCCGAATTGTAATACGACTCACTATAGGGAAGCAGAGCAACAAAGTCG
ARNt$^{Asp(GUC)}$ n_R	TAAAGAATTCAAAAAAACGCCGTTGCCG
ARNt$^{Gln(CUG)}$ n_F	TAAAGGATCCGAATTGTAATACGACTCACTATAGGCAACAGCTATAGCAAAAAGG

ARNt$^{Gln(CUG)}$ n_R	TAAAGAATTCAAAAAAAGAGGTTCTACCGAG
prorp2_F	TAAACCATGGGAGCTGCTTCTGATCAACACCG
prorp2_R	TAAACTCGAGAGGAATCTTCCCATTACTCTTAGGT
prorp3_F	TAAACCATGGGAGCTGGTACTGATAACCGCCG
prorp3_R	TAAACTCGAGTGAACTCTGCCTTGTAATGCAC

VIGS

POP1 _F	TAATCTAGATCTGAGCGTTTGAACAAGGA
POP1_R	TAACTCGAGGCCTGAGAAACTGGTGGTTC
POP4_F	TAAAGGATCCGGCATCACAGAAGGATTCAA
POP4_R	TAAACTCGAGCTCTTGAAATAAACTTGTCCCCAT
PRORP1_F	TAAGGATCCAAATCGACTTTTGATATGATTGAGG
PRORP1_R	TAACTCGAGATCATCGTTTGAGCCAGGAG
PRORP2_F	TAATCTAGATGAATCTGTGATGAAGGAACTTTA
PRORP2_R	TAACTCGAGTGTCTGCTAATACACATCCAAGTT
PDS_F	TAAGAATTCCTTCCGTAGTGCTCCTCGTC
PDS_R	TAATCTAGACCATGTCAGCATCTCGTTGT

miARN

A6_miR-I	GATACATTATAATGGTACTGGGTTCTCTCTTTTGTATTCC
A6_miR-II	GAACCCAGTACCATTATAATGTATCAAAGAGAATCAATGA
A6_miR*III	GAACACAGTACCATTTTAATGTTTCACAGGTCGTGATATG
A6_miR*IV	GAAACATTAAAATGGTACTGTGTTCTACATATATATTCCT
A29_miR-I	GATTTTACACGACTCTTGTGGAGTCTCTCTTTTGTATTCC
A29_miR-II	GACTCCACAAGAGTCGTGTAAAATCAAAGAGAATCAATGA
A29_miR*III	GACTACACAAGAGTCCTGTAAATTCACAGGTCGTGATATG
A29_miR*IV	GAATTTACAGGACTCTTGTGTAGTCTACATATATATTCCT
B44_miR-I	GATACTAACGCAGCCAACACATATCTCTCTTTTGTATTCC
B44_miR-II	GATATGTGTTGGCTGCGTTAGTATCAAAGAGAATCAATGA
B44_miR*III	GATACGTGTTGGCTGGGTTAGTTTCACAGGTCGTGATATG
B44_miR*IV	GAAACTAACCCAGCCAACACGTATCTACATATATATTCCT
B45_miR-I	GATACTAACGCAGCCAACACAATTCTCTCTTTTGTATTCC
B45_miR-II	GAATTGTGTTGGCTGCGTTAGTATCAAAGAGAATCAATGA
B45_miR*III	GAATCGTGTTGGCTGGGTTAGTTTCACAGGTCGTGATATG

B45_miR*IV	GAAACTAACCCAGCCAACACGATTCTACATATATATTCCT
Oligo A	CTGCAAGGCGATTAAGTTGGGTAAC
Oligo B	GCGGATAACAATTTCACACAGGAAACAG

CONTROLE PAR PCR

PLASMIDE

promoteurT7_F	TAATACGACTCACTATAGGG
promoteurSP6_F	ATTTAGGTGACACTATAG
M13_F	GTAAAACGACGGCCAGT
M13_R	CAGGAAACAGCTATGAC
EYFP-F_R	GCTGAACTTGTGGCCGTTTA
pTRV2.1_F	TGTTTTTATGTTCAGGCGG
pTRV2.1_R	GACTTACCGATCAATCAAGATC
pTRV2.2_F	TTACTCAAGGAAGCACGATG
pTRV2.2_R	GAACCTAAAACTTCAGACACGG
pER8_F	CATTTGGAGAGGACACGCTGAAG
pER8_R	ACGTTGTCGAAACCGATGATACGG
pET28_R	GCTAGTTATTGCTCAGCGG

ZYGOTIE

LB4 (Versailles)	CGTGTGCCAGGTGCCCACGGAATAGT
Lba1 (SALK)	TGGTTCACGTAGTGGGCCATCG
LB1a (GABI)	GGGCTACACTGAATTGGTAGCTC
RB4 (Versailles)	TCACGGGTTGGGGTTTCTACAGGAC
RBa (SALK)	TAAAACGGCTTGTCCCGCGTCATC
RB2 (SALK)	AATAGCCTCTCCACCCAAGC
RB1a (GABI)	ACAGTGGTCCCAAAGATGGA
prorp1-1_F	CTTTCCACCTTGGGAAGAAAC
prorp1-1_R	GCCCTTTTTGGTTTCTGTAGG
prorp1-2_F	GGTCTATAAGACGTTGCAGCG
prorp1-2_R	ACCTGCAAACATTGAATCCAG
prorp2 _F	TTCATAGTCATGGCATTAGCTT
prorp2 _R	TTTTCGCATGAGACCGGGAAA

prorp3_F	ACACAGCCTAAGAATCGGGTC
prorp3_R	AGCCGGAGCATAAGTTCTCA

INTERNE AUX GENES prorp

306.P2/3int-F_seq	GAGGCTGAAAAGGGTTAT
306.P2/3int-R_seq	TATAGCTTCGTAATCTCCATG
306.P1int-F_seq	GCTTAGTAGGGGTTTTGATA
306.P1int-R_seq	TTAGCCTTGACTTCTCTTTC
P2/3front_RV	GACTTATCCCCGAACTAACCATAC
P2dif F	GGCATCTGTTACTTCAGTGGCTA
P2dif R	TTTTCGCATGAGACCGGGAAA
P3dif F	ATCATCTGTCACTGCAGTTGCAC
P3dif R	TCTTGACTCGTGATTGGAACG

cRT-PCR

ARNt$^{Asp(GUC)}$ n_F	GTGACCCGGGTTCGATCC
ARNt$^{Asp(GUC)}$ n_R	GGGAATACTTACCACTATACTACAACG
ARNt$^{Gln(CUG)}$ n_F	CAGTAACCCGAGTTCAATTCTC
ARNt$^{Gln(CUG)}$ n_R	TCCAATGTCCTAACCGCTAGA

q-RTPCR

PRORP1_F	CCAAGGTGGAAAGAGAAGCA
PRORP1_R	ATGCCAGGTTCCATCTTCAG
PRORP2_F	GCACCAGCTTATTGAGTTCATGTA
PRORP2_R	GAAAAGTTCAGTCCCTTATTATAGAACC
POP1_F	TCCAGAAGACCTTACATCCTGC
POP1_R	CAGCAACGTTTTTAACCTCAGC
POP4_F	CCGTCCATCTTTGATTCGTC
POP4_R	TTTGCAACAGCAAGCCTTC
EXP1_F	GAGCTGAAGTGGCTTCCATGAC
EXP1_R	GGTCCGACATACCCATGATCC
SAND_F	AACTCTATGCAGCATTTGATCCACT

SAND_R	TGATTGCATATCTTTATCGCCATC
TIP41_F	GTGAAAACTGTTGGAGAGAAGCAA
TIP41_R	TCAACTGGATACCCTTTCGCA
EXP2_F	TTGAAAATTGGAGTACCGTACCAA
EXP2_R	TCCCTCGTATACATCTGGCCA

NORTHERN BLOT

VIGS

ARNt$^{Phe(GAA)}$ p	CCTCTGCTCTACCAACTGAGCTATC
ARNt$^{Ile(GAU)}$ p	GCGCGCTCTACCACTGAGC
ARNt$^{Val(GAC)}$ p	TGACACTCTACCGCTGAGTTATATCC
ARNt$^{Ser(GCU)}$ m	CAATGCCTTAAGCCACTCAGC
ARNt$^{Lys(UUU)}$ m	CCAATGCTCTACCAACTGAGC
ARNt$^{Gln(CUG)}$ n	GGTCCCACCGAGACTTGAAC
ARNt$^{Asp(GUC)}$ n	GGGAATACTTACCACTATACTACAACG
ARNt$^{Ser(AGA)}$ n	GTCATGCCCGATAACCACTC
ARNt$^{Lys(CUU)}$ n	TTGCGCTCTACCAACTGAGCT
ARNt$^{Asn(GUU)}$ n	CTAGAGAAAGGGGTCGAACC
ARNt$^{His(GUG)}$ n	CAGGTCCCCACGGCCACAAC

miARN

306A6north	ACCCAGTACCATTATAATGTA
306A29north	CTCCACAAGAGTCGTGTAAAA
306B44north	TATGTGTTGGCTGCGTTAGTA
306B45north	ATTGTGTTGGCTGCGTTAGTA

3.1.5 PLASMIDES DISPONIBLES

Le clone d'expression des protéines (pBad-PRORP1) utilisé comme substrat pour la mutagenèse ainsi que les clones utilisés pour la localisation cellulaire (pART7-PRORP1-EYFP, pART7-PRORP2-EYFP et pART7-PRORP3-EYFP) ont été préalablement obtenus au laboratoire par Anthony Gobert.

Un plasmide renfermant la séquence du t-element de l'*orf138* de *B. napus* a été fourni par Françoise Budar (INRA de Versailles). Il correspond à la séquence comprise entre les nucléotides 910 et 1101 en aval de l'ATG de l'*orf138*.

3.2 METHODES

3.2.1 ADN

3.2.1.1 AMPLIFICATION PAR PCR

La technique de réaction de polymérisation en chaîne (PCR) repose sur l'amplification exponentielle d'une séquence d'ADN présente entre deux amorces spécifiques au brin matriciel. Pour cela, trois phases répétées sont nécessaires : la dénaturation de l'ADN, l'hybridation des amorces puis l'élongation par une ADN polymérase ADN dépendante thermorésistante.

Pour une telle réaction, le volume réactionnel sera composé de **tampon de PCR 1x**, 200 µM de chaque dNTP, 1 U d'ADN polymérase ADN dépendante, 5 pmol de chaque amorce, 2.5 mM $MgCl_2$ et d'une quantité appropriée d'ADN matrice ou de bactéries dans le cas de PCR sur colonie.

Il existe plusieurs types d'ADN polymérase ADN dépendante utilisées suivant l'expérience proposée. Pour le clonage et la mutagenèse, une enzyme à haut degré de fidélité sera privilégiée comme la *Phusion* (*Pyrococcus*-like : clontech) ou la *GoTaq* (Promega). Elle possède une activité exonucléase 3'-5' qui implique une activité de relecture. Au contraire pour le génotypage et l'analyse de colonie bactérienne une *Taq* ADN polymérase thermorésistance (de *Thermophilus aquaticus*) sera utilisée.

La PCR débutera par une étape de dénaturation de 3 à 5 min à 95°C afin de casser la paroi bactérienne et de dénaturer la matrice. S'en suivra un cycle répété 35 fois et comprenant les trois phases décrites précédemment : 30 sec à 95°C, 30 sec à la température d'hybridation des amorces (5°C en dessous de la température de fusion des amorces), puis 1 min/kb d'ADN matriciel à amplifier à 72°C, la température d'activité

de l'ADN polymérase thermorésistante. Une phase d'élongation finale de 5 min à 72°C fera suite aux 35 cycles, afin de terminer l'élongation de tous les brins.

Tampon de PCR 10x : 20 mM tris-HCl pH 8.3 ; 50 mM KCl ; 2 mM $MgCl_2$

3.2.1.2 ELECTROPHORESE SUR GEL D'AGAROSE

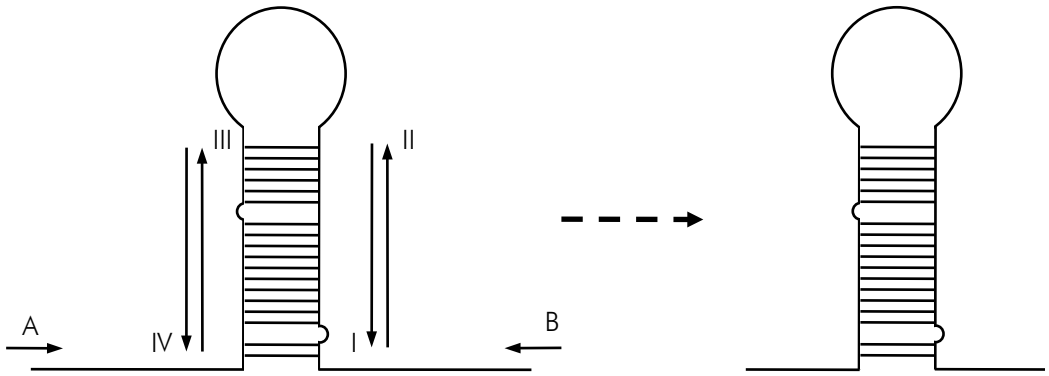
Cette technique permet de séparer des molécules d'ADN selon leur taille. Une solution d'ADN associée à du **tampon de charge 1x**, sera chargée sur un **gel d'agarose** (0.8 à 2% (p/v) dans du **tampon TAE 0.5x**) auquel 0.5 µg/ml de bromure d'éthidium a été incorporé.
L'électrophorèse est effectuée dans du **tampon TAE 0.5x** sous un voltage de 100 V. Le bromure d'éthidium, s'intercalant entre les bases, permet d'observer aux ultraviolets les molécules d'ADN.

Gel d'agarose à x % : x % (p/v) d'agarose ; TAE 0.5x
Tampon de charge 10x : 50 % (v/v) glycérol ; 0.1 M EDTA ; 0.1 % (p/v) SDS ; 0.05 % (p/v) bleu de bromophénol ; 0.05 % (p/v) xylène cyanol
Tampon TAE 1x : 40 mM Tris-acétate pH 8.0 ; 1 mM EDTA

3.2.1.3 DIGESTION ET LIGATION DE FRAGMENTS D'ADN

Digestion
Les digestions des fragments d'ADN par des enzymes de restriction sont réalisées selon les conditions du fournisseur.

Ligation
Une pmol d'insert et 0.1 pmol de vecteur sont mises en présence dans un volume réactionnel de 10 µl présentant l'ADN ligase du phage T4 (1 U/µg de vecteur), 1 mM ATP et 1x tampon ligase (fourni par Fermentas). La réaction a lieu pendant 1 h à température ambiante ou 16 h à 15°C.



	AMORCE SENS	AMORCE ANTISENS	MATRICE
a	A	IV	pRS300
b	III	II	pRS300
c	I	B	pRS300
d	A	B	a + b + c

Figure 17 : PRINCIPE DU CLONAGE DES MIARN ARTIFICIELS (http://wmd2.weigelworld.org/cgi-bin/mirnatools.pl)

Le miARN artificiel contenant la séquence d'intérêt est obtenu par quatre PCR successives. Les PCR a, b et c consistent en une mutagenèse du miR319a présent dans le plasmide pRS300. Les amorces I, II, III et IV sont choisies avec le programme wmd2 et les amorces A et B ont comme séquence celle du plasmide.

3.2.1.4 CONSTRUCTION AMIARN

La séquence du amiARN (miARN artificiel) a été choisie grâce au logiciel WMD2 (http://wmd2.weigelworld.org/cgi-bin/mirnatools.pl) tout en tenant compte des règles suivantes :

position 1 : U

position 2 à 12 : 1 mésappariement au maximum avec ARNm

position 10 et 11 : pas de mésappariement avec l'ARNm

position 10 : A

extrémité 3' : 1 à 3 mésappariements avec l'ARNm

Dans un premier temps, trois mutagenèses par PCR sont réalisées à partir du miR319a présent sur le plasmide pRS300 (figure 17). On amplifiera séparément le bras 5' (A-IV), la boucle (III-II) et le bras 3' (I-B) afin de remplacer le miR319a par le pre-amiARN.

Dans un second temps, les trois fragments PCR sont fusionnés en un fragment PCR, grâce à une PCR ayant pour matrice les trois fragments et pour amorce les oligonucléotides A et B.

Le fragment obtenu est dans un premier temps cloné dans le vecteur pGEM-Teasy, permettant le séquençage des constructions et facilitant le second clonage dans le vecteur pER8.

3.2.1.5 TRANSFORMATION DE BACTERIES PAR ELECTROPORATION

40 µl de bactéries électrocompétentes sont mises en présence de 50 à 200 ng d'ADN plasmidique représentant un volume de 1 µl. Ce mélange transféré dans une cuve d'électroporation va subir une décharge électrique (200 Ω ; 2,5 kV ; 25 µF), permettant à l'ADN de traverser la paroi bactérienne transitoirement déstabilisée.

Après électroporation, 400 µl de **milieu LB** sont ajoutés au mélange qui est incubé 1 h à 37°C sous agitation. Après cette période nécessaire à la régénération de la paroi bactérienne et à l'expression du gène de sélection, 200 µl de la solution sont étalés sur une boîte de Pétri contenant du **milieu LB** auquel est associé l'antibiotique approprié.

Milieu LB : 1 % (p/v) de bacto-tryptone, 0.5 % (p/v) d'extrait de levure, 0.5 % (p/v) de NaCl, pH ajusté à 7.2 (ajouter 1 % (p/v) d'agar pour obtenir du milieu LB solide).

3.2.1.6 TRANSFORMATION DES BACTERIES PAR CHOC THERMIQUE

50 µl de bactéries thermocompétentes sont mises en présence de 15 ng à 1 µg de plasmide ou de ligation puis incubées pendant 30 min dans de la glace. Ce mélange va ensuite subir un choc thermique durant 30 sec à 42°C permettant l'entrée du plasmide dans les bactéries, avant d'être replacé sur la glace durant 2 min.
400 µl de **milieu LB** sont ajoutés au mélange qui sera incubé 1 h, délai nécessaire pour la régénération des membranes et à la première division cellulaire. 100 à 200 µl seront ensuite étalés sur boite de Pétri contenant du milieu LB agar en présence de l'antibiotique approprié.

Milieu LB : 1 % (p/v) de bacto-tryptone, 0.5 % (p/v) d'extrait de levure, 0.5 % (p/v) de NaCl, pH ajusté à 7.2 (ajouter 1 % (p/v) d'agar pour obtenir du milieu LB solide).

3.2.1.7 PREPARATION D'ADN PLASMIDIQUE PAR LYSE ALCALINE

Elle a été effectuée selon les méthodes standards (Sambrook and Russell, 2001).

3.2.1.8 SEQUENÇAGE D'ADN

Les réactions de séquençage sont réalisées par M. Alioua (service de séquençage de l'IBMP). Cette technique utilise une PCR asymétrique basée sur la méthode de Sanger (Sanger et al., 1977). Les acides nucléiques sont ensuite séparés par électrophorèse capillaire à l'aide d'un appareil "Applied Biosystems 3100" (Perkin Elmer).

3.2.1.9 EXTRACTION DE L'ADN TOTAL DE FEUILLE D'*ARABIDOPSIS*

<u>Méthode 1 suivie par une PCR avec des amorces de haute affinité</u>
Cette technique commence par la lyse des cellules par l'intermédiaire d'un broyage du matériel végétal. Pour cela 100 mg de feuilles sont mis en présence de billes de verre et de 400 µl de **tampon d'extraction** sans SDS afin de lyser les tissus avec l'agitateur-broyeur Precellys (3 cycles de broyage de 30 s à 25000 g). 10 mM de SDS sont ensuite ajoutés à l'extrait obtenu. Ce détergent permettra de dénaturer les protéines et d'éviter que les membranes ne s'agrègent avec l'ADN. Après 30 min d'incubation à température

ambiante, les membranes sont culottées par une centrifugation de 1 min à 16000 g. Le surnageant est mis en présence de 300 µl d'isopropanol afin de précipiter l'ADN. Après un temps d'incubation de 2 min à température ambiante, les tubes sont centrifugés à 16000 g pendant 10 min. Le surnageant éliminé puis le culot séché, le culot d'ADN est repris dans 100 µl d'H_2O.

Tampon d'extraction : 200mM Tris-HCl pH 7.5, 250mM NaCl, 25 mM EDTA, 10mM SDS

Méthode 2 suivie par une PCR avec des amorces de faible affinité

Deux feuilles de 1 cm sont mises en présence de billes de verre et de 500 µl de **tampon d'extraction**. Le matériel végétal est ensuite broyé à l'agitateur-broyeur Precellys (2 cycles de broyage de 30 s à 25000 g avec une pose de 30 sec). 4 mM de SDS sont ensuite ajoutés au broyat obtenu. Ce détergent permettra de dénaturer les protéines et d'éviter que les membranes ne s'agrègent avec l'ADN. Après 15 min d'incubation à température ambiante, les membranes sont culottées par une centrifugation de 10 min à 16000 g et à 20°C. Au surnageant est ajouté 1.5 M d'acétate de potassium pH 8 afin de précipiter le SDS et les protéines associées. Après avoir délicatement homogénéisé la solution, le SDS et les protéines sont culottées par une centrifugation de 15 min à 16000 g et à 4°C. Au surnageant est ajouté 0.7 V d'isopropanol afin de précipiter l'ADN. Après un temps d'incubation de 15 min à -20°C, l'ADN est culotté par une centrifugation de 15 min à 16000 g et à 4°C. Une fois le surnageant éliminé, l'ADN est repris dans 50 µl de **R40**.

Tampon d'extraction : 0.1 M Tris-HCl ph8, 0.05 M EDTA, 0.5 M NaCl, 1 % (p/v) PVP40

R40 : 10 mM Tris/HCl pH8, 1 mM EDTA, 40 µg/ml de RNase A

3.2.1.10 MUTAGENESE DIRIGEE

La mutagenèse dirigée est réalisée avec le kit "QuikChange II Site-directed Mutagenesis Kit" (Stratagene) (figure 18). Elle consiste en une synthèse par PCR de deux brins complets de plasmide à l'aide de deux amorces complémentaires portant la mutation à introduire. Les cycles suivants vont permettre la formation de brin d'ADN linéaire complémentaire au brin parental circulaire. Les produits obtenus seront digérés

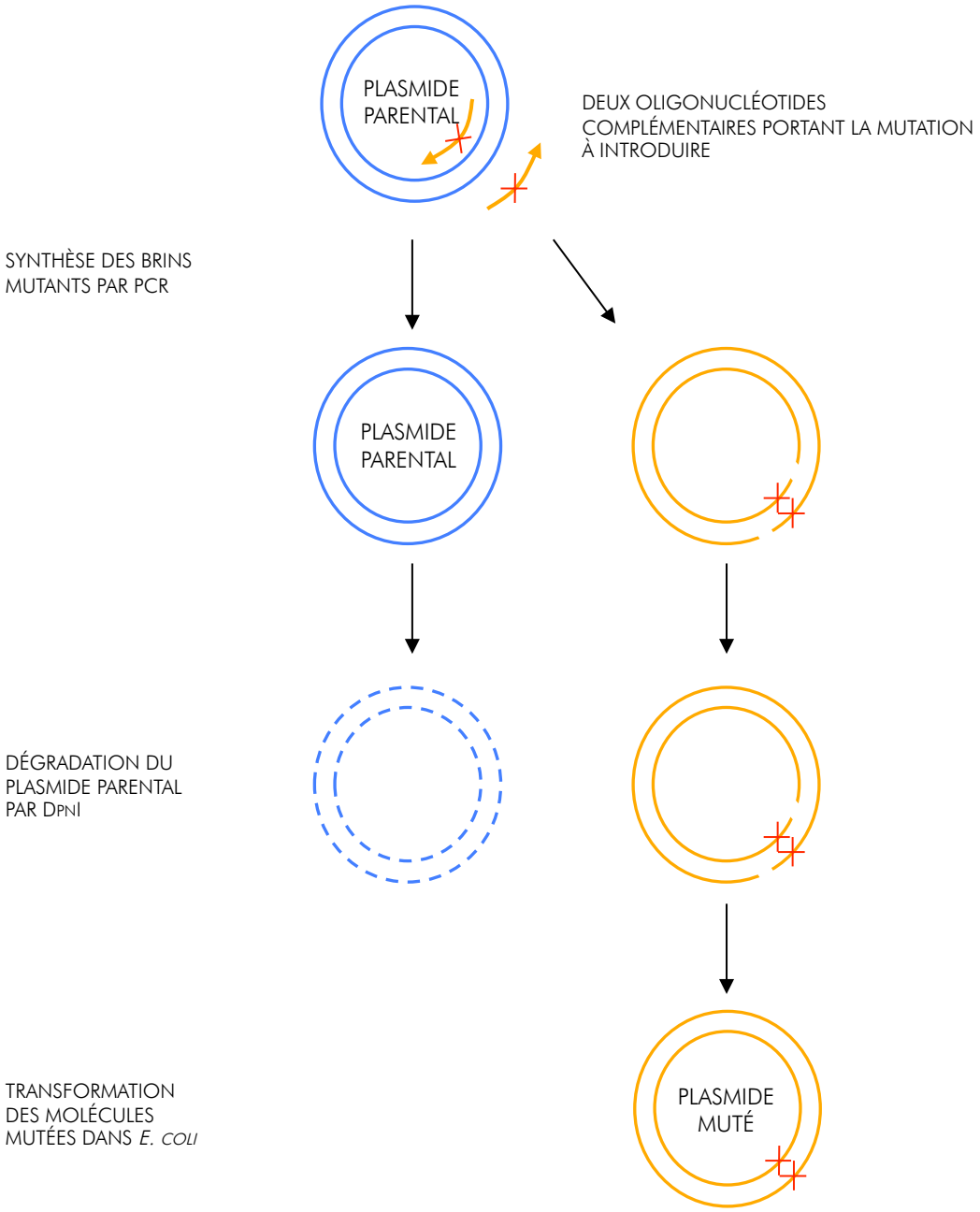


Figure 18 : PRINCIPE DE LA MUTAGENÈSE DIRIGÉE ("QUIKCHANGE II SITE-DIRECTED MUTAGENESIS KIT").
En bleu sont représentés les plasmides parentaux purifiés à partir de bactéries et en orange les plasmides néosynthétisés. Une étape de digestion par l'enzyme *DpnI* est réalisée. Elle a pour but de dégrader les brins parentaux méthylés dans les bactéries. Les brins néosynthétisés portant la mutation sont transformés dans *E. coli*. Ces brins linéaires sont recircularisés par le système de réparation bactérien de l'ADN.

par l'enzyme de restriction DpnI qui ne clive que les brins d'ADN méthylés. Les molécules plasmidiques initiales, extraites de bactéries, possèdent de telles modifications, ce qui n'est pas le cas des molécules synthétisées *in vitro*. Les plasmides non digérés possédant la mutation seront transformés dans *E. coli*. Ils seront circularisés *in vivo* grâce aux mécanismes de réparation de l'ADN des bactéries.

Le milieu réactionnel de 50 µl a été réalisé en prenant 30 ng de l'ADN plasmidique codant pour le gène à modifier, 1 µl de dNTP du mix fourni par le kit, 0.5 µl de chaque amorces (62.5 ng), 5 µl de tampon de PCR 10x et 2.5 U de PfuUltra HF ADN polymérase (*Pyrococcus furiosus*).

Les cycles de PCR ont été programmés comme suit :

première étape de 30 sec à 95°C

un cycle répété 15 fois de 30 sec à 95°C, 1 min à 55°C, 6 min 30 à 68°C.

Par la suite, 5 U de DpnI sont ajoutés au milieu de PCR puis le mélange est incubé 1h à 37°C afin de digérer les brins parentaux non mutés. Les brins néosynthétisés contenant la mutation seront circularisés *in vivo* après transformation dans *E. coli*.

3.2.2 ARN

3.2.2.1 EXTRACTION D'ARN TOTAL DE PLANTE

<u>RNeasy Plant Mini kit</u>

L'extraction est réalisée selon le protocole du fournisseur (Invitrogen).

<u>Extraction au trizol</u>

Le matériel végétal congelé est broyé dans un mortier reposant dans de l'azote liquide. À 150 mg de broyat obtenu est ajouté 1.5 ml de réactif Trizol, qui va permettre de dénaturer les protéines, les séparant des acides nucléiques. Après homogénéisation, le mélange est laissé 5 min à température ambiante. 0.2 V de chloroforme est ensuite ajouté, permettant de séparer les solvants présents dans le Trizol en deux phases, une phase aqueuse contenant les ARN, une phase organique renfermant les protéines et entre, une interphase, contenant l'ADN. Ce mélange est incubé à température ambiante pendant 5 min, puis les ARN sont séparés des protéines et des ADN par une centrifugation à 16000 g et à 4°C pendant 15 min. À la phase aqueuse est ensuite ajouté 0.5 V d'isopropanol, afin de précipiter les ARN. Après 10 min à température

ambiante, les ARN sont culottés par une centrifugation de 5 min à 16000 g et à 4°C. Le culot est ensuite lavé à l'éthanol 70 % et repris dans 20 à 50 µl d'H_2O.

3.2.2.2 TRANSCRIPTION *IN VITRO*

Le plasmide servant de matrice à la transcription *in vitro* comprend le promoteur de la T7 ARN polymerase. Ce plasmide est tout d'abord digéré par un enzyme de restriction à la position précise correspondant à la fin du transcrit que l'on veut obtenir.

RiboMAX™ (Promega)

Pour la transcription, le volume réactionnel de 10 µl est constitué de 150 à 200 ng d'ADN plasmidique, 7.5 mM de rNTP, 1 µl du mélange enzymatique contenant l'ARN polymérase ADN dépendant du phage T7 (fourni par RiboMAX™ Large Scale, Promega) et le tampon de l'ARN polymérase 1x. Après 4 h d'incubation à 37°C, 1 U de DNase RQ1 (fourni par RiboMAX™ Large Scale, Promega) est ajoutée afin de digérer l'ADN plasmidique pendant 15 min à 37°C. Afin de purifier l'ARN obtenu, deux extractions au phénol-chloroforme sont effectuées selon les méthodes standards (Sambrook and Russell, 2001).

TranscriptAid™ (Fermentas)

La transcription est réalisée dans un volume réactionnel de 10 µl composé de tampon "TranscriptAid" 1x, 10 mM rNTP, 1 µg de matrice et 1 µl de "TranscriptAid™ Enzyme Mix", contenant l'ARN polymérase ADN dépendant du phage T7. Après 6 h d'incubation à 37°C, la matrice est digérée par 1 U de DNase I (fourni par TranscriptAid™, Fermentas). Afin de purifier l'ARN obtenu, deux extractions au phénol-chloroforme sont effectuées selon les méthodes standards (Sambrook and Russell, 2001).

3.2.2.3 MARQUAGE EN 5' DE L'ARN PAR LA POLYNUCLEOTIDYLE KINASE

Cette technique consiste à échanger un phosphate à l'extrémité 5' du transcrit par un phosphate radioactif.

Déphosphorylation des transcrits

Le volume réactionnel comporte 150 à 200 ng de transcrits, 20 U de RNaseOUT (Invitrogen) pour éviter une dégradation de l'ARN, 1 U de phosphatase alkaline (FastAP : fermentas) et son tampon 1x.
La réaction est incubée pendant 30 min à 37°C.
Les ARN déphosphorylés sont ensuite purifiés par deux extractions au phénol/chloroforme suivi de deux extractions au chloroforme, puis précipités à l'éthanol.

Phosphorylation radioactive
Elle est réalisée par la polynucleotidyl kinase (PNK) qui ajoute un phosphate en 5' d'une molécule d'ARN déphosphorylée à partir d'une molécule d'ATP comportant un ^{32}P en position γ. Le volume réactionnel est constitué de 100 à 150 ng de transcrit, 4 U de RNaseOUT, 50 μCi ^{32}P ATP, 1 U de PNK et le tampon A 1x de la PNK (Fermentas). La réaction est effectuée pendant 1 h à 37°C.

3.2.2.4 ELECTROPHORESE SUR GEL DE POLYACRYLAMIDE

Le gel utilisé pour analyser les ARN est un **gel de polyacrylamide dénaturant** (8 % à 17 %). Après dépôt des échantillons associés à du **tampon de charge 1x**, la migration est réalisée dans du **tampon TBE 1x** à 200 V pour un gel 8 % et 350 V pour un gel 17 %, sous un ampérage constant de 25 mA.

Gel d'acrylamide dénaturant à x % pour ARN : x % acrylamide/bisacrylamide 19/1, 7 M urée, 1x TBE
Tampon TBE 10x : 900 mM Tris, 25 mM EDTA, 900 mM borate
Tampon de charge 2x : 95 % (v/v) formamide, 20 mM EDTA, 0.05 % (p/v) bleu de bromophénol, 0.05 % (p/v) xylène cyanol

3.2.2.5 ELECTROPHORESE SUR GEL D'AGAROSE

Le gel utilisé pour séparer les ARN de grande taille est un **gel d'agarose dénaturant 1,2 %**. Après dépôt des échantillons additionnés à du **tampon de charge**, la migration est exécutée dans du **tampon de migration** à 100 V.

Gel d'agarose dénaturant à x % pour ARN : **x %** : x % (p/v) d'agarose, 1x MOPS pH 7, 5 % (v/v) formaldéhyde

Tampon de charge : 50 % (v/v) formamide, 20 % (v/v) formaldéhyde, 1x MOPS, 10 % (v/v) glycérol, 0.0025 % (p/v) BET, 0.01 % (p/v) bleu de bromophénol

Tampon de migration : 1x MOPS pH 7, 5 % (v/v) formaldéhyde

3.2.2.6 ANALYSE DE TYPE "NORTHERN BLOT"

Electrotransfert des ARN sur membrane

Les ARN séparés sur un gel acrylamide 7 % urée, sont électrotransférés sur une membrane de nylon chargée positivement (Hybond^{TM}-N+, 0.45 µm, Amersham). Il est réalisé à l'aide d'un appareil de transfert liquide en présence de **tampon TAE 0.25x**, pendant 45 min à 500 mA. La fixation des acides nucléiques sur la membrane est effectuée par un pontage covalent aux UV, grâce à l'appareil pour réticulation Stratalinker 2400 (Stratagene).

Transfert des ARN par capillarité

Les ARN séparés sur gel d'agarose dénaturant, sont transférés par capillarité sur une membrane de nylon chargée positivement (Hybond^{TM}-N+, 0.45 µm, Amersham). Cette diffusion est réalisée durant 18 h, grâce à du tampon 10x SSC. La fixation des acides nucléiques sur la membrane est effectuée par un pontage covalent aux UV grâce à l'appareil pour réticulation Stratalinker 2400 (Stratagene).

Préparation de la sonde

Le marquage en 5' des sondes est réalisé dans un volume réactionnel de 10 µl, comprenant du tampon kinase A 1x, 20 µM oligonucléotide (Sigma et Eurofins), 25 µCi ^{32}P γATP et 10 U PNK (Fermentas). La réaction a lieu durant 1 h à 37°C. 50 µl d'H_2O sont ajoutés avant de séparer la sonde des nucléotides non incorporés par filtration sur gel de type G50.

Hybridation de type northern

La membrane est incubée 14 h à 45°C dans le **tampon d'hybridation** en présence de la sonde radiomarquée. Après 2 lavages de 10 min avec du **SSC 2x** et un lavage avec une solution SSC 2x, 0.1 % SDS, la membrane est exposée sur une plaque

radiosensible phosphorescente. Le signal obtenu est ensuite observé par le phosphorimager FLA-7000.

Tampon TAE : 40 mM Tris-acétate pH 8.0 ; 1 mM EDTA
20x SSC : 3 M chlorure de sodium, 300 mM citrate de trisodium, pH 7 ajusté à l'HCl
Tampon d'hybridation : 5x SSC, 0.1 % (p/v) SDS

3.2.2.7 PURIFICATION DES TRANSCRITS EXTRAITS SUR GEL DE POLYACRYLAMIDE

Les transcrits sont séparés par électrophorèse sur un gel de polyacrylamide. Les bandes correspondant aux transcrits d'intérêt sont coupées et les ARN sont élués pendant 16 h à température ambiante, en présence d'un **tampon d'élution** avant d'être précipités à l'éthanol.

Tampon d'élution : 0.5 M NH_4Ac, 0.1 % (p/v) SDS, 5 mM EDTA

3.2.2.8 RT-PCR POUR L'OBTENTION D'ADNc TOTAL

Cette technique consiste à synthétiser des brins d'ADN complémentaire (ADNc) aux ARN. Ces ADNc seront obtenus grâce à une ADN polymérase ARN dépendante.
Pour l'obtention d'ADNc représentant les ARN totaux, le volume réactionnel initial de 13 µl contient (concentrations ajustées pour un volume réactionnel de 20 µl) 1.25 µM d'oligo$(dT)_{20}$, 0.5 mM dNTP, 0.5 ng/µl random primer et 1 µg d'ARN. La solution est incubée 5 min à 65°C, puis 5 min dans la glace. Puis sont ajoutés 5 mM DTT, 1x tampon "First-Strand", 40 U RNaseOUT (Invitrogen), 200 U SuperScript III RT (Invitrogen) pour un volume final de 20 µl. Le milieu réactionnel est alors incubé 5 min à 25°C, 1 h à 50°C puis l'enzyme est inactivée 15 min à 70°C. 1 µl d'ADNc obtenu est utilisé pour l'amplification par PCR.

3.2.2.9 cRT-PCR

Cette technique est divisée en trois étapes : une circularisation de l'ARN, une rétrotranscription et une PCR.

La circularisation est réalisée dans un volume réactionnel de 25 µl comprenant 500 ng à 1.5 µg d'ARN, de 5 U T4 RNA ligase (New England Biolabs), 20 U RNaseOUT (Invitrogen). Cette réaction est réalisée 1 h à 37°C puis l'enzyme est inactivée 15 min à 65°C.

L'ARN circularisé est utilisé comme matrice pour la rétrotranscription spécifique. Le volume réactionnel de 13 µl comporte 2 pmol d'amorce gène-spécifique, 0.5 mM dNTP (concentration dans 20 µl) et 10 µl d'ARN circularisé. La solution est incubée 5 min à 65°C, puis 5 min dans la glace. Puis sont ajoutés 5 mM DTT, 1x tampon "First-Strand", 40 U RNaseOUT (Invitrogen), 200 U SuperScript III RT (Invitrogen) pour un volume final de 20 µl. Le milieu réactionnel est alors incubé 1 h à 52°C, puis l'enzyme est inactivée 15 min à 70°C.

1 µl de l'ADNc obtenu est utilisé comme matrice pour l'amplification par PCR.

3.2.2.10 QRT-PCR

Cette technique consiste à mesurer l'expression d'un gène, i.e. à mesurer la quantité d'un ARN cible par l'intermédiaire d'un ADNc (voir précédemment pour la synthèse des ADNc totaux). Le volume réactionnel de 10 µl comporte 1x SYBR Green I Master, 0.25 µM d'amorces sens et antisens et 1 µl d'ADNc. Un triplicat technique est réalisé pour chaque triplicat biologique.

La qRT-PCR est réalisée par le LightCycler 180 Real-Time PCR System (Roche). La PCR commence par une dénaturation initiale de 5 min à 95°C, suivie de 40 cycles à 95°C de 10 sec, à 60°C de 15 sec et à 72°C de 15 sec. À chaque cycle la quantité de produit est mesurée, permettant de mesurer le Ct (crossing threshold), qui est le point pour lequel le signal fluorescent est significativement supérieur au bruit de fond. Il apparaît au début de la phase exponentielle.

À la fin des cycles d'amplification, une courbe de fusion est obtenue en chauffant les produits PCR de 55 à 95°C, permettant ainsi de vérifier qu'il n'y a qu'un amplifia.

Les résultats obtenus sont normalisés en utilisant des gènes représentant des contrôles endogènes et analysés en utilisant les programmes *geNorm* et *NormFinder*. L'efficacité de PCR est calculée par le programme *LinRegPCR*. Les coefficients de corrélation obtenus pour chaque paire de primer excèdent 0.99. Une fois les efficacités déterminées la quantité d'ADN est obtenue par la formule suivante :

$E_{\text{référence}}^{\Delta Ct\ \text{référence}} / E_{\text{cible}}^{\Delta Ct\ \text{cible}}$.

3.2.3 PROTÉINES

3.2.3.1 EXPRESSION DANS *E. COLI* ET PURIFICATION

Préculture

L'expression de protéines nécessaires aux tests d'activité *in vitro* est réalisée dans les souches d'*E. coli* BL21(DE3), transformées par une construction où le gène d'intérêt est cloné dans le vecteur pET28 (Novagen). Une préculture de 14 h à 37°C est réalisée dans 10 ml de LB, 1 % glucose (afin d'éviter l'expression de la protéine recombinante) et 50 µg/ml kanamycine. La préculture est ensuite stockée à 4°C afin de ralentir la croissance bactérienne et d'éviter d'avoir une préculture trop dense.

Induction

2.5 ml de préculture sont utilisés pour ensemencer 200 ml de milieu LB, 1 % glucose et 50 µg/ml kanamycine, qui seront mis en culture pendant 6 h à 37°C. Les bactéries sont ensuite sédimentées par une centrifugation à 5000 g et 4°C pendant 15 min. Les bactéries sont lavées avec 20 ml de milieu LB, puis centrifugées pendant 5 min à 5000 g et 4°C. Le culot est ensuite repris dans 200 ml de milieu LB, 50 µg/ml kanamycine et 1mM IPTG. L'induction a lieu pendant 14 h à 17°C sous agitation.

Lyse des cellules

Les bactéries ayant surproduit la protéine d'intérêt sont centrifugées 15 min à 5000 g et 4°C. Le culot est remis en suspension dans 50 ml de **PBS 1x**. Les bactéries sont à nouveau sédimentées par une centrifugation de 15 min à 5000 g et 4°C. Cette étape est répétée une seconde fois. Le culot est ensuite remis en suspension dans 15 ml de $\mathbf{LB_{pET}}$, puis les cellules sont culottées dans les mêmes conditions de centrifugation. Les cellules sont reprises dans 5.5 ml de $\mathbf{LB2_{pET}}$, avant d'être lysées par trois cycles de dépression de 800 bar avec une presse de type "French".

Purification des protéines

Le surnageant est filtré à 0.2 µm, puis le pH est ajusté à 7.8. À ce pH, l'histidine est chargée négativement et peut être fixée par des métaux comme le Ni^{2+}. Les ARN sont supprimés en présence de 10 µg/ml de RNase A (50 U/ml). L'extrait protéique est ensuite mis en présence de 200 µl de billes de Ni-NTA (nickel-nitrilotriacetic acid resin) (Qiagen). L'ensemble est incubé pendant 16 heures à 4°C sur un agitateur rotatif, puis transféré dans une colonne, préalablement équilibrée avec 1 ml de $\mathbf{LB2_{pET}}$.

La colonne est alors rincée avec 5 ml de **$WB1_{pET}$**, puis 5 ml de **$WB2_{pET}$**. Les protéines fixées sur les billes de Ni-NTA sont ensuite éluées avec 4 ml de **$EL1_{pET}$** et 4 ml de **$EL2_{pET}$**.

Pour PRORP1 seconde purification avec colonne échangeuse de cations HiTrap SP HQ

La colonne HiTrap SP HQ (Amersham) est équilibrée avec 10 V de **EL_{50}**. L'élution **$EL1_{pET}$**, diluée avec 2 V de **EL_{150}**, est chargée sur la colonne. La colonne est rincée avec 5 V de **EL_{100}**, puis les protéines sont éluées avec 5 V de **EL_{750}**.

PBS 10x : 1,36 M NaCl ; 26,8 mM KCl ; 85 mM Na_2HPO_4 ; 14 mM KH_2PO_4

LB_{pET} : 20 mM MOPS pH 7.8, 500 mM NaCl, 15 % (v/v) glycérol

$LB2_{pET}$: dans LB_{pET}, 0.2 mM DTT, 1/175 protease inhibitor, 50 mM imidazole

$WB1_{pET}$: 20 mM MOPS pH 7.8, 150 mM NaCl, 15 % (v/v) glycérol, 50 mM imidazole

$WB2_{pET}$: 20 mM MOPS pH 7.8, 250 mM NaCl, 15 % (v/v) glycérol, 75 mM imidazole

$EL1_{pET}$: 20 mM MOPS pH 7.4, 250 mM NaCl, 15 % (v/v) glycérol, 200 mM imidazole

$EL2_{pET}$: 20 mM MOPS pH 7.4, 250 mM NaCl, 15 % (v/v) glycérol, 500 mM imidazole

EL_{50} : 20 mM MOPS pH 7.4, 50 mM NaCl, 15 % (v/v) glycérol

EL_{100} : 20 mM MOPS pH 7.4, 100 mM NaCl, 15 % (v/v) glycérol

EL_{150} : 20 mM MOPS pH 7.4, 150 mM NaCl, 15 % (v/v) glycérol

EL_{750} : 20 mM MOPS pH 7.4, 750 mM NaCl, 15 % (v/v) glycérol

3.2.3.2 DOSAGE DES PROTEINES PAR LA METHODE DE BRADFORD

Le dosage des protéines a été réalisé soit par Bradford, soit à l'aide d'un NanoDrop 1000 (Thermo scientific).

La technique de Bradford permet de mesurer la quantité de protéines dans une solution. Elle repose sur un dosage colorimétrique se traduisant par un changement d'absorbance à 595 nm.

On ajoute 900 µl de **réactif de Bradford** à 100 µl de solution de protéines et l'absorbance à 595 nm est mesurée. La comparaison avec une gamme étalon à base de BSA permet de déterminer la quantité de protéines dans la solution.

Réactif de Bradford : 0.01% (p/v) bleu de Coomasie G-250, 4.7 % (v/v) éthanol, 8.5 % (v/v) acide orthophosphorique

3.2.3.3 ELECTROPHORESE SUR GEL DE POLYACRYLAMIDE EN CONDITIONS DENATURANTES (SDS-PAGE)

Les protéines obtenues sont analysées par SDS-PAGE. Cette technique permet de séparer les protéines selon leur taille. Les protéines dans du **tampon de charge 1x** migrent, sous ampérage constant de 25 mA, sur un gel divisé en deux parties : un **gel de concentration** et un **gel de séparation** (8 à 12 %). La migration s'effectue dans du **tampon Laemmli 1x**. À la fin de la migration, le gel peut être coloré ou transféré sur membrane.

Tampon de charge x2 : 100 mM Tris-HCl pH 6.8 ; 2 % (p/v) SDS ; 10 % (v/v) glycérol ; 3 % (v/v) β-mercaptoéthanol ; 0,1 % (p/v) bleu de bromophénol
Gel de concentration : 5 % acrylamide/bisacrylamide 37.5/1 ; 0,125 M Tris-HCl pH6.8 ; 0,1% (p/v) SDS ; 0,1 % (v/v) APS ; 0,01 % TEMED
Gel de séparation à x % : x % acrylamide/bisacrylamide 37.5/1 ; 0.375 M Tris-HCl pH 8.8 ; 0.1 % (p/v) SDS ; 0,1 % (v/v) APS ; 0,01 % TEMED
Tampon Laemmli 10x : 250 mM Tris HCl pH 8.3 ; 2 M glycine ; 1 % (p/v) SDS

3.2.3.4 COLORATION DU GEL DE POLYACRYLAMIDE

Le gel est incubé dans une solution de 12 % (v/v) TCA pendant 30 min, afin de fixer les protéines. Après cela le gel est coloré grâce à une **solution de coloration** pendant une nuit, puis est décoloré pendant 20 min dans une solution de 25 % (v/v) méthanol.

Solution de coloration des gels : 3 % (v/v) acide orthophosphorique, 6 % (p/v) sulfate d'ammonium, 0.1 % (v/v) Coomassie G250

3.2.3.5 TRANSFERT DES PROTEINES SUR MEMBRANE

Transfert semi-sec

Les protéines, séparées sur SDS-PAGE, sont électrotransférées sur une membrane hydrophobe de PVDF (Immobilon-P, 0.45 µm, Millipore). Ceci est réalisé dans un appareil de transfert semi sec (TRANS-BLOT SDr, BIO-RAD) en présence de **tampon de transfert semi-sec,** pendant 1 h à 25 V et 70 mA. La membrane peut être ensuite colorée dans une solution de **coloration de membrane**, puis lavée avec une solution de **décoloration de membrane**.

Transfert liquide

Les protéines, séparées sur SDS-PAGE, sont électrotransférées sur une membrane hydrophobe de PVDF (Immobilon-P, 0.45 µm, Millipore). Ceci est réalisé dans un appareil de transfert liquide (Mini Trans-Blot® Cell, BIO-RAD) en présence de **tampon de transfert liquide,** pendant 1 h, à 260 mA et à 4°C. La membrane peut être ensuite colorée dans une solution de **coloration de membrane**, puis lavée avec une solution de **décoloration de membrane**.

Tampon de transfert semi-sec : 38 mM glycine, 48 mM Tris base, 3.74 % (v/v) SDS
Tampon de transfert liquide : 25 mM Tris, 192 mM glycine, 15 % méthanol
Solution de coloration de membrane : 0.1 % (p/v) bleu de Coomassie, 50 % (v/v) méthanol, 7 % (v/v) acide acétique
Solution de décoloration des membranes : 50 % (v/v) méthanol, 7 % (v/v) acide acétique

3.2.3.6 IMMUNODETECTION DE TYPE WESTERN

Les sites de fixation aspécifique de la membrane sont bloqués par une solution de **PBS-Tween** et 5 % (m/v) de lait. La membrane est ensuite incubée 1 h avec des anticorps monoclonaux de souris orientés contre l'étiquette histidine de la protéine de fusion, utilisés à une dilution au 1/ 5000 dans du PBS-Tween et 2 % (m/v) de lait. Après trois lavages de 5 min avec du PBS-Tween, la membrane est incubée 30 min avec un anticorps anti-immunoglobuline de souris couplé à la peroxydase et dilué au 1/5000 dans du PBS-Tween. La membrane est lavée trois fois avant l'ajout de 10 µl/cm^2 de substrat de la peroxydase ("Enhanced Chemical Luminscence". Amersham

Phamacia Biotech). La protéine d'intérêt est révélée grâce à l'activité de la peroxydase qui produit une émission lumineuse impressionnant un film photographique.

PBS-Tween : PBS 1x, 0.2 % (v/v) Tween 20

3.2.3.7 TEST D'ACTIVITE

Deux types de test d'activité ont été réalisés : un radioactif afin de détecter l'extrémité 5' libérée et un non radioactif pour observer l'ARNt mature. Après le test d'activité, les protéines sont supprimées par une réaction au phénol-chloroforme (Sambrook and Russell, 2001), puis les ARN sont déposés sur un gel 8 % acrylamide, 7 % urée, 1x TBE et visualisés par bromure d'éthidium ou radiographie.

Test d'activité en présence d'ARNt radiomarqués

La réaction est réalisée dans 10 µl avec 100 ng de protéines recombinantes purifiées (voir section 3.2.3.1.) et 100 ng d'ARN radiomarqué en 5', dans un tampon contenant 20 mM Tris HCl pH 8, 30 mM KCl, 4.5 mM $MgCl_2$, 20 µg/ml BSA et 2 mM DTT. La réaction est réalisée durant 15 min à 20°C.

Test d'activité en présence d'ARNt non radiomarqués

100 ng de protéines recombinantes purifiées (voir section 3.2.3.1.) est mis en présence de 1 µg d'ARN. La réaction a lieu pendant 15 min à température ambiante dans 20 mM Tris HCl pH 8, 30 mM KCl, 4.5 mM $MgCl_2$, 20 µg/ml BSA et 2 mM DTT.

3.2.4 PLANTES

3.2.4.1 EXPRESSION TRANSITOIRE DANS DES PROTOPLASTES ISSUS DE CELLULES MESOPHYLLE D'*A. THALIANA*

Préparation des protoplastes

Des feuilles d'*A. thaliana* sont sectionnées en fines bandelettes dans un **tampon de plasmolyse** avant d'être infiltrées sous vide par une **solution enzymatique** pendant 5 min. Après 3 h d'incubation à 23°C, les protoplastes sont filtrés à travers une

membrane en nylon puis dilués dans 0.5 V de $CaCl_2$ 200 mM. Les protoplastes sont sédimentés par une centrifugation de 2 min à 100 g. Ils sont ensuite lavés avec 10 mL de **solution W5**. Après une centrifugation de 1 min à 100 g, les protoplastes purifiés sont à nouveau resuspendus dans la **solution W5**, puis incubés dans la glace pendant 30 min.

Tampon de plasmolyse : 500 mM mannitol, 10 mM $CaCl_2$, 10 mM MES-KOH pH 5.6

Solution enzymatique : 1 % cellulase, 0.3 % macerase, 0.4 M mannitol ; 20 mM KCl, 20 mM MS-KOH. *10 min à 55°C*. 10 mM $CaCl_2$, 0.1 % BSA

Solution W5 : 125 mM $CaCl_2$, 154 mM NaCl, 2 mM MES-KOH pH 5.6, 5 mM glucose, 5 mM KCl.

Transformation des protoplastes

Les protoplastes sont sédimentés par une centrifugation de 1 min à 100 µg, puis resuspendus dans la **solution mannitol-$MgCl_2$**. 100 µl de la suspension de protoplaste est transfectée par 10 µg d'ADN en présence de 110 µl de **PEG-CMS**. Après une incubation de 30 min à 23°C, la solution de transfection est diluée progressivement par 3.7 ml de **solution W5**. Après une centrifugation de 2 min à 100 g, le culot est resuspendu avec 3 ml de **milieu de culture pour protoplastes**. Après une nouvelle centrifugation de 2 min à 100 g, les protoplastes sont repris dans 1 ml de **milieu de culture pour protoplastes**. Cette solution est ensuite transférée dans une boîte de Pétri de 60 mm de diamètre. Les protoplastes sont mis en culture 24 à 48 h à 23°C dans le noir.

Solution mannitol $MgCl_2$: 0.4 M mannitol, 15 mM $MgCl_2$, 5 mM MES-KOH pH 5.6.

Solution PEG-CMS : 50 % (p/v) PEG4000 (Fluka 95904), 0.25 M mannitol, 0.125 M $Ca(NO_3)_2$

Milieu de culture pour protoplastes (pH 5.8) : 0.245 % (p/v) MS salts + vitamines (Duchefar Biochemie), 350 mM mannitol, 50 mM glucose, 3 mM CaCl2, 2.5 % (p/v) chloramphénicol.

3.2.4.2 VIRUS INDUCED GENE SILENCING (VIGS)

Un protocole de VIGS optimisé pour *Arabidopsis* a été utilisé ici (Burch-Smith et al., 2006). La construction pTRV1, portant les gènes essentiels au virus TRV, est transformée dans la souche d'agrobactéries C58C1 et les constructions pTRV2, possédant la séquence d'intérêt à déréguler, dans la souche d'agrobactéries GV3101.

Les agrobactéries sont mises en préculture dans 1 ml de LB en présence de 50 µg/ml kanamycine et 50 µg/ml rifampicine pendant 24 h à 28°C à 200 rpm. La préculture est ensuite utilisée pour ensemencer le **milieu de culture** et mise en culture pendant 15 h à 28°C sous agitation à 200 rpm. Les agrobactéries sont culottées durant 10 min à 4000 rpm et à température ambiante. Elles sont ensuite lavées dans 10 ml d'eau miliQ, puis sédimentées dans les mêmes conditions, les agrobactéries sont resuspendues dans 5 ml de **milieu d'infiltration**. Le volume est ajusté afin d'obtenir une DO_{600} de 1. Les souches d'agrobactéries C58C1 et GV3101 sont mélangées volume à volume pour chacune des constructions, puis la suspension bactérienne est incubée à température ambiante pendant 1 h. Les plantes d'*Arabidopsis* sont agroinfiltrées avec cette solution au stade 2 à 4 feuilles.

Deux contrôles sont réalisés :

- un contrôle négatif en présence du vecteur vide afin de vérifier que l'effet observé est dû à la construction d'intérêt
- un contrôle positif en présence d'une construction ciblant l'ARNm de la phytoène désaturase pour vérifier l'efficacité de l'agroinfiltration en observant l'apparition de zones blanches caractéristiques de la dérégulation de la phytoène désaturase.

Milieu de culture : LB, 50 µg/ml kanamycine, 50 µg/ml rifampicine, AcSG 20 µM
Milieu d'infiltration : MES 10 mM, $MgCl_2$ 10mM et AcSG 150 µM

3.2.4.3 TRANSFORMATION DE PLANTE PAR FLORALDIP

Une préculture est réalisée pendant 10 h à 28°C et 200 rpm dans 3 ml de LB avec les antibiotiques appropriés. La préculture est utilisée pour ensemencer 150 ml de LB avec antibiotiques qui sont mis en culture pendant 20 h à 28°C et 200 rpm.

Les bactéries sont sédimentées par une centrifugation de 10 min à 5000 rpm. Le culot est ensuite resuspendu dans le **milieu d'infiltration**. Les hampes florales des plantes à

agroinfiltrer sont trempées 30 sec dans la solution bactérienne, puis les plantes sont mises au noir pendant 48 h.

Milieu d'infiltration : ½ MS medium (MO255 avec vitamines et MES (Duchefa)), 5 % (p/v) sucrose, 0.5 ml/L Silvet L-77, 200 µM AcSG

3.2.4.4 STERILISATION DES GRAINES

Cette technique fait intervenir deux étapes de stérilisation. Lors de la première étape, 4 V d'éthanol 96 % sont ajoutés aux graines à stériliser. Les graines sont ensuite mises sous agitation pendant 1 min puis culottées par une centrifugation de 5 sec à 1000 g. Afin de réaliser la deuxième étape de stérilisation, les graines sont reprises dans 6 V d'1/3 Javel. Après une agitation de 5 min, les graines sont culottées par une centrifugation dans les mêmes conditions.
Elles sont ensuite lavées avec 5 V d'H_2O, puis mises sous agitation pendant 5 sec avant d'être culottées par une centrifugation dans les mêmes conditions. Cette étape est répétée 5 fois.

3.2.5 MICROSCOPIE

3.2.5.1 OBSERVATION AU MICROSCOPE CONFOCAL DES PROTOPLASTES TRANSFORMES AFIN DE DETERMINER LA LOCALISATION DES PROTEINES DE FUSION PRORP-EYFP

Les protoplastes sont visualisés au microscope confocal (Zeiss LSM510). Les fluorochromes sont excités à l'aide du laser argon 488 (E-YFP, mitotracker, chlorophylles) ou d'une source UV (DAPI), après sélection de la longueur d'onde d'excitation. L'émission de fluorescence spécifique des différents fluorochromes est filtrée afin d'isoler la lumière émise. La fluorescence du mitotracker ou du DAPI ainsi que l'autofluorescence de la chlorophylle permettront de déterminer la localisation des protéines de fusion en superposant les images obtenues pour les différents signaux.

3.2.5.2 OBSERVATION AU STEREOMICROSCOPE

L'observation des siliques a été réalisée grâce à un stéréomicroscope (SZ2-ILST, Olympus).

4 RESULTATS

4.1 LOCALISATION DES PROTEINES PRORP ET FONCTION DE PRORP1

4.1.1 PRÉAMBULE

La RNase P intervient dans la maturation des ARNt en clivant l'extrémité 5' leader des ARNt précurseurs. Jusqu'à très récemment cette enzyme était considérée comme étant universellement soutenue par un complexe ribonucléoprotéique dont l'ARN, un ribozyme, porte l'activité catalytique (Lai et al., 2012a).

Chez les plantes, l'activité RNase P est présente dans les trois compartiments possédant une expression génétique, i.e. les chloroplastes, les mitochondries et le noyau (Marchfelder, 1995; Schon, 1995). Des études préliminaires, dans les années 1990 ont suggéré que l'activité RNase P dans les plastes d'épinard requiert uniquement la présence de protéine et ne contient pas d'ARN (Thomas et al., 2000). Cependant aucun candidat n'a été identifié au niveau moléculaire. Ces résultats, allant à l'encontre du dogme établi de la conservation universelle de la RNase P ribonucléoprotéique, ont été reçus avec scepticisme.

Ce dogme a été contredit avec la caractérisation de la RNase P mitochondriale humaine (Holzmann et al., 2008). Cette enzyme est composée uniquement de trois protéines, nommé MRPP1 (Mitochondrial RNase P Protein 1), MRPP2 et MRPP3. La présence de ces trois protéines est indispensable pour reconstituer une activité RNase P *in vitro*.

4.1.2 PUBLICATION 1 : A SINGLE *ARABIDOPSIS* ORGANELLAR PROTEIN HAS RNASE P ACTIVITY

Nous avons recherché la présence d'orthologue des protéines MRPP chez *A. thaliana* par des analyses bioinformatiques. Trois orthologues d'une seule des trois protéines, MRPP3, ont pu être identifiés. Ces orthologues, nommés au laboratoire PRORP (pour PROteinaceous RNase P), sont localisés expérimentalement dans les organites (PRORP1) et le noyau (PRORP2 et 3). Les protéines PRORP possèdent en N-terminal des motifs PPR qui pourraient être impliqués dans la liaison à l'ARN et en C-terminal un domaine de type métallonucléase qui pourrait être responsable de l'activité catalytique de ces protéines.

Dans ce travail nous nous sommes ensuite focalisés sur la fonction de la protéine PRORP1, la seule protéine PRORP organellaire. Dans le cas de la RNase P mitochondriale humaine, la présence des trois protéines MRPP est nécessaire pour une activité *in vitro*. Au contraire, la protéine PRORP1 recombinante, seule, catalyse *in vitro* la maturation des précurseurs d'ARNt chloroplastiques et mitochondriaux en clivant au niveau du site canonique de la RNase P. De plus, elle clive des "t-elements" qui sont des structures ressemblant à des ARNt présents sur certains précurseurs d'ARNm mitochondriaux. Enfin les résultats *in vitro* ont permis d'identifier par mutagenèse deux acides aminés, présents dans le domaine métallonucléase, comme essentiels à l'activité de la protéine.

In vivo PRORP1 posséderait également une activité RNase P. En effet la protéine PRORP1 est capable de complémenter une lignée *E. coli* dérégulée pour l'ARN de la RNase P bactérienne. Ainsi nous avons proposé que PRORP1 soit seule responsable de l'activité RNase P dans les organites de plantes. Si tel est le cas alors la fonction de PRORP1 devrait être essentielle. Cette hypothèse a été confirmée par une étude de deux mutants d'insertion ADN-T dans le gène *prorp1*.

Cette première étude établit qu'une protéine seule, peut avoir une activité RNase P. Elle constitue ainsi une avancée importante pour la recherche sur les RNase P. Toutefois les conclusions présentées ici ont été basées uniquement sur des analyses *in vitro* et des complémentations bactériennes.

L'étude de la fonction "*in planta*" des protéines PRORP fait l'objet du deuxième chapitre.

Pour cette étude, mon implication personnelle a plus particulièrement porté sur :

- la purification des protéines PRORP1
- le clonage, la transcription *in vitro* des t-element et l'étude de leur clivage
- la préparation des protoplastes en vue de la localisation des protéines PRORP par microscopie confocale
- la sélection par génotypage des mutants ADN-T

Gobert, A., Gutmann, B., Taschner, A., Gossringer, M., Holzmann, J., Hartmann, R.K., Rossmanith, W., and Giege, P. (2010). A single *Arabidopsis* organellar protein has RNase P activity. Nat Struct Mol Biol 17, 740-744.

A single *Arabidopsis* organellar protein has RNase P activity

Anthony Gobert[1], Bernard Gutmann[1,4], Andreas Taschner[2,4], Markus Gößringer[3,4], Johann Holzmann[2], Roland K Hartmann[3], Walter Rossmanith[2,5] & Philippe Giegé[1,5]

[1]Institut de Biologie Moléculaire des Plantes du Centre National de la Recherche Scientifique, Strasbourg, France.

[2]Center for Anatomy & Cell Biology, Medical University of Vienna, Vienna, Austria.

[3]Institute for Pharmaceutical Chemistry, Philipps-Universität Marburg, Marburg, Germany.

[4]These authors contributed equally to this work. 5These authors contributed equally to this work. Correspondence should be addressed to W.R. (walter.rossmanith@meduniwien.ac.at) or P.G. (philippe.giege@ibmp-cnrs.unistra.fr).

Received 22 September 2009; accepted 19 March 2010; published online 16 May 2010; doi:10.1038/nsmb.1812

The ubiquitous endonuclease RNase P is responsible for the 5′ maturation of tRNA precursors. Until the discovery of human mitochondrial RNase P, these enzymes had typically been found to be ribonucleoproteins, the catalytic activity of which is associated with the RNA component. Here we show that, in *Arabidopsis thaliana* mitochondria and plastids, a single protein called ‘proteinaceous RNase P’ (PRORP1) can perform the endonucleolytic maturation of tRNA precursors that defines RNase P activity. In addition, PRORP1 is able to cleave tRNA-like structures involved in the maturation of plant mitochondrial mRNAs. Finally, we show that *Arabidopsis* PRORP1 can replace the bacterial ribonucleoprotein RNase P in *Escherichia coli* cells. PRORP2 and PRORP3, two paralogs of PRORP1, are both localized in the nucleus.

4.2 FONCTION *IN VITRO* DES PROTEINES PRORP NUCLEAIRES ET *IN PLANTA* DES TROIS PROTEINES PRORP

4.2.1 PRÉAMBULE

Aucun ARN et aucune sous-unité spécifique de la RNase P ribonucléoprotéique n'a été identifié dans les génomes des plantes (Hartmann and Hartmann, 2003). Ceci a suggéré qu'un autre type de RNase P puisse être présent chez les plantes. Chez *Arabidopsis*, nous avons identifié trois protéines qui sont des candidats pour l'activité RNase P. Nous avons montré que la protéine PRORP1, la seule des trois protéines PRORP localisée dans les organites, possède *in vitro* une activité RNase P (Gobert et al., 2010). Nous avons donc cherché à savoir si les deux protéines nucléaires PRORP2 et 3 étaient impliquées dans l'activité RNase P nucléaire. Cependant une autre enzyme ribonucléoprotéique, la RNase MRP, impliquée dans la maturation des ARN ribosomiques se trouve dans le noyau des plantes. Elle présente chez les animaux des sous-unités communes à la RNase P. Ainsi, nous ne pouvions exclure que la RNase MRP ait pu acquérir chez les plantes l'activité RNase P dans le noyau. Nous avons donc aussi étudié la RNase MRP chez les plantes pour chercher à savoir si elle était impliquée ou non dans la maturation des ARNt.

4.2.2 PUBLICATION 2 : PRORP PROTEINS SUPPORT RNASE P ACTIVITY IN BOTH ORGANELLES AND THE NUCLEUS IN *ARABIDOPSIS*

Les protéines PRORP2 et PRORP3, localisées dans le noyau, présentent une forte similarité de séquence. Leur fonction est redondante comme le montre la présence de plantes homozygotes dans les lignées mutantes ADN-T *prorp2* et *prorp3*. De plus cette fonction reste essentielle à la cellule. En effet aucune plante double mutante homozygote pour *prorp2 prorp3* n'a été isolée.

Les protéines PRORP2 et PRORP3 sont des paralogues de PRORP1, ainsi nous avons prédit qu'à l'égal de PRORP1, elles seules possèderaient également l'activité endonucléolytique caractérisant la RNase P. Cette hypothèse a été confirmée *in vitro* en testant l'activité des protéines recombinantes en présence d'ARNt nucléaires.

La stratégie de dérégulation par la technique VIGS a permis de confirmer la fonction des protéines PRORP *in planta* et d'identifier d'autres substrats que les ARNt pour ces enzymes. Ainsi, *in vivo*, PRORP1 est impliquée dans la maturation des ARNt, chloroplastiques et mitochondriaux, et de certains ARNm mitochondriaux en clivant au niveau des t-elements. PRORP2 et PRORP3 sont impliquées dans la maturation des ARNt

nucléaires et, indirectement, dans celle de certains snoARN. En effet il existe chez *Arabidopsis* un précurseur dicistronique ARNt^{Gly}-snoR43. Le clivage en 5' de l'ARNt^{Gly} par la protéine PRORP2 ou PRORP3 est nécessaire pour avoir un clivage par la RNase Z entre l'ARNt et le snoARN.

Des lignées de dérégulation ont été aussi utilisées pour vérifier que l'activité de la RNase MRP est liée à la maturation des ARNr et non à celle des ARNt. Pour cela nous avons réalisé et étudié deux lignés de dérégulation contre les protéines POP1 et POP4, deux sous-unités essentielles à la RNase P et la RNase MRP chez la levure.

Ainsi nos résultats montrent que les plantes représentent le premier exemple d'organismes décrits comme ayant une activité RNase P réalisée uniquement par des protéines et ne nécessitant pas d'enzyme RNP pour cette activité.

L'étude des protéines PRORP2 et PRORP3 a constitué le cœur de ma thèse. Durant ce travail j'ai étudié *in vitro* la fonction des deux protéines nucléaires, qui a été complété par une analyse de cinétique de clivage réalisée par Anthony Gobert. J'ai poursuivi l'étude *in vivo* au niveau des trois protéines, en utilisant les mutants ADN-T et en développant et analysant des lignées de dérégulation par VIGS.

Gutmann, B., Gobert, A., and Giege, P. (2012). PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. Genes Dev **26,** 1022-1027.

PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*

Bernard Gutmann,[1] Anthony Gobert,[1] and Philippe Giegé[2]

Institut de Biologie moléculaire des plantes du CNRS, University of Strasbourg, 67084 Strasbourg, France

[1]These authors contributed equally to this work.

[2]Corresponding author.

E-mail philippe.giege@ibmp-cnrs.unistra.fr.

Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.189514.112.

RNase P is an essential enzyme that cleaves the 5' leader sequence of tRNA precursors. RNase Ps were believed until now to occur universally as ribonucleoproteins in organisms performing RNase P activity. Here we find that protein-only RNase P enzymes called PRORP (for proteinaceous RNase P) support RNase P activity in vivo in both organelles and the nucleus in Arabidopsis. Beyond tRNA, PRORP proteins are involved in the maturation of small nucleolar RNA (snoRNA) and mRNA. Finally, ribonucleoprotein RNase MRP is not involved in tRNA maturation in plants. Altogether, our results indicate that ribonucleoprotein enzymes have been entirely replaced by proteins for RNase P activity in plants.

4.3 AUTRES OUTILS DEVELOPPES LORS DES TRAVAUX DE THESE

Les résultats présentés dans ces deux dernières publications portent sur l'activité RNase P ouvrant des perspectives intéressantes pour la compréhension de l'évolution du monde aux ARN putatif au monde actuel. Toutefois certaines questions fondamentales restent ouvertes. Le premier axe d'étude est la compréhension de l'histoire évolutive des protéines PRORP, leur distribution chez les eucaryotes et par le même biais leur absence chez certains eucaryotes. Un des éléments de réponse pourrait être apporté par l'étude des substrats.

En effet, un défi majeur consistera à déterminer l'ensemble du spectre de substrat des protéines PRORP au niveau du transcriptome entier

Un autre aspect très important consistera à comparer le mode d'action des protéines PRORP avec celui des RNase P RNP. Ce dernier point donnera sûrement des indices déterminants pour la compréhension de l'évolution et la diversification de l'action de la RNase P dans le vivant.

Des travaux complémentaires, non publiés, réalisés lors de ma thèse et décrits ci-dessous ont consisté à développer des outils qui pourraient aider à répondre à ces questions.

4.3.1 OUTILS POUR L'IDENTIFICATION DES SUBSTRATS PRORP

4.3.1.1 VIGS

Nous avons vu que les protéines PRORP intervenaient dans la maturation des ARNt et d'ARNm mitochondriaux. Cependant chez certaines espèces, il existe un spectre de substrat beaucoup plus large, comme par exemple chez les bactéries, où la RNase P est impliquée dans la maturation de l'ARNr 4.5 S, de certains riboswitchs, ou encore d'ARNt-like présents sur certains bactériophages. Il est donc possible qu'il existe une situation similaire chez les plantes.

Cette analyse pourra par exemple être réalisée par séquençage à haut débit de type RNA seq, en comparant les plantes sauvages avec les lignées de dérégulation. Nous disposons d'une lignée de dérégulation pour la protéine PRORP1 et pour les deux protéines nucléaires PRORP2 et 3, grâce à la technique du VIGS.

Cependant l'impact de la dérégulation de PRORP2 dans les lignées ADN-T *prorp3*, conduit à un faible effet sur les ARNt. Au contraire, il est possible qu'une dérégulation de l'expression de PRORP3 dans un mutant *prorp2* KO conduit à un effet moléculaire plus fort. Cette hypothèse va être testée avec un vecteur viral ciblant PRORP3 sur deux lignées ADN-T *prorp2* d'écotypes différents.

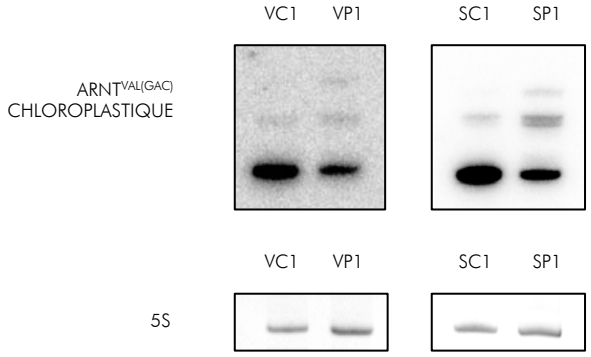


Figure 19 : ACCUMULATION DE PRÉCURSEURS D'ARNT CHLOROPLASTIQUES DANS LES PLANTES SUREXPRIMANT LA PROTÉINE PRORP1

PRORP1 est dérégulée par VIGS par surexpression. Pour cela des plantes sauvages col0 ont été infectées soit par le vecteur viral vide (VC1), soit par le vecteur viral ciblant PRORP1 (VP1), conduisant à une diminution de l'accumulation de PRORP1 (Gutmann et al., 2012a). En parallèle des plantes sauvages col0 ont été transformées de manière stable soit par le gène codant pour la protéine PRORP1 en fusion à l'E-YFP sous la dépendance du promoteur fort 35S (SP1), soit par l'E-YFP seule sous la dépendance du promoteur fort 35S (SC1).
L'impact de cette dérégulation sur l'ARNt$^{Val(GAC)}$ chloroplastique est analysé par northern blot avec une sonde spécifique à l'ARNt$^{Val(GAC)}$. Le niveau de chargement est vérifié grâce à l'ARNr cytosolique 5S.

4.3.1.2 LIGNEES DE SUREXPRESSION

Une lignée surexpresseur prorp1 est à l'étude au laboratoire. Elle résulte d'une transformation stable de plante *A. thaliana* sauvage Col-0 par un ADN-T possédant le gène codant pour la protéine de fusion PRORP1-EYFP sous la dépendance du promoteur fort 35S. Anthony Gobert a observé par western blot que certaines plantes surexprimaient la protéine PRORP1, tandis que cette dernière était fortement diminuée chez certaines plantes résultant probablement d'événements de silencing (Napoli et al., 1990). De telles plantes représenteraient donc une stratégie alternative à l'utilisation de la technique du VIGS.

J'ai alors analysé par northern blot l'impact de cette dérégulation sur l'ARNt$^{Val(GAC)}$ chloroplastique (figure 19). Comme pour les lignées VIGS, une dérégulation de la protéine PRORP1 conduit à une diminution de la maturation de l'ARNt$^{Val(GAC)}$.

Des lignées surexpresseurs pour les deux protéines nucléaires sont en cours de développement au laboratoire.

4.3.1.3 LIGNEES MIARN

Une autre stratégie alternative aux lignées de dérégulation par VIGS est l'utilisation des amiARN (miARN artificiels). Dans cette stratégie, j'ai choisi 4 séquences de 21 nt qui sont soit complémentaires à PRORP1 (A6 et A29), soit complémentaires à PRORP2 et PRORP3 (B44 et B45). L'ADN codant pour ces amiARN a été cloné sous la dépendance d'un promoteur inductible à l'œstradiol et transformé de manière stable dans des plantes sauvage d'*A. thaliana* Col-0.

Les premiers résultats montrent que contrairement aux lignés VIGS, l'induction du amiARN à l'œstradiol pendant 7 jours ne conduit pas à un phénotype macroscopique. Cependant l'expression des amiARN au niveau moléculaire et la dérégulation des ARNm des gènes prorp n'ont pas encore été étudiées. L'évaluation de ces lignées de dérégulation se poursuit au laboratoire.

4.3.2 OUTIL POUR L'IDENTIFICATION DU MODE D'ACTION DES PROTEINES PRORP

La comparaison de tous les substrats pourra aider à la compréhension du mode de reconnaissance des protéines PRORP. Nous avons vu précédemment qu'elles étaient

impliquées dans la maturation des ARNt. Cependant, elles pourraient être également impliquées dans la maturation des ARN SINE. La séquence de ces ARN dérive de celle des ARNt mais ces ARN ont perdu la structure tridimensionnelle en trèfle.

Par ailleurs, une collection de protéines PRORP de séquence mutée est en cours de développement au laboratoire. De même, nous construisons une collection de substrats ARN mutés. Ces outils aideront à la détermination du mode d'action des protéines PRORP.

Durant ma thèse j'ai aussi participé, en collaboration avec Laurence Drouard, à l'élaboration d'une banque informatique regroupant la séquence des ARNt, des ARN SINE et des pseudogènes de cinq plantes à fleur (*Arabidopsis thaliana*, *Medicago truncatula*, *Populus trichocarpa*, *Oryza sativa*, *Brachypordium distachyon*) et d'une algue verte (*Chlamydomonas reinhardtii*). Je me suis plus particulièrement concentré sur l'annotation des ARNt de *Populus trichocarpa* (http://plantrna.ibmp.cnrs.fr/) (manuscrit en cours de préparation, voir en annexe le résumé).

5 DISCUSSION

5.1 LA VIE SANS RNASE P RIBONUCLEOPROTEIQUE

5.1.1 LES PROTÉINES PRORP ONT UNE ACTIVITÉ RNASE P

La RNase P est une endonucléase qui intervient dans la maturation de l'extrémité 5' des ARNt précurseurs. Il y a environ 30 ans Sidney Altman a décrit la RNase P d'*E. coli* comme possédant un ARN responsable de l'activité catalytique de l'enzyme (Guerrier-Takada et al., 1983). Cette découverte, associée à l'étude de l'épissage de l'ARN chez *Tetrahymena* (Cech et al., 1981), a conduit à l'obtention du prix Nobel de chimie en 1989 par Altman et Cech. La RNase P a ensuite rapidement été identifiée comme un complexe ribonucléoprotéique chez différents procaryotes et de nombreux eucaryotes à l'exception notamment des mitochondries chez les animaux et également des organites et du noyau chez les végétaux. En 2008, la RNase P mitochondriale humaine a été caractérisée au niveau moléculaire comme un complexe composé uniquement de trois protéines qui sont essentielles, *in vitro*, à l'activité catalytique du complexe (Holzmann et al., 2008).

Chez les plantes aucune sous-unité protéique et aucun ARN de la RNase P RNP n'est codé par le génome nucléaire ou organellaire. Cependant trois orthologues pour l'une des trois protéines de la RNase P mitochondriale humaine sont codés par le génome nucléaire. Ils sont localisés soit dans le noyau (PRORP2 et 3) soit dans les organelles (PRORP1).

Lors de ma thèse, j'ai réalisé une étude fonctionnelle des trois orthologues. Pour cela, je me suis appuyé sur des tests d'activité *in vitro* ainsi que sur des études *in vivo* de lignées ADN-T et de lignées de dérégulation. Les résultats ont permis d'identifier, pour la première fois, les composants au niveau moléculaire de la RNase P chez les plantes. De plus, contrairement à la RNase P mitochondriale humaine, les protéines PRORP possèdent seules l'activité RNase P.

5.1.2 LES PROTÉINES PRORP SONT ESSENTIELLES

Si la protéine PRORP1 possède seule l'activité RNase P dans les mitochondries et les chloroplastes, on s'attend à ce que sa fonction soit essentielle. Cette hypothèse est confirmée par les résultats obtenus *in vivo* sur des lignées de plantes mutantes. Le caractère essentiel de prorp1 s'explique de toute manière par l'activité de type RNase P de

PRORP1. En effet, sans cette activité, la maturation des ARNt mitochondriaux et chloroplastiques ne peut être complète et des ARNt fonctionnels nécessaires à la traduction ne sont plus produits. Or une absence de traduction mitochondriale et chloroplastique résultent en une absence de mécanisme indispensable à la cellule tel que la phosphorylation oxydative ou la photosynthèse.

De plus une mutation du gène *prorp1* conduit à une létalité embryonnaire au stade globulaire. Les gamètes sont donc viables ainsi que l'embryon durant les premières divisions cellulaires. Ce constat irait à priori à l'encontre du caractère essentiel de PRORP1. Cependant il pourrait s'expliquer par la stabilité des ARNt. Il est possible qu'une quantité suffisante d'ARNt ait pu être synthétisée avant les premières étapes de la méiose. De plus on ne peut pas exclure que PRORP1 ait une demi-vie suffisamment longue pour être encore présente dans les cellules lors des premières étapes de l'embyogenèse.

Les protéines PRORP2 et 3 sont toutes deux localisées dans le noyau et présentent une forte similarité de séquence (77% d'identité des acides aminés), il n'est donc pas étonnant que leurs fonctions soient redondantes. Cette hypothèse a été confirmée par les lignées ADN-T mutantes. De plus le croisement des lignées homozygotes pour ces deux gènes a confirmé le caractère essentiel de la fonction portée par ces deux protéines PRORP nucléaires.

5.1.3 LES PROTEINES PRORP SONT SEULES RESPONSABLES DE L'ACTIVITE RNASE P

Chez la levure et les animaux, la RNase P possède certaines sous-unités communes avec la RNase MRP. Il a donc été suggéré que la RNase MRP ait pu acquérir au cours de l'évolution l'activité RNase P (Krehan et al., 2012). Les résultats que j'ai obtenus avec les lignées de dérégulation dirigées contre POP1 et POP4 réfutent cette hypothèse, en montrant que la RNase MRP intervient dans la maturation des ARNr et non des ARNt. De plus elle semble suggérer que la RNase MRP clive au site A3 qui correspond au site canonique de clivage de la RNase MRP chez la levure et les animaux.

Ainsi les protéines PRORP seraient seules responsables de l'activité RNase P chez les plantes. Ce qui implique que *A. thaliana* serait le premier organisme décrit comme ne possédant pas de ribozymes pour l'activité RNase P.

5.2 PRORP REFLETE-T-ELLE UNE DIVERSITE DE MODE D'ACTION PARMI LES PROTEINES PPR ?

5.2.1 UNE NOUVELLE CATÉGORIE DE PROTÉINES PPR

La famillle des protéines PPR est divisée en deux grandes sous-familles, la sous-famille P présente chez tous les eucaryotes et la sous-famille PLS uniquement présente chez les plantes. Cette dernière famille se distingue de la sous-famille PLS par la présence en C-terminal de motifs autres que les motifs PPR (Lurin et al., 2004). L'expansion de ces deux familles se serait faite par des duplications soit des bloc P, soit des bloc PLS (Rivals et al., 2006).

Ces duplications auraient pu avoir lieu dans des gènes codant par exemple pour des nucléases ou des polymérases, conduisant à une protéine de fusion, constituée en N-terminal de motifs PPR et en C-terminal d'un domaine catalytique. Ces caractéristiques correspondent aux protéines PRORP ou à l'ARN polymerase mitochondriale chez les levures et les animaux. Contrairement aux sous-famille P et PLS, ces protéines présentent peu de motifs PPR, ou des motifs PPR non canoniques, qui ne sont pas forcément répétés en tandem. Ainsi l'ensemble des motifs PPR ne formeraient pas une super-hélice et leur mode de reconnaissance de l'ARN pourrait diverger des protéines PPR canoniques P et PLS.

5.2.2 MODE DE RECONNAISSANCE DES ARN PAR PRORP

Les protéines PPR sont impliquées dans la maturation ou la stabilité des ARNm. Elles peuvent reconnaître les structures secondaires présentes au niveau de ces ARN. Ainsi la protéine PPR10 présente dans les chloroplastes du maïs va reconnaître 17 nt présents en partie dans une tige boucle. Cette reconnaissance, nécessaire à la stabilité et à la traduction des protéines PPR, est analogue à celui du motif Puf, i.e. 1 nt pour 1 motif PPR (Prikryl et al., 2011).

Cependant les études menées *in vitro* suggèrent que les protéines PPR canoniques aient une plus grande affinité pour l'ARN simple brin par rapport à l'ARN double brin (Lurin et al., 2004). La protéine PPR5 comprenant 10 motifs PPR reconnaît l'intron de groupe II de l'ARN truG UCC non épissé (Williams-Carrier et al., 2008). La séquence de reconnaissance est constituée d'un 51-mer, dont la partie centrale forme une tige boucle. La boucle, contrairement à la tige, n'est pas importante pour la reconnaissance. Ainsi, si

l'on assume que cette protéine PPR ne peut reconnaître une séquence plus longue que 40 nt, il est probable que la protéine recouvre la séquence simple brin de l'ARN en sautant la tige boucle.

Les résultats de clivage *in vitro*, qui montrent que PRORP est capable de cliver tout ARNt de structure canonique, laissent supposer que les protéines PRORP pourraient reconnaître la structure et non la séquence de l'ARN. Ainsi PRORP pourrait peut-être reconnaître l'ARN sous forme double brin.

Cependant les ARNt possèdent une structure secondaire constituée de tige-boucle. Les boucles constituées d'ARN simple brin pourraient former une plate-forme d'interaction pour les motifs PPR. Une étude cristallographique du complexe ARNt-PRORP, menée en collaboration avec l'équipe de Claude Sauter (IBMC-CNRS Strasbourg - UPR 9002), devrait permettre de comprendre comment la reconnaissance PRORP/ARN a lieu.

5.2.3 SUBSTRAT MINIMAL

L'étude du mode de reconnaissance des protéines PRORP est aussi intéressante à étudier dans le contexte de l'évolution de l'activité RNase P. Est-il possible que la protéine PPR PRORP et la RNase P RNP partagent des similarités de mode de reconnaissance de l'ARN ? Cela nous amènerait à penser que la protéine PRORP ait pu évoluer pour mimer la RNP. Pour répondre à cette question, la détermination du substrat minimal de PRORP sera importante.

Les tests d'activité *in vitro* ont montré que PRORP1 était capable de cliver les t-elements de nad6 et d'orf138. Une étude complémentaire a été réalisée sur le t-element de ccmC de 71 nt dont les boucles et les tiges D et anticodon sont absentes (résultat non publié). Aucune activité n'est observée *in vitro* pour PRORP1 en présence de ce t-element. Ainsi PRORP1 ne clive pas, ou du moins *in vitro*, un t-element dont la structure s'éloigne beaucoup d'un ARNt canonique. Par contre, pour le t-element de *nad6*, un clivage est observé. Ceci suggère qu'au moins la tige boucle de l'anticodon n'est pas essentielle à la reconnaissance par PRORP1.

L'importance des boucles D et TΨC pourrait s'expliquer par la présence de positions conservées qui seront reconnues par les motifs PPR des protéines PRORP. Ce type de reconnaissance est observé pour la RNase P de type bactérien. En effet les positions G19 et C56, présentent respectivement dans les boucles D et TΨC, interagissent avec l'ARN

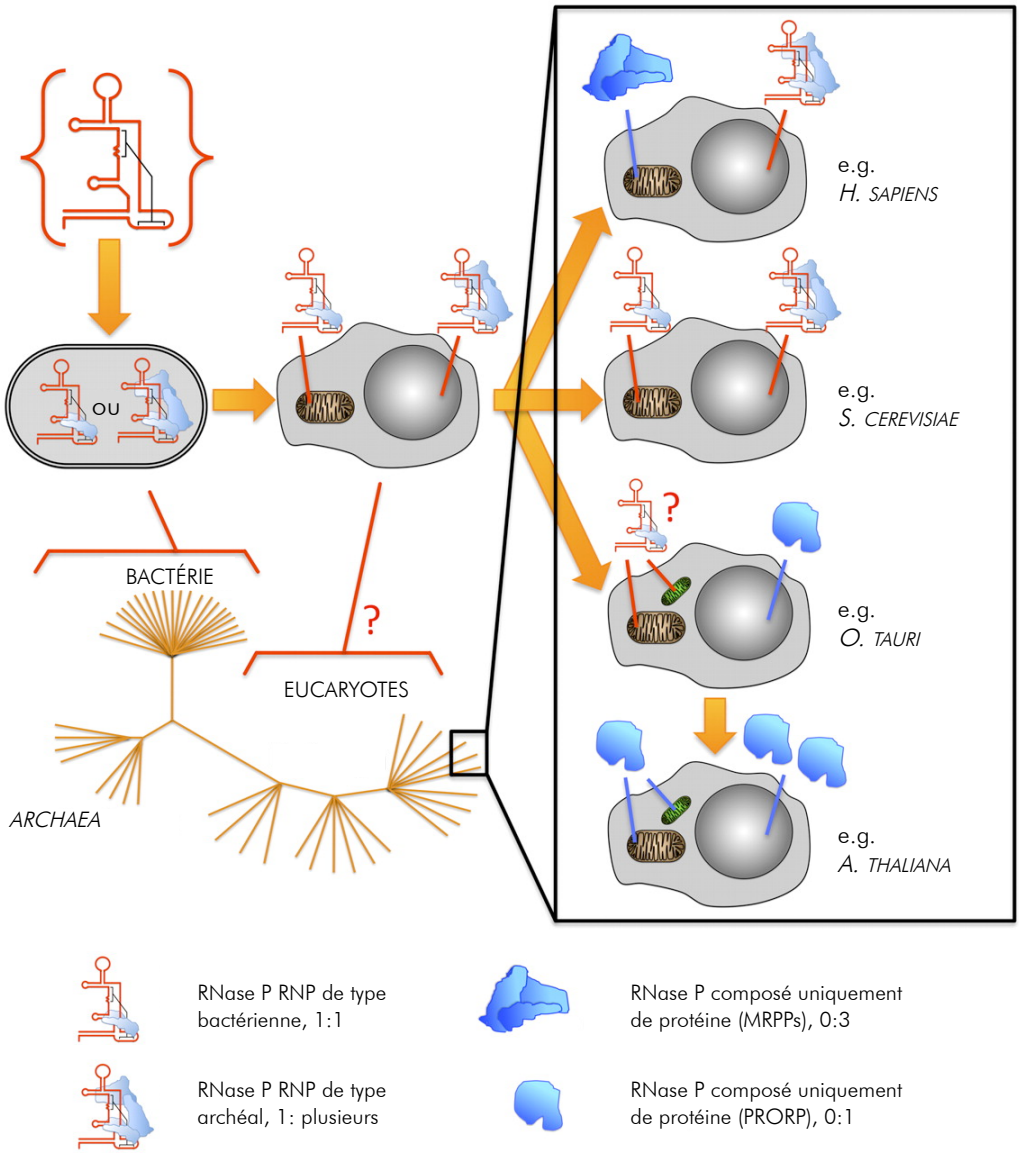


Figure 20 : DIVERSITÉ DE LA RNASE P DANS LE VIVANT (Goldfarb et al., 2012)

L'arbre phylogénétique représente les trois domaines du vivant, i.e. les bactéries, les archaea et les eucaryotes. En rouge est représenté l'ARN de la RNase P de types RNP et en bleu les protéines composant soit la RNase P RNP soit la RNase P protéique. Les flèches en orange représentent l'évolution de la RNase P. Le rectangle de droite représente les différents modèles pour l'activité RNase P chez les eucaryotes.

Le ratio ARN:protéine est représenté pour chaque cas dans la légende.

de la RNase P RNP (Reiter et al., 2010). De même chez les plantes les positions G19 et C56 sont conservées et pourraient être nécessaires à la reconnaissance par la RNase P.

5.3 L'EVOLUTION DES RNASE P

5.3.1 LA PERTE DU RIBOZYME DE LA RNASE P

Chez les eucaryotes, l'activité RNase P est réalisée soit par une RNP, soit par un complexe protéique (figure 20). La RNase P ribonucléoprotéique, dérivant du monde putatif à ARN, est apparue longtemps avant la formation des eucaryotes. Ainsi le génome procaryotique code pour l'ARN et les sous-unités protéiques de la RNase P. La mitochondrie, qui dérive de l'endosymbiose d'une α-protéobactérie, possède elle aussi une RNP chez la levure, le génome mitochondrial codant pour le ribozyme et le génome nucléaire pour l'unique sous-unité protéique. Dans le noyau des animaux et des levures, le ribozyme de la RNase P a évolué et nécessite la présence d'environ 10 protéines supplémentaires pour son activité.
Contrairement à la RNP, la RNase P protéique est uniquement présente chez les eucaryotes, aussi bien dans le monde végétal qu'animal et serait donc apparue tôt dans l'histoire évolutive des eucaryotes.

La levure, les animaux et les plantes pourraient représenter trois étapes de la transition d'une activité RNase P de type ribonucléoprotéique vers une activité RNase P de type protéique. Cette transition pourrait s'expliquer par deux hypothèses.
Nous avons vu précédemment que le ribozyme de la RNase P mitochondriale de la levure était codé par le génome mitochondrial. Il est possible que ce gène ait disparu après une recombinaison du génome mitochondrial ou après un transfert du gène dans le génome nucléaire. Une nucléase présente dans le protéome mitochondrial aurait alors remplacé l'activité de cette enzyme.
Il serait aussi possible que les deux enzymes aient coexisté dans le même compartiment et que la protéine PRORP ou un complexe renfermant la protéine PRORP ait fourni un avantage sélectif par rapport au ribozyme. La RNase P ribonucléoprotéique aurait alors soit disparue comme dans le cas des végétaux, soit évoluée afin de fournir une autre fonction à la mitochondrie. Ce serait peut être le cas chez les animaux, dont l'ARN de la RNase P pourrait être importé dans les mitochondries (Doersen et al., 1985) (Puranam and

Attardi, 2001). Cependant la présence même de ce ribozyme dans les mitochondries reste controversée (Rossmanith, 2012).

Comme pour la RNase P de type RNP, la RNase P de type protéique aurait évolué en s'associant à d'autres protéines. Cette évolution est visible dans les mitochondries humaines avec la formation d'un complexe de trois protéines, probablement nécessaires à la reconnaissance des ARNt non canoniques codés par le génome mitochondrial. Chez les plantes une telle association n'est pas nécessaire pour la maturation des ARNt qui sont tous de type canonique. Cependant on ne peut exclure que la protéine PRORP de plante ne s'associe à d'autres protéines pour la reconnaissance d'autres substrats.

5.3.2 *ARABIDOPSIS* EST-ELLE UNE EXCEPTION CHEZ LES PLANTES TERRESTRES ?

Les résultats de localisation montrent que la protéine PRORP1 est organellaire alors que les protéines PRORP2 et 3 sont nucléaires. La colocalisation de PRORP2 et 3 peut porter à interrogation. La présence de plantes homozygotes *prorp2* et *prorp3* pour l'insertion ADN-T ainsi que l'absence de différence de phénotype entre les plantes KO et les plantes sauvages semblent suggérer que les deux protéines nucléaires ne présentent pas une différence d'expression entre les différents organes et les différents moments du développement. Cependant les deux protéines pourraient être exprimées différemment dans les cellules ou présenter une différence d'activité durant le cycle circadien.
Il est possible que l'activité des protéines PRORP nucléaires soit redondante pour des substrats essentiels, comme les ARNt, alors que l'activité de PRORP2 et 3 pourrait présenter des spécificités pour d'autres substrats non essentiels.

Cependant il est plus vraisemblable qu'*Arabidopsis* représente une exception chez les plantes. En effet les études phylogénétiques réalisées au laboratoire montrent que, dans les espèces n'appartenant pas aux *Brassicaceae*, les protéines PRORP2 et PRORP3 sont plus éloignées et présentent des séquences d'adressage vers des compartiments différents. Par exemple, des études menées au laboratoire ont montré que la protéine PRORP2 de *Brachypodium distachyon* est localisée dans les chloroplatstes. Ainsi chez les plantes, il y a toujours trois protéines PRORP dont, dans la plupart des cas, une serait nucléaire et deux organellaires.

5.3.3 DIVERSITE DE L'EVOLUTION DE LA RNASE P ?

Les études menées jusqu'à présent montrent que différentes évolutions ont permis de maintenir l'activité RNase P chez les eucaryotes.

Chez les plantes et dans les mitochondries des animaux, le ribozyme de la RNase P a été remplacé par la protéine PRORP qui possède l'activité RNase P grâce à son domaine métallonucléase et qui pourrait reconnaître le substrat grâce aux motifs PPR (Holzmann et al., 2008; Gobert et al., 2010).

Chez les archaea et dans le noyau des levures et des animaux l'ARN de la RNase P a perdu des domaines essentiels à son activité comme ceux impliqués dans la reconnaissance du substrat (Ellis and Brown, 2009). Cette perte a été suppléée par la formation d'un complexe comprenant plusieurs sous-unités protéiques.

Dans les mitochondries de levure, la RNase P ne possède pas les caractéristiques de la RNase P de type bactérien. Bien que la RNase P soit constituée, comme chez les bactéries, d'un ribozyme associé à une sous-unité protéique, l'ARN de la RNase P a fortement divergé par rapport à celui de la bactérie (Seif et al., 2003). De plus, la seule sous-unité protéique associée au ribozyme n'a aucune homologie de séquence avec celle de la RNase P bactérienne (Lai et al., 2010a).

Dans les mitochondries d'un grand nombre de champignons non ascomycètes, les acteurs de l'activité RNase P n'ont pas encore été identifiés. Or le ribozyme de la RNase P est absent et le génome nucléaire ne code pas pour des protéines PRORP. Il est possible que l'ARN de la RNase P ait divergé à tel point qu'on ne puisse l'identifier. Cependant on ne peut pas exclure qu'un autre type de RNase P, RNP ou protéique, complètement différent de ceux déjà identifiés, existe aussi.

5.4 INTEGRATION DE LA FONCTION DES PROTEINES PRORP AVEC D'AUTRES MECANISMES D'EXPRESSION GENETIQUE

L'activité RNase P est indispensable pour la maturation des ARNt. Cependant d'autres enzymes, comme la RNase Z, les enzymes de modification de l'ARN, et la CCase, interviennent aussi dans cette maturation. De plus, certaines modifications sont indispensables pour que d'autres soient réalisées. Chez les plantes, c'est par exemple le cas de l'édition des précurseurs d'ARNt^{Phe} et ARNt^{His} dans les mitochondries de mélèze (Fey et al., 2002). En effet, une absence d'édition empêche le repliement structural des ARNt, préalable à la maturation des extrémités 5' et 3'. On peut citer également la maturation du

transcrit dicistronique ARNt^{Gly}-snoR43, dont la coupure endonucléolytique par la RNase Z entre l'ARNt et le snoARN requiert la maturation préalable par la RNase P (Kruszka et al., 2003; Gutmann et al., 2012a).

Ainsi les enzymes de maturation pourraient former des complexes nécessaires à la coordination des différentes activités. Les protéines PRORP, qui possèdent une activité RNase P, pourraient ainsi faire partie d'un ou plusieurs complexes de maturation des ARNt et/ou des ARNm.

Un tel complexe a déjà été identifié dans les mitochondries de levure. En effet un supercomplexe, stable, a été isolé dans les mitochondries. Il comprend des sous-unités protéiques de la RNase P, de la RNase Z et du dégradosome des ARN, ainsi que l'ARN de la RNase P et les ARNr de la petite et de la grande sous-unité du ribosome (Daoud et al., 2012).

Un autre résultat va dans le sens de cette hypothèse dans le noyau des plantes, dont l'activité RNase P coimmunoprécipite avec la protéine POP1, une sous-unité de la RNase MRP (Krehan et al., 2012).

De même chez les plantes, PRORP et d'autres protéines impliquées dans l'expression génétique sont trouvées dans une fraction enrichie en nucléoide de chloroplaste de maïs (Majeran et al., 2012). De plus les protéines PRORP1 sont regroupées en grande quantité dans différentes zones des chloroplastes (Gobert et al., 2010). Ainsi ces protéines pourraient être localisées dans des zones particulières regroupant les différentes machineries impliquées dans l'expression des génomes. Cependant ces observations ont été réalisées avec des protoplastes transformés de manière transitoire, sous le contrôle de promoteurs forts. Il se pourrait donc que les localisations observées ne reflètent pas entièrement la situation *in vivo*.

Afin de rechercher la présence de protéines PRORP dans des complexes dans les différents compartiments des cellules végétales, il serait intéressant d'identifier les protéines interagissant avec les protéines PRORP. Cette identification pourrait se faire par immunoprécipitation des complexes putatifs en utilisant les anticorps spécifiques disponibles et en identifiant les composants des complexes par spectrométrie de masse. Des méthodes génétiques dans la levure peuvent être aussi utilisées, telles que le double-hybride ou la technique du split-ubiquitin afin d'identifier des interactants membranaires. Les outils permettant l'identification des partenaires protéiques sont en cours de développement. Cependant, des résultats préliminaires suggèrent déjà que la protéine PRORP1 serait en

effet associée avec d'autres protéines impliquées dans l'expression génétique dans les organelles de plante.

La découverte des acteurs moléculaires de l'activité RNase P dans les cellules des plantes terrestres représente une avancée intéressante pour la compréhension de l'évolution des systèmes d'expression génétique. En effet la comparaison du modèle *Arabidopsis*, utilisant uniquement des protéines pour l'activité RNase P, avec des modèles possédant soit uniquement des RNP (e.g. *S. cerevisiae*) soit des RNP ou des protéines (e.g. *H. sapiens*) aidera à comprendre comment la transition s'est faite entre le monde prébiotique à ARN et le monde actuel.

ANNEXES

Cognat, V., Pawlak, G., Duchêne, A.M., Daujat, M., Gigant, A., Salinas, T., Michaud, M., Gutmann, B., Giegé, P., Gobert, A., and Maréchal-Drouard, L. (2012). PlantRNA, a database for tRNAs of photosynthetic eukaryotes. Nucleic Acids Res (in press).

Gutmann, B., Gobert, A. and Giegé, P. (2012b) Mitochondrial genome evolution and the emergence of PPR proteins. Advances in botanical research. Vol. Mitochondrial genome evolution, Drouard L. ed (in press)

PlantRNA, a database for tRNAs of photosynthetic eukaryotes

Valérie Cognat, Gaël Pawlak, Anne-Marie Duchêne, Magali Daujat, Anaïs Gigant, Thalia Salinas, Morgane Michaud, Bernard Gutmann, Philippe Giegé, Anthony Gobert and Laurence Maréchal-Drouard*

Institut de Biologie Moléculaire des Plantes, UPR 2357-CNRS, Université de Strasbourg, 12 rue du Général Zimmer, F-67084 Strasbourg Cedex, France

*To whom correspondence should be addressed. Tel: +33 3 67 15 53 98; Fax: +33 3 88 61 44 42; Email: laurence.drouard@ibmp-cnrs.unistra.fr

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

Received August 1, 2012; Revised September 6, 2012; Accepted September 16, 2012

PlantRNA database (http://plantrna.ibmp.cnrs.fr/) compiles transfer RNA (tRNA) gene sequences retrieved from fully annotated plant nuclear, plastidial and mitochondrial genomes. The set of annotated tRNA gene sequences has been manually curated for maximum quality and confidence. The novelty of this database resides in the inclusion of biological information relevant to the function of all the tRNAs entered in the library. This includes 5'- and 3'-flanking sequences, A and B box sequences, region of transcription initiation and poly(T) transcription termination stretches, tRNA intron sequences, aminoacyl-tRNA synthetases and enzymes responsible for tRNA maturation and modification. Finally, data on mitochondrial import of nuclear-encoded tRNAs as well as the bibliome for the respective tRNAs and tRNA-binding proteins are also included. The current annotation concerns complete genomes from 11 organisms: five flowering plants (*Arabidopsis thaliana*, *Oryza sativa*, *Populus trichocarpa*, *Medicago truncatula* and *Brachypodium distachyon*), a moss (*Physcomitrella patens*), two green algae (*Chlamydomonas reinhardtii* and *Ostreococcus tauri*), one glaucophyte (*Cyanophora paradoxa*), one brown alga (*Ectocarpus siliculosus*) and a pennate diatom (*Phaeodactylum tricornutum*). The database will be regularly updated and implemented with new plant genome annotations so as to provide extensive information on tRNA biology to the research community.

Mitochondrial Genome Evolution and the Emergence of PPR Proteins

Bernard Gutmann, Anthony Gobert and Philippe Giegé[1]

Institut de Biologie Moléculaire des Plantes du CNRS, University of Strasbourg, Strasbourg, France

[1]Corresponding author. E-mail: philippe.giege@ibmp-cnrs.unistra.fr

Abstract

The structure of mitochondrial genomes has greatly diverged throughout evolution. They can be very compact in metazoans, whereas they are much larger with considerably lower gene density in higher plants. These changes in structure have occurred in conjunction with the evolution of specific gene expression processes. For example, splicing and RNA editing are typical in plants but never found in animal mitochondria. Most of the organelle-specific gene expression processes rely on the function of nuclear-encoded proteins that do not originate from the organelle bacterial ancestor but have rather evolved during eukaryote history. It has become increasingly evident that pentatricopeptide repeat (PPR) proteins play a major role in mitochondrial gene expression processes. Similar to the mitochondrial genome structure, the number and diversity of PPR proteins have greatly varied during eukaryote evolution. There are, e.g. 80 times more PPR genes in higher plants than in animal genomes. This diversification correlates with the evolution of specific gene expression processes. For instance, the occurrence of PPR genes containing an additional DYW domain, which was proposed to support RNA editing activity, correlates precisely with the occurrence of RNA editing in the respective eukaryote phyla.

BIBLIOGRAPHIE

Abelson, J., Trotta, C.R., and Li, H. (1998). tRNA splicing. J Biol Chem **273,** 12685-12688.

Alifano, P., Rivellini, F., Piscitelli, C., Arraiano, C.M., Bruni, C.B., and Carlomagno, M.S. (1994). Ribonuclease E provides substrates for ribonuclease P-dependent processing of a polycistronic mRNA. Genes Dev **8,** 3021-3031.

Allen, J.F., and Raven, J.A. (1996). Free-radical-induced mutation vs redox regulation: costs and benefits of genes in organelles. J Mol Evol **42,** 482-492.

Allmang, C., Henry, Y., Morrissey, J.P., Wood, H., Petfalski, E., and Tollervey, D. (1996). Processing of the yeast pre-rRNA at sites A(2) and A(3) is linked. Rna **2,** 63-73.

Altman, S., Kirsebom, L., and Talbot, S. (1993). Recent studies of ribonuclease P. Faseb J **7,** 7-14.

Altman, S., Wesolowski, D., Guerrier-Takada, C., and Li, Y. (2005). RNase P cleaves transient structures in some riboswitches. Proc Natl Acad Sci U S A **102,** 11284-11289.

Anantharaman, V., and Aravind, L. (2006). The NYN domains: novel predicted RNAses with a PIN domain-like fold. RNA Biol **3,** 18-27.

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408,** 796-815.

Aubourg, S., Boudet, N., Kreis, M., and Lecharny, A. (2000). In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. Plant Mol Biol **42,** 603-613.

Banerjee, A.K. (1980). 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. Microbiol Rev **44,** 175-205.

Beattie, T.L., Olive, J.E., and Collins, R.A. (1995). A secondary-structure model for the self-cleaving region of *Neurospora* VS RNA. Proc Natl Acad Sci U S A **92,** 4686-4690.

Bell, P.J. (2001). Viral eukaryogenesis: was the ancestor of the nucleus a complex DNA virus? J Mol Evol **53,** 251-256.

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 A resolution. Science **334,** 1524-1529.

Benne, R. (1994). RNA editing in trypanosomes. Eur J Biochem **221,** 9-23.

Bennett, M.D., Leitch, I.J., Price, H.J., and Johnston, J.S. (2003). Comparisons with *Caenorhabditis* (approximately 100 Mb) and *Drosophila* (approximately 175 Mb) using flow cytometry show genome size in Arabidopsis to be approximately 157 Mb and thus approximately 25% larger than the Arabidopsis genome initiative estimate of approximately 125 Mb. Ann Bot **91,** 547-557.

Berr, A., Shafiq, S., and Shen, W.H. (2011). Histone modifications in transcriptional activation during plant development. Biochim Biophys Acta **1809,** 567-576.

Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays **21,** 932-939.

Bogenhagen, D.F., Sakonju, S., and Brown, D.D. (1980). A control region in the center of the 5S RNA gene directs specific initiation of transcription: II. The 3' border of the region. Cell **19,** 27-35.

Brannvall, M., and Kirsebom, L.A. (1999). Manganese ions induce miscleavage in the Escherichia coli RNase P RNA-catalyzed reaction. J Mol Biol **292,** 53-63.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., and Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. Science **320,** 1185-1190.

Brown, J.W. (1999). The Ribonuclease P Database. Nucleic Acids Res **27,** 314.

Brown, J.W., and Pace, N.R. (1992). Ribonuclease P RNA and protein subunits from bacteria. Nucleic Acids Res **20,** 1451-1456.

Brown, J.W., Haas, E.S., James, B.D., Hunt, D.A., Liu, J.S., and Pace, N.R. (1991). Phylogenetic analysis and evolution of RNase P RNA in proteobacteria. J Bacteriol **173,** 3855-3863.

Brusca, E.M., True, H.L., and Celander, D.W. (2001). Novel RNA-binding properties of Pop3p support a role for eukaryotic RNase P protein subunits in substrate recognition. J Biol Chem **276,** 42543-42548.

Burch-Smith, T.M., Schiff, M., Liu, Y., and Dinesh-Kumar, S.P. (2006). Efficient virus-induced gene silencing in Arabidopsis. Plant Physiol **142,** 21-27.

Castandet, B., Choury, D., Begu, D., Jordana, X., and Araya, A. (2010). Intron RNA editing is essential for splicing in plant mitochondria. Nucleic Acids Res **38,** 7112-7121.

Cech, T.R. (1990). Self-splicing of group I introns. Annu Rev Biochem **59,** 543-568.

Cech, T.R., Zaug, A.J., and Grabowski, P.J. (1981). In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. Cell **27,** 487-496.

Chamberlain, J.R., Pagan, R., Kindelberger, D.W., and Engelke, D.R. (1996). An RNase P RNA subunit mutation affects ribosomal RNA processing. Nucleic Acids Res **24,** 3158-3166.

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 Dev **12,** 1678-1690.

Chang, C.C., Sheen, J., Bligny, M., Niwa, Y., Lerbs-Mache, S., and Stern, D.B. (1999). Functional analysis of two maize cDNAs encoding T7-like RNA polymerases. Plant Cell **11,** 911-926.

Chateigner-Boutin, A.L., and Small, I. (2007). A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. Nucleic Acids Res **35,** e114.

Chaudhuri, S., Carrer, H., and Maliga, P. (1995). Site-specific factor involved in the editing of the psbL mRNA in tobacco plastids. EMBO J **14,** 2951-2957.

Chellappan, P., Xia, J., Zhou, X., Gao, S., Zhang, X., Coutino, G., Vazquez, F., Zhang, W., and Jin, H. (2010). siRNAs from miRNA sites mediate DNA methylation of target genes. Nucleic Acids Res **38,** 6883-6894.

Chen, J.X., and Yan, S.S. (2010). Role of mitochondrial amyloid-beta in Alzheimer's disease. J Alzheimers Dis **20 Suppl 2,** S569-578.

Chen, L.S., Du-Cuny, L., Vethantham, V., Hawke, D.H., Manley, J.L., Zhang, S., and Gandhi, V. (2010). Chain termination and inhibition of mammalian poly(A) polymerase by modified ATP analogues. Biochem Pharmacol **79,** 669-677.

Chen, Z.H., Bao, M.L., Sun, Y.Z., Yang, Y.J., Xu, X.H., Wang, J.H., Han, N., Bian, H.W., and Zhu, M.Y. (2011). Regulation of auxin response by miR393-targeted transport inhibitor response protein 1 is involved in normal development in *Arabidopsis*. Plant Mol Biol **77,** 619-629.

Cho, Y., Qiu, Y.L., Kuhlman, P., and Palmer, J.D. (1998). Explosive invasion of plant mitochondria by a group I intron. Proc Natl Acad Sci U S A **95,** 14244-14249.

Chu, S., Zengel, J.M., and Lindahl, L. (1997). A novel protein shared by RNase MRP and RNase P. RNA **3,** 382-391.

Cochrane, J.C., Lipchock, S.V., and Strobel, S.A. (2007). Structural investigation of the GlmS ribozyme bound to Its catalytic cofactor. Chem Biol **14,** 97-105.

Collins, J.A., Irnov, I., Baker, S., and Winkler, W.C. (2007). Mechanism of mRNA destabilization by the glmS ribozyme. Genes Dev **21,** 3356-3368.

Cooper, M.P., Qu, L., Rohas, L.M., Lin, J., Yang, W., Erdjument-Bromage, H., Tempst, P., and Spiegelman, B.M. (2006). Defects in energy homeostasis in Leigh syndrome French Canadian variant through PGC-1alpha/LRP130 complex. Genes Dev **20,** 2996-3009.

Copenhaver, G.P., and Pikaard, C.S. (1996). Two-dimensional RFLP analyses reveal megabase-sized clusters of rRNA gene variants in *Arabidopsis thaliana*, suggesting local spreading of variants as the mode for gene homogenization during concerted evolution. Plant J **9,** 273-282.

Coughlin, D.J., Pleiss, J.A., Walker, S.C., Whitworth, G.B., and Engelke, D.R. (2008). Genome-wide search for yeast RNase P substrates reveals role in maturation of intron-encoded box C/D small nucleolar RNAs. Proc Natl Acad Sci U S A **105,** 12218-12223.

Covello, P.S., and Gray, M.W. (1989). RNA editing in plant mitochondria. Nature **341,** 662-666.

Crary, S.M., Niranjanakumari, S., and Fierke, C.A. (1998). The protein component of *Bacillus subtilis* ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-$tRNA^{Asp}$. Biochemistry **37,** 9409-9416.

Dang, Y.L., and Martin, N.C. (1993). Yeast mitochondrial RNase P. Sequence of the RPM2 gene and demonstration that its product is a protein subunit of the enzyme. J Biol Chem **268,** 19791-19796.

Daoud, R., Forget, L., and Lang, B.F. (2012). Yeast mitochondrial RNase P, RNase Z and the RNA degradosome are part of a stable supercomplex. Nucleic Acids Res **40,** 1728-1736.

de Longevialle, A.F., Meyer, E.H., Andres, C., Taylor, N.L., Lurin, C., Millar, A.H., and Small, I.D. (2007). The pentatricopeptide repeat gene OTP43 is required for trans-splicing of the mitochondrial nad1 Intron 1 in *Arabidopsis thaliana*. Plant Cell **19,** 3256-3265.

Deppmann, C.D., Alvania, R.S., and Taparowsky, E.J. (2006). Cross-species annotation of basic leucine zipper factor interactions: Insight into the evolution of closed interaction networks. Mol Biol Evol **23,** 1480-1492.

Dichtl, B., and Tollervey, D. (1997). Pop3p is essential for the activity of the RNase MRP and RNase P ribonucleoproteins in vivo. EMBO J **16,** 417-429.

Doersen, C.J., Guerrier-Takada, C., Altman, S., and Attardi, G. (1985). Characterization of an RNase P activity from HeLa cell mitochondria. Comparison with the cytosol RNase P activity. J Biol Chem **260,** 5942-5949.

Doniwa, Y., Ueda, M., Ueta, M., Wada, A., Kadowaki, K., and Tsutsumi, N. (2010). The involvement of a PPR protein of the P subfamily in partial RNA editing of an Arabidopsis mitochondrial transcript. Gene **454,** 39-46.

Doudna, J.A. (1998). Ribozymes: the hammerhead swings into action. Curr Biol **8,** R495-497.

Drager, R.G., Higgs, D.C., Kindle, K.L., and Stern, D.B. (1999). 5' to 3' exoribonucleolytic activity is a normal component of chloroplast mRNA decay pathways. Plant J **19,** 521-531.

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*. Proc Natl Acad Sci U S A **102,** 16484-16489.

Edmonds, M., and Abrams, R. (1960). Polynucleotide biosynthesis: formation of a sequence of adenylate units from adenosine triphosphate by an enzyme from thymus nuclei. J Biol Chem **235,** 1142-1149.

Ellis, J.C., and Brown, J.W. (2009). The RNase P family. RNA Biol **6,** 362-369.

Fang, Y., and Spector, D.L. (2007). Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living *Arabidopsis* plants. Curr Biol **17,** 818-823.

Fasken, M.B., and Corbett, A.H. (2005). Process or perish: quality control in mRNA biogenesis. Nat Struct Mol Biol **12,** 482-488.

Fey, J., Weil, J.H., Tomita, K., Cosset, A., Dietrich, A., Small, I., and Marechal-Drouard, L. (2002). Role of editing in plant mitochondrial transfer RNAs. Gene **286,** 21-24.

Forner, J., Weber, B., Thuss, S., Wildum, S., and Binder, S. (2007). Mapping of mitochondrial mRNA termini in Arabidopsis thaliana: t-elements contribute to 5' and 3' end formation. Nucleic Acids Res **35,** 3676-3692.

Forterre, P. (2011). A new fusion hypothesis for the origin of *Eukarya*: better than previous ones, but probably also wrong. Res Microbiol **162,** 77-91.

Franklin, S.E., Zwick, M.G., and Johnson, J.D. (1995). Characterization and partial purification of two pre-tRNA 5'-processing activities from *Daucus carrota* (carrot) suspension cells. Plant J **7,** 553-563.

Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., Fritchman, R.D., Weidman, J.F., Small, K.V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T.R., Saudek, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.F., Dougherty, B.A., Bott, K.F., Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O., Hutchison, C.A., 3rd, and Venter, J.C. (1995). The minimal gene complement of *Mycoplasma genitalium.* Science **270,** 397-403.

Fuerst, J.A. (2005). Intracellular compartmentation in planctomycetes. Annu Rev Microbiol **59,** 299-328.

Galli, G., Hofstetter, H., and Birnstiel, M.L. (1981). Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. Nature **294,** 626-631.

Giege, P., and Brennicke, A. (1999). RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. Proc Natl Acad Sci U S A **96,** 15324-15329.

Gimple, O., and Schon, A. (2001). In vitro and in vivo processing of cyanelle tmRNA by RNase P. Biol Chem **382,** 1421-1429.

Gobert, A., Gutmann, B., Taschner, A., Gossringer, M., Holzmann, J., Hartmann, R.K., Rossmanith, W., and Giege, P. (2010). A single *Arabidopsis* organellar protein has RNase P activity. Nat Struct Mol Biol **17,** 740-744.

Goldfarb, K.C., Borah, S., and Cech, T.R. (2012). RNase P branches out from RNP to protein: organelle-triggered diversification? Genes Dev **26,** 1005-1009.

Gopalan, V., Baxevanis, A.D., Landsman, D., and Altman, S. (1997). Analysis of the functional role of conserved residues in the protein subunit of ribonuclease P from *Escherichia coli.* J Mol Biol **267,** 818-829.

Gray, M.W. (1999). Evolution of organellar genomes. Curr Opin Genet Dev **9,** 678-687.

Gualberto, J.M., Weil, J.H., and Grienenberger, J.M. (1990). Editing of the wheat coxIII transcript: evidence for twelve C to U and one U to C conversions and for sequence similarities around editing sites. Nucleic Acids Res **18,** 3771-3776.

Gualberto, J.M., Lamattina, L., Bonnard, G., Weil, J.H., and Grienenberger, J.M. (1989). RNA editing in wheat mitochondria results in the conservation of protein sequences. Nature **341,** 660-662.

Guerrier-Takada, C., McClain, W.H., and Altman, S. (1984). Cleavage of tRNA precursors by the RNA subunit of *E. coli* ribonuclease P (M1 RNA) is influenced by 3'-proximal CCA in the substrates. Cell **38,** 219-224.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell **35,** 849-857.

Gupta, R. (1984). *Halobacterium volcanii* tRNAs. Identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs. J Biol Chem **259,** 9461-9471.

Gutmann, B., Gobert, A., and Giege, P. (2012a). PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. Genes Dev **26,** 1022-1027.

Gutmann, B., Gobert, A. and Giegé, P. (2012b) Mitochondrial genome evolution and the emergence of PPR proteins. Advances in botanical research. Vol. Mitochondrial genome evolution, Drouard L. ed (in press)

Haas, E.S., and Brown, J.W. (1998). Evolutionary variation in bacterial RNase P RNAs. Nucleic Acids Res **26,** 4093-4099.

Hall, T.A., and Brown, J.W. (2002). Archaeal RNase P has multiple protein subunits homologous to eukaryotic nuclear RNase P proteins. RNA **8,** 296-306.

Hammani, K., Gobert, A., Hleibieh, K., Choulier, L., Small, I., and Giege, P. (2011). An *Arabidopsis* dual-localized pentatricopeptide repeat protein interacts with nuclear proteins involved in gene expression regulation. Plant Cell **23,** 730-740.

Hampel, A., and Cowan, J.A. (1997). A unique mechanism for RNA catalysis: the role of metal cofactors in hairpin ribozyme cleavage. Chem Biol **4,** 513-517.

Hardt, W.D., Warnecke, J.M., Erdmann, V.A., and Hartmann, R.K. (1995). Rp-phosphorothioate modifications in RNase P RNA that interfere with tRNA binding. EMBO J **14,** 2935-2944.

Harris, J.K., Haas, E.S., Williams, D., Frank, D.N., and Brown, J.W. (2001). New insight into RNase P RNA structure from comparative analysis of the archaeal RNA. RNA **7,** 220-232.

Harris, M.E., and Pace, N.R. (1995). Identification of phosphates involved in catalysis by the ribozyme RNase P RNA. RNA **1,** 210-218.

Hartmann, E., and Hartmann, R.K. (2003). The enigma of ribonuclease P evolution. Trends Genet **19,** 561-569.

Hartmann, R.K., Heinrich, J., Schlegl, J., and Schuster, H. (1995). Precursor of C4 antisense RNA of bacteriophages P1 and P7 is a substrate for RNase P of *Escherichia coli*. Proc Natl Acad Sci U S A **92,** 5822-5826.

Hedtke, B., Borner, T., and Weihe, A. (2000). One RNA polymerase serving two genomes. EMBO Rep **1,** 435-440.

Herrin, D.L., and Nickelsen, J. (2004). Chloroplast RNA processing and stability. Photosynth Res **82,** 301-314.

Hiesel, R., Wissinger, B., Schuster, W., and Brennicke, A. (1989). RNA editing in plant mitochondria. Science **246,** 1632-1634.

Holec, S., Lange, H., Kuhn, K., Alioua, M., Borner, T., and Gagliardi, D. (2006). Relaxed transcription in *Arabidopsis* mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. Mol Cell Biol **26,** 2869-2876.

Hollingsworth, M.J., and Martin, N.C. (1986). RNase P activity in the mitochondria of *Saccharomyces cerevisiae* depends on both mitochondrion and nucleus-encoded components. Mol Cell Biol **6,** 1058-1064.

Holzle, A., Jonietz, C., Torjek, O., Altmann, T., Binder, S., and Forner, J. (2011). A RESTORER OF FERTILITY-like PPR gene is required for 5'-end processing of the nad4 mRNA in mitochondria of *Arabidopsis thaliana*. Plant J **65,** 737-744.

Holzmann, J., and Rossmanith, W. (2009). tRNA recognition, processing, and disease: hypotheses around an unorthodox type of RNase P in human mitochondria. Mitochondrion **9,** 284-288.

Holzmann, J., Frank, P., Loffler, E., Bennett, K.L., Gerner, C., and Rossmanith, W. (2008). RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell **135,** 462-474.

Hopper, A.K., and Phizicky, E.M. (2003). tRNA transfers to the limelight. Genes Dev **17,** 162-180.

Hori, Y., Tanaka, T., and Kikuchi, Y. (2000a). The catalytic RNA of RNase P from *Escherichia coli* cleaves *Drosophila* 2S ribosomal RNA in vitro: a new type of naturally occurring substrate for the ribozyme. FEBS Lett **472,** 187-190.

Hori, Y., Tanaka, T., and Kikuchi, Y. (2000b). In vitro cleavage of *Drosophila* 2S rRNA by M1 RNA. Nucleic Acids Symp Ser**,** 93-94.

Houser-Scott, F., Xiao, S., Millikin, C.E., Zengel, J.M., Lindahl, L., and Engelke, D.R. (2002). Interactions among the protein and RNA subunits of *Saccharomyces cerevisiae* nuclear RNase P. Proc Natl Acad Sci U S A **99,** 2684-2689.

Huang, Y., and Carmichael, G.G. (1996). Role of polyadenylation in nucleocytoplasmic transport of mRNA. Mol Cell Biol **16,** 1534-1542.

Ichikawa, K., Shimizu, A., Okada, R., Satbhai, S.B., and Aoki, S. (2008). The plastid sigma factor SIG5 is involved in the diurnal regulation of the chloroplast gene psbD in the moss *Physcomitrella patens.* FEBS Lett **582,** 405-409.

Ichinose, M., Tasaki, E., Sugita, C., and Sugita, M. (2012). A PPR-DYW protein is required for splicing of a group II intron of cox1 pre-mRNA in *Physcomitrella patens.* Plant J **70,** 271-278.

Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., Kobori, M., Takeba, G., Nakahira, Y., and Shiina, T. (2005). A nuclear-encoded sigma factor, *Arabidopsis* SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. Plant J **42,** 133-144.

Isono, K., Shimizu, M., Yoshimoto, K., Niwa, Y., Satoh, K., Yokota, A., and Kobayashi, H. (1997a). Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of sigma70 factors of bacterial RNA polymerases in *Arabidopsis thaliana.* Proc Natl Acad Sci U S A **94,** 14948-14953.

Isono, M., Wakabayashi, Y., Fujiki, M., Inoue, R., Asakuno, K., and Hori, S. (1997b). Evidence of apoptotic procedure in deafferented striatum after cortical injury in young adult rats. Brain Res **765,** 319-323.

Iyer, N.J., Jia, X., Sunkar, R., Tang, G., and Mahalingam, R. (2012). microRNAs responsive to ozone-induced oxidative stress in *Arabidopsis thaliana*. Plant Signal Behav **7**.

Jackman, J.E., Montange, R.K., Malik, H.S., and Phizicky, E.M. (2003). Identification of the yeast gene encoding the tRNA m1G methyltransferase responsible for modification at position 9. RNA **9,** 574-585.

Jarrous, N., and Altman, S. (2001). Human ribonuclease P. Methods Enzymol **342,** 93-100.

Jarrous, N., and Reiner, R. (2007). Human RNase P: a tRNA-processing enzyme and transcription factor. Nucleic Acids Res **35,** 3519-3524.

Jonietz, C., Forner, J., Hildebrandt, T., and Binder, S. (2011). RNA PROCESSING FACTOR3 is crucial for the accumulation of mature ccmC transcripts in mitochondria of *Arabidopsis* accession Columbia. Plant Physiol **157,** 1430-1439.

Jonietz, C., Forner, J., Holzle, A., Thuss, S., and Binder, S. (2010). RNA PROCESSING FACTOR2 is required for 5' end processing of nad9 and cox3 mRNAs in mitochondria of *Arabidopsis thaliana.* Plant Cell **22,** 443-453.

Joyce, G.F. (1989). RNA evolution and the origins of life. Nature **338,** 217-224.

Kachouri, R., Stribinskis, V., Zhu, Y., Ramos, K.S., Westhof, E., and Li, Y. (2005). A surprisingly large RNase P RNA in *Candida glabrata.* RNA **11,** 1064-1072.

Kasai, K., Kawagishi-Kobayashi, M., Teraishi, M., Ito, Y., Ochi, K., Wakasa, K., and Tozawa, Y. (2004). Differential expression of three plastidial sigma factors, OsSIG1, OsSIG2A, and OsSIG2B, during leaf development in rice. Biosci Biotechnol Biochem **68,** 973-977.

Kazantsev, A.V., and Pace, N.R. (2006). Bacterial RNase P: a new view of an ancient enzyme. Nat Rev Microbiol **4,** 729-740.

Keene, J.D., Komisarow, J.M., and Friedersdorf, M.B. (2006). RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. Nat Protoc **1,** 302-307.

Kikovska, E., Svard, S.G., and Kirsebom, L.A. (2007). Eukaryotic RNase P RNA mediates cleavage in the absence of protein. Proc Natl Acad Sci U S A **104,** 2062-2067.

Kim, S., Yang, J.Y., Xu, J., Jang, I.C., Prigge, M.J., and Chua, N.H. (2008). Two cap-binding proteins CBP20 and CBP80 are involved in processing primary MicroRNAs. Plant Cell Physiol **49,** 1634-1644.

Kirsebom, L.A., and Vioque, A. (1995). RNase P from bacteria. Substrate recognition and function of the protein subunit. Mol Biol Rep **22,** 99-109.

Kiss, T., Marshallsay, C., and Filipowicz, W. (1992). 7-2/MRP RNAs in plant and mammalian cells: association with higher order structures in the nucleolus. EMBO J **11,** 3737-3746.

Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K., and Inokuchi, H. (1994). A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli.* Proc Natl Acad Sci U S A **91,** 9223-9227.

Kotera, E., Tasaka, M., and Shikanai, T. (2005). A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature **433,** 326-330.

Kouzuma, Y., Mizoguchi, M., Takagi, H., Fukuhara, H., Tsukamoto, M., Numata, T., and Kimura, M. (2003). Reconstitution of archaeal ribonuclease P from RNA and four protein components. Biochem Biophys Res Commun **306,** 666-673.

Krehan, M., Heubeck, C., Menzel, N., Seibel, P., and Schon, A. (2012). RNase MRP RNA and RNase P activity in plants are associated with a Pop1p containing complex. Nucleic Acids Res. (in press)

Kruszka, K., Barneche, F., Guyot, R., Ailhas, J., Meneau, I., Schiffer, S., Marchfelder, A., and Echeverria, M. (2003). Plant dicistronic tRNA-snoRNA genes: a new mode of expression of the small nucleolar RNAs processed by RNase Z. EMBO J **22,** 621-632.

Kufel, J., and Kirsebom, L.A. (1998). The P15-loop of Escherichia coli RNase P RNA is an autonomous divalent metal ion binding domain. RNA **4,** 777-788.

Kuhn, K., Bohne, A.V., Liere, K., Weihe, A., and Borner, T. (2007). Arabidopsis phage-type RNA polymerases: accurate in vitro transcription of organellar genes. Plant Cell **19,** 959-971.

Kuhn, K., Richter, U., Meyer, E.H., Delannoy, E., de Longevialle, A.F., O'Toole, N., Borner, T., Millar, A.H., Small, I.D., and Whelan, J. (2009). Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of *Arabidopsis thaliana*. Plant Cell **21,** 2762-2779.

Kuhn, U., and Wahle, E. (2004). Structure and function of poly(A) binding proteins. Biochim Biophys Acta **1678,** 67-84.

Kuiper, M.T., and Lambowitz, A.M. (1988). A novel reverse transcriptase activity associated with mitochondrial plasmids of *Neurospora*. Cell **55,** 693-704.

Kunzmann, A., Brennicke, A., and Marchfelder, A. (1998). 5' end maturation and RNA editing have to precede tRNA 3' processing in plant mitochondria. Proc Natl Acad Sci U S A **95,** 108-113.

Kurihara, Y., and Watanabe, Y. (2004). Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. Proc Natl Acad Sci U S A **101,** 12753-12758.

Kurihara, Y., Takashi, Y., and Watanabe, Y. (2006). The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA **12,** 206-212.

Kurz, J.C., Niranjanakumari, S., and Fierke, C.A. (1998). Protein component of *Bacillus subtilis* RNase P specifically enhances the affinity for precursor-$tRNA^{Asp}$. Biochemistry **37,** 2393-2400.

Lai, L.B., Vioque, A., Kirsebom, L.A., and Gopalan, V. (2010a). Unexpected diversity of RNase P, an ancient tRNA processing enzyme: challenges and prospects. FEBS Lett **584,** 287-296.

Lai, L.B., Chan, P.P., Cozen, A.E., Bernick, D.L., Brown, J.W., Gopalan, V., and Lowe, T.M. (2010b). Discovery of a minimal form of RNase P in *Pyrobaculum*. Proc Natl Acad Sci U S A **107,** 22493-22498.

Lang, B.F., Burger, G., O'Kelly, C.J., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., and Gray, M.W. (1997). An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature **387,** 493-497.

Lange, H., Sement, F.M., and Gagliardi, D. (2011). MTR4, a putative RNA helicase and exosome co-factor, is required for proper rRNA biogenesis and development in *Arabidopsis thaliana*. Plant J **68,** 51-63.

Lee, B.J., de la Pena, P., Tobian, J.A., Zasloff, M., and Hatfield, D. (1987). Unique pathway of expression of an opal suppressor phosphoserine tRNA. Proc Natl Acad Sci U S A **84,** 6384-6388.

Lee, J.Y., Rohlman, C.E., Molony, L.A., and Engelke, D.R. (1991). Characterization of RPR1, an essential gene encoding the RNA component of *Saccharomyces cerevisiae* nuclear RNase P. Mol Cell Biol **11,** 721-730.

Lee, T.S., Silva Lopez, C., Giambasu, G.M., Martick, M., Scott, W.G., and York, D.M. (2008). Role of Mg^{2+} in hammerhead ribozyme catalysis from molecular simulation. J Am Chem Soc **130,** 3053-3064.

Lee, Y.C., Lee, B.J., Hwang, D.S., and Kang, H.S. (1996). Purification and characterization of mitochondrial ribonuclease P from *Aspergillus nidulans*. Eur J Biochem **235,** 289-296.

Li, Y., and Altman, S. (2003). A specific endoribonuclease, RNase P, affects gene expression of polycistronic operon mRNAs. Proc Natl Acad Sci U S A **100,** 13213-13218.

Li, Z., Pandit, S., and Deutscher, M.P. (1998). 3' exoribonucleolytic trimming is a common feature of the maturation of small, stable RNAs in *Escherichia coli*. Proc Natl Acad Sci U S A **95,** 2856-2861.

Liere, K., Weihe, A., and Borner, T. (2011). The transcription machineries of plant mitochondria and chloroplasts: Composition, function, and regulation. J Plant Physiol **168,** 1345-1360.

Lilley, D.M. (2004). The Varkud satellite ribozyme. RNA **10,** 151-158.

Lin, X., Kaul, S., Rounsley, S., Shea, T.P., Benito, M.I., Town, C.D., Fujii, C.Y., Mason, T., Bowman, C.L., Barnstead, M., Feldblyum, T.V., Buell, C.R., Ketchum, K.A., Lee, J., Ronning, C.M., Koo, H.L., Moffat, K.S., Cronin, L.A., Shen, M., Pai, G., Van Aken, S., Umayam, L., Tallon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H., Goodman, H.M., Somerville, C.R., Copenhaver, G.P., Preuss, D., Nierman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M., and Venter, J.C. (1999). Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana. Nature **402,** 761-768.

Lipinski, K.A., Puchta, O., Surendranath, V., Kudla, M., and Golik, P. (2011). Revisiting the yeast PPR proteins--application of an Iterative Hidden Markov Model algorithm reveals new members of the rapidly evolving family. Mol Biol Evol **28,** 2935-2948.

Logan, D.C., and Knight, M.R. (2003). Mitochondrial and cytosolic calcium dynamics are differentially regulated in plants. Plant Physiol **133,** 21-24.

Lopez-Garcia, P., and Moreira, D. (2006). Selective forces for the origin of the eukaryotic nucleus. Bioessays **28,** 525-533.

Loria, A., and Pan, T. (2000). The 3' substrate determinants for the catalytic efficiency of the Bacillus subtilis RNase P holoenzyme suggest autolytic processing of the RNase P RNA in vivo. RNA **6,** 1413-1422.

Lorkovic, Z.J., Wieczorek Kirk, D.A., Lambermon, M.H., and Filipowicz, W. (2000). Pre-mRNA splicing in higher plants. Trends Plant Sci **5,** 160-167.

Lundberg, U., and Altman, S. (1995). Processing of the precursor to the catalytic RNA subunit of RNase P from Escherichia coli. RNA **1,** 327-334.

Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., Lecharny, A., Le Ret, M., Martin-Magniette, M.L., Mireau, H., Peeters, N., Renou, J.P., Szurek, B., Taconnat, L., and Small, I. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell **16,** 2089-2103.

Lygerou, Z., Mitchell, P., Petfalski, E., Seraphin, B., and Tollervey, D. (1994). The POP1 gene encodes a protein component common to the RNase MRP and RNase P ribonucleoproteins. Genes Dev **8,** 1423-1433.

Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., Watkins, K.P., Barkan, A., and van Wijk, K.J. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. Plant Physiol **158,** 156-189.

Marchfelder, A. (1995). Plant mitochondrial RNase P. Mol Biol Rep **22,** 151-156.

Marechal-Drouard, L., Kumar, R., Remacle, C., and Small, I. (1996). RNA editing of larch mitochondrial $tRNA^{His}$ precursors is a prerequisite for processing. Nucleic Acids Res **24,** 3229-3234.

Marin-Navarro, J., Manuell, A.L., Wu, J., and P. Mayfield, S. (2007). Chloroplast translation regulation. Photosynth Res **94,** 359-374.

Marquez, S.M., Chen, J.L., Evans, D., and Pace, N.R. (2006). Structure and function of eukaryotic Ribonuclease P RNA. Mol Cell **24,** 445-456.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D. (2002). Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci U S A **99,** 12246-12251.

McClain, W.H., Guerrier-Takada, C., and Altman, S. (1987). Model substrates for an RNA enzyme. Science **238,** 527-530.

Michaud, M., Cognat, V., Duchene, A.M., and Marechal-Drouard, L. (2011). A global picture of tRNA genes in plant genomes. Plant J **66,** 80-93.

Mingler, M.K., Hingst, A.M., Clement, S.L., Yu, L.E., Reifur, L., and Koslowsky, D.J. (2006). Identification of pentatricopeptide repeat proteins in *Trypanosoma brucei*. Mol Biochem Parasitol **150,** 37-45.

Mitchell, S.F., Walker, S.E., Algire, M.A., Park, E.H., Hinnebusch, A.G., and Lorsch, J.R. (2010). The 5'-7-methylguanosine cap on eukaryotic mRNAs serves both to stimulate canonical translation initiation and to block an alternative pathway. Mol Cell **39,** 950-962.

Moore, M.J., and Proudfoot, N.J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. Cell **136,** 688-700.

Morales, M.J., Dang, Y.L., Lou, Y.C., Sulo, P., and Martin, N.C. (1992). A 105-kDa protein is required for yeast mitochondrial RNase P activity. Proc Natl Acad Sci U S A **89,** 9875-9879.

Moss, T., Langlois, F., Gagnon-Kugler, T., and Stefanovsky, V. (2007). A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. Cell Mol Life Sci **64,** 29-49.

Nagashima, A., Hanaoka, M., Shikanai, T., Fujiwara, M., Kanamaru, K., Takahashi, H., and Tanaka, K. (2004a). The multiple-stress responsive plastid sigma factor, SIG5, directs activation of the psbD blue light-responsive promoter (BLRP) in *Arabidopsis thaliana.* Plant Cell Physiol **45,** 357-368.

Nagashima, A., Hanaoka, M., Motohashi, R., Seki, M., Shinozaki, K., Kanamaru, K., Takahashi, H., and Tanaka, K. (2004b). DNA microarray analysis of plastid gene expression in an *Arabidopsis* mutant deficient in a plastid transcription factor sigma, SIG2. Biosci Biotechnol Biochem **68,** 694-704.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell **2,** 279-289.

Nesbitt, S., Hegg, L.A., and Fedor, M.J. (1997). An unusual pH-independent and metal-ion-independent mechanism for hairpin ribozyme catalysis. Chem Biol **4,** 619-630.

Niranjanakumari, S., Stams, T., Crary, S.M., Christianson, D.W., and Fierke, C.A. (1998). Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA. Proc Natl Acad Sci U S A **95,** 15212-15217.

Nojima, T., Hirose, T., Kimura, H., and Hagiwara, M. (2007). The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. J Biol Chem **282,** 15645-15651.

O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C., Sugita, M., and Small, I. (2008). On the expansion of the pentatricopeptide repeat gene family in plants. Mol Biol Evol **25,** 1120-1128.

Ogle, J.M., Brodersen, D.E., Clemons, W.M., Jr., Tarry, M.J., Carter, A.P., and Ramakrishnan, V. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. Science **292,** 897-902.

Ohta, N., Matsuzaki, M., Misumi, O., Miyagishima, S.Y., Nozaki, H., Tanaka, K., Shin, I.T., Kohara, Y., and Kuroiwa, T. (2003). Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae.* DNA Res **10,** 67-77.

Okuda, K., Nakamura, T., Sugita, M., Shimizu, T., and Shikanai, T. (2006). A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. J Biol Chem **281,** 37661-37667.

Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K., and Shikanai, T. (2007). Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. Proc Natl Acad Sci U S A **104,** 8178-8183.

Pagan-Ramos, E., Lee, Y., and Engelke, D.R. (1996a). Mutational analysis of *Saccharomyces cerevisiae* nuclear RNase P: randomization of universally conserved positions in the RNA subunit. RNA **2,** 441-451.

Pagan-Ramos, E., Lee, Y., and Engelke, D.R. (1996b). A conserved RNA motif involved in divalent cation utilization by nuclear RNase P. RNA **2,** 1100-1109.

Panagiotidis, C.A., Drainas, D., and Huang, S.C. (1992). Modulation of ribonuclease P expression in *Escherichia coli* by polyamines. Int J Biochem **24,** 1625-1631.

Pannucci, J.A., Haas, E.S., Hall, T.A., Harris, J.K., and Brown, J.W. (1999). RNase P RNAs from some *Archaea* are catalytically active. Proc Natl Acad Sci U S A **96,** 7803-7808.

Peck-Miller, K.A., and Altman, S. (1991). Kinetics of the processing of the precursor to 4.5 S RNA, a naturally occurring substrate for RNase P from *Escherichia coli.* J Mol Biol **221,** 1-5.

Perreault, J., Weinberg, Z., Roth, A., Popescu, O., Chartrand, P., Ferbeyre, G., and Breaker, R.R. (2011). Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. PLoS Comput Biol **7,** e1002031.

Pfalz, J., Bayraktar, O.A., Prikryl, J., and Barkan, A. (2009). Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. EMBO J **28,** 2042-2052.

Prikryl, J., Rojas, M., Schuster, G., and Barkan, A. (2011). Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. Proc Natl Acad Sci U S A **108,** 415-420.

Pring, D., Brennicke, A., and Schuster, W. (1993). RNA editing gives a new meaning to the genetic information in mitochondria and chloroplasts. Plant Mol Biol **21,** 1163-1170.

Puranam, R.S., and Attardi, G. (2001). The RNase P associated with HeLa cell mitochondria contains an essential RNA component identical in sequence to that of the nuclear RNase P. Mol Cell Biol **21,** 548-561.

Pusnik, M., Small, I., Read, L.K., Fabbro, T., and Schneider, A. (2007). Pentatricopeptide repeat proteins in *Trypanosoma brucei* function in mitochondrial ribosomes. Mol Cell Biol **27,** 6876-6888.

Qin, D., Liu, Q., Devaraj, A., and Fredrick, K. (2012). Role of helix 44 of 16S rRNA in the fidelity of translation initiation. RNA **18,** 485-495.

Rackham, O., and Filipovska, A. (2011). The role of mammalian PPR domain proteins in the regulation of mitochondrial gene expression. Biochim Biophys Acta.

Randau, L., Schroder, I., and Soll, D. (2008). Life without RNase P. Nature **453,** 120-123.

Raven, J.A., and Allen, J.F. (2003). Genomics and chloroplast evolution: what did cyanobacteria do for plants? Genome Biol **4,** 209.

Reilly, T.H., and Schmitt, M.E. (1995). The yeast, *Saccharomyces cerevisiae*, RNase P/MRP ribonucleoprotein endoribonuclease family. Mol Biol Rep **22,** 87-93.

Reiter, N.J., Osterman, A., Torres-Larios, A., Swinger, K.K., Pan, T., and Mondragon, A. (2010). Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. Nature **468,** 784-789.

Rivals, E., Bruyere, C., Toffano-Nioche, C., and Lecharny, A. (2006). Formation of the *Arabidopsis* pentatricopeptide repeat family. Plant Physiol **141,** 825-839.

Rodnina, M.V., Beringer, M., and Wintermeyer, W. (2007). How ribosomes make peptide bonds. Trends Biochem Sci **32,** 20-26.

Ronemus, M., Vaughn, M.W., and Martienssen, R.A. (2006). MicroRNA-targeted and small interfering RNA-mediated mRNA degradation is regulated by argonaute, dicer, and RNA-dependent RNA polymerase in *Arabidopsis*. Plant Cell **18,** 1559-1574.

Rossmanith, W. (2012). Of P and Z: Mitochondrial tRNA processing enzymes. Biochim Biophys Acta, in press. Published online November 23, 2011. 10.1016/j.bbagrm.2011.11.003

Rossmanith, W., and Karwan, R.M. (1998). Characterization of human mitochondrial RNase P: novel aspects in tRNA processing. Biochem Biophys Res Commun **247,** 234-241.

Rupert, P.B., and Ferre-D'Amare, A.R. (2001). Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. Nature **410,** 780-786.

Sakonju, S., Bogenhagen, D.F., and Brown, D.D. (1980). A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5' border of the region. Cell **19,** 13-25.

Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., Billault, A., Brottier, P., Camus, J.C., Cattolico, L., Chandler, M., Choisne, N., Claudel-Renard, C., Cunnac, S., Demange, N., Gaspin, C.,

Lavie, M., Moisan, A., Robert, C., Saurin, W., Schiex, T., Siguier, P., Thebault, P., Whalen, M., Wincker, P., Levy, M., Weissenbach, J., and Boucher, C.A. (2002). Genome sequence of the plant pathogen *Ralstonia solanacearum*. Nature **415,** 497-502.

Salinas, K., Wierzbicki, S., Zhou, L., and Schmitt, M.E. (2005). Characterization and purification of *Saccharomyces cerevisiae* RNase MRP reveals a new unique protein component. J Biol Chem **280,** 11352-11360.

Salinas, T., Duchene, A.M., and Marechal-Drouard, L. (2008). Recent advances in tRNA mitochondrial import. Trends Biochem Sci **33,** 320-329.

Salone, V., Rudinger, M., Polsakiewicz, M., Hoffmann, B., Groth-Malonek, M., Szurek, B., Small, I., Knoop, V., and Lurin, C. (2007). A hypothesis on the identification of the editing enzyme in plant organelles. FEBS Lett **581,** 4132-4138.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., and Tabata, S. (1999). Complete structure of the chloroplast genome of *Arabidopsis thaliana.* DNA Res **6,** 283-290.

Schedl, P., and Primakoff, P. (1973). Mutants of *Escherichia coli* thermosensitive for the synthesis of transfer RNA. Proc Natl Acad Sci U S A **70,** 2091-2095.

Schein, A., Sheffy-Levin, S., Glaser, F., and Schuster, G. (2008). The RNase E/G-type endoribonuclease of higher plants is located in the chloroplast and cleaves RNA similarly to the *E. coli* enzyme. Rna **14,** 1057-1068.

Schmid, M., and Jensen, T.H. (2008). The exosome: a multipurpose RNA-decay machine. Trends Biochem Sci **33,** 501-510.

Schmitt, M.E., and Clayton, D.A. (1994). Characterization of a unique protein component of yeast RNase MRP: an RNA-binding protein with a zinc-cluster domain. Genes Dev **8,** 2617-2628.

Schmitz-Linneweber, C., Williams-Carrier, R.E., Williams-Voelker, P.M., Kroeger, T.S., Vichas, A., and Barkan, A. (2006). A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 pre-mRNA. Plant Cell **18,** 2650-2663.

Schon, A. (1995). Ribonuclease P from plant nuclei and photosynthetic organelles. Mol Biol Rep **22,** 139-145.

Schuster, G., and Stern, D. (2009). RNA polyadenylation and decay in mitochondria and chloroplasts. Prog Mol Biol Transl Sci **85,** 393-422.

Scott, I., and Logan, D.C. (2008). Mitochondria and cell death pathways in plants: Actions speak louder than words. Plant Signal Behav **3,** 475-477.

Scott, W.G., Finch, J.T., and Klug, A. (1995). The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. Cell **81,** 991-1002.

Seif, E.R., Forget, L., Martin, N.C., and Lang, B.F. (2003). Mitochondrial RNase P RNAs in ascomycete fungi: lineage-specific variations in RNA secondary structure. RNA **9,** 1073-1083.

Seo, P.J., Kim, S.G., and Park, C.M. (2008). Membrane-bound transcription factors in plants. Trends Plant Sci **13,** 550-556.

Sharp, S., DeFranco, D., Dingermann, T., Farrell, P., and Soll, D. (1981). Internal control regions for transcription of eukaryotic tRNA genes. Proc Natl Acad Sci U S A **78,** 6657-6661.

Sharwood, R.E., Halpert, M., Luro, S., Schuster, G., and Stern, D.B. (2011). Chloroplast RNase J compensates for inefficient transcription termination by removal of antisense RNA. RNA **17,** 2165-2176.

Shevelev, E.L., Bryant, D.A., Loffelhardt, W., and Bohnert, H.J. (1995). Ribonuclease-P RNA gene of the plastid chromosome from *Cyanophora paradoxa*. DNA Res **2,** 231-234.

Shimotohno, K., Kodama, Y., Hashimoto, J., and Miura, K.I. (1977). Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis. Proc Natl Acad Sci U S A **74,** 2734-2738.

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

Simillion, C., Vandepoele, K., Van Montagu, M.C., Zabeau, M., and Van de Peer, Y. (2002). The hidden duplication past of *Arabidopsis thaliana*. Proc Natl Acad Sci U S A **99,** 13627-13632.

Small, I.D., and Peeters, N. (2000). The PPR motif - a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci **25,** 46-47.

Smith, D., Burgin, A.B., Haas, E.S., and Pace, N.R. (1992). Influence of metal ions on the ribonuclease P reaction. Distinguishing substrate binding from catalysis. J Biol Chem **267,** 2429-2436.

Srisawat, C., Houser-Scott, F., Bertrand, E., Xiao, S., Singer, R.H., and Engelke, D.R. (2002). An active precursor in assembly of yeast nuclear ribonuclease P. RNA **8,** 1348-1360.

Stams, T., Niranjanakumari, S., Fierke, C.A., and Christianson, D.W. (1998). Ribonuclease P protein structure: evolutionary origins in the translational apparatus. Science **280,** 752-755.

Steitz, T.A., and Steitz, J.A. (1993). A general two-metal-ion mechanism for catalytic RNA. Proc Natl Acad Sci U S A **90,** 6498-6502.

Stolc, V., and Altman, S. (1997). Rpp1, an essential protein subunit of nuclear RNase P required for processing of precursor tRNA and 35S precursor rRNA in *Saccharomyces cerevisiae*. Genes Dev **11,** 2926-2937.

Stolc, V., Katz, A., and Altman, S. (1998). Rpp2, an essential protein subunit of nuclear RNase P, is required for processing of precursor tRNAs and 35S precursor rRNA in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A **95,** 6716-6721.

Stribinskis, V., and Ramos, K.S. (2007). Rpm2p, a protein subunit of mitochondrial RNase P, physically and genetically interacts with cytoplasmic processing bodies. Nucleic Acids Res **35,** 1301-1311.

Stribinskis, V., Gao, G.J., Sulo, P., Dang, Y.L., and Martin, N.C. (1996). Yeast mitochondrial RNase P RNA synthesis is altered in an RNase P protein subunit mutant: insights into the biogenesis of a mitochondrial RNA-processing enzyme. Mol Cell Biol **16,** 3429-3436.

Stribinskis, V., Gao, G.J., Sulo, P., Ellis, S.R., and Martin, N.C. (2001). Rpm2p: separate domains promote tRNA and Rpm1r maturation in *Saccharomyces cerevisiae* mitochondria. Nucleic Acids Res **29,** 3631-3637.

Stribinskis, V., Heyman, H.C., Ellis, S.R., Steffen, M.C., and Martin, N.C. (2005). Rpm2p, a component of yeast mitochondrial RNase P, acts as a transcriptional activator in the nucleus. Mol Cell Biol **25,** 6546-6558.

Symons, R.H. (1997). Plant pathogenic RNAs and RNA catalysis. Nucleic Acids Res **25,** 2683-2689.

Szarzynska, B., Sobkowiak, L., Pant, B.D., Balazadeh, S., Scheible, W.R., Mueller-Roeber, B., Jarmolowski, A., and Szweykowska-Kulinska, Z. (2009). Gene structures and processing of *Arabidopsis thaliana* HYL1-dependent pri-miRNAs. Nucleic Acids Res **37,** 3083-3093.

Takenaka, M., and Brennicke, A. (2003). In vitro RNA editing in pea mitochondria requires NTP or dNTP, suggesting involvement of an RNA helicase. J Biol Chem **278,** 47526-47533.

Tallsjo, A., and Kirsebom, L.A. (1993). Product release is a rate-limiting step during cleavage by the catalytic RNA subunit of *Escherichia coli* RNase P. Nucleic Acids Res **21,** 51-57.

Theologis, A., Ecker, J.R., Palm, C.J., Federspiel, N.A., Kaul, S., White, O., Alonso, J., Altafi, H., Araujo, R., Bowman, C.L., Brooks, S.Y., Buehler, E., Chan, A., Chao, Q., Chen, H., Cheuk, R.F., Chin, C.W., Chung, M.K., Conn, L., Conway, A.B., Conway, A.R., Creasy, T.H., Dewar, K., Dunn, P., Etgu, P., Feldblyum,

T.V., Feng, J., Fong, B., Fujii, C.Y., Gill, J.E., Goldsmith, A.D., Haas, B., Hansen, N.F., Hughes, B., Huizar, L., Hunter, J.L., Jenkins, J., Johnson-Hopson, C., Khan, S., Khaykin, E., Kim, C.J., Koo, H.L., Kremenetskaia, I., Kurtz, D.B., Kwan, A., Lam, B., Langin-Hooper, S., Lee, A., Lee, J.M., Lenz, C.A., Li, J.H., Li, Y., Lin, X., Liu, S.X., Liu, Z.A., Luros, J.S., Maiti, R., Marziali, A., Militscher, J., Miranda, M., Nguyen, M., Nierman, W.C., Osborne, B.I., Pai, G., Peterson, J., Pham, P.K., Rizzo, M., Rooney, T., Rowley, D., Sakano, H., Salzberg, S.L., Schwartz, J.R., Shinn, P., Southwick, A.M., Sun, H., Tallon, L.J., Tambunga, G., Toriumi, M.J., Town, C.D., Utterback, T., Van Aken, S., Vaysberg, M., Vysotskaia, V.S., Walker, M., Wu, D., Yu, G., Fraser, C.M., Venter, J.C., and Davis, R.W. (2000). Sequence and analysis of chromosome 1 of the plant *Arabidopsis thaliana.* Nature **408,** 816-820.

Thomas, B.C., Li, X., and Gegenheimer, P. (2000). Chloroplast ribonuclease P does not utilize the ribozyme-type pre-tRNA cleavage mechanism. RNA **6,** 545-553.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y., and Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet **5,** 123-135.

Tous, C., Vega-Palas, M.A., and Vioque, A. (2001). Conditional expression of RNase P in the cyanobacterium *Synechocystis* sp. PCC6803 allows detection of precursor RNAs. Insight in the in vivo maturation pathway of transfer and other stable RNAs. J Biol Chem **276,** 29059-29066.

Tracy, R.L., and Stern, D.B. (1995). Mitochondrial transcription initiation: promoter structures and RNA polymerases. Curr Genet **28,** 205-216.

Turmel, M., Otis, C., and Lemieux, C. (1999). The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes. Proc Natl Acad Sci U S A **96,** 10248-10253.

Unseld, M., Marienfeld, J.R., Brandt, P., and Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nat Genet **15,** 57-61.

Vaillant, I., Tutois, S., Cuvillier, C., Schubert, I., and Tourmente, S. (2007). Regulation of *Arabidopsis thaliana* 5S rRNA Genes. Plant Cell Physiol **48,** 745-752.

Vaucheret, H., Vazquez, F., Crete, P., and Bartel, D.P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev **18,** 1187-1197.

Walter, M., Kilian, J., and Kudla, J. (2002). PNPase activity determines the efficiency of mRNA 3'-end processing, the degradation of tRNA and the extent of polyadenylation in chloroplasts. EMBO J **21,** 6905-6914.

Wang, M.J., Davis, N.W., and Gegenheimer, P. (1988). Novel mechanisms for maturation of chloroplast transfer RNA precursors. EMBO J **7,** 1567-1574.

Wang, Z., Zou, Y., Li, X., Zhang, Q., Chen, L., Wu, H., Su, D., Chen, Y., Guo, J., Luo, D., Long, Y., Zhong, Y., and Liu, Y.G. (2006). Cytoplasmic male sterility of rice with boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. Plant Cell **18,** 676-687.

Waugh, D.S., Green, C.J., and Pace, N.R. (1989). The design and catalytic properties of a simplified ribonuclease P RNA. Science **244,** 1569-1571.

Weihe, A. (2004). The transcription of plant organelle genomes. In Molecular Biology and Biotechnology of Plant Organelles. Chloroplasts and Mitochondria, Daniell H., Chase C. eds (Dordrecht: Springer), pp. 213-237.

Williams-Carrier, R., Kroeger, T., and Barkan, A. (2008). Sequence-specific binding of a chloroplast pentatricopeptide repeat protein to its native group II intron ligand. RNA **14,** 1930-1941.

Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A., and Breaker, R.R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. Nature **428,** 281-286.

Wu, L., Zhou, H., Zhang, Q., Zhang, J., Ni, F., Liu, C., and Qi, Y. (2010). DNA methylation mediated by a microRNA pathway. Mol Cell **38,** 465-475.

Yoshizawa, S., Fourmy, D., and Puglisi, J.D. (1999). Recognition of the codon-anticodon helix by ribosomal RNA. Science **285,** 1722-1725.

Yuan, Y., and Altman, S. (1995). Substrate recognition by human RNase P: identification of small, model substrates for the enzyme. EMBO J **14,** 159-168.

Yusupova, G.Z., Yusupov, M.M., Cate, J.H., and Noller, H.F. (2001). The path of messenger RNA through the ribosome. Cell **106,** 233-241.

Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S.H., Liou, L.W., Barefoot, A., Dickman, M., and Zhang, X. (2011). *Arabidopsis* Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. Cell **145,** 242-256.

Ziehler, W.A., Day, J.J., Fierke, C.A., and Engelke, D.R. (2000). Effects of 5' leader and 3' trailer structures on pre-tRNA processing by nuclear RNase P. Biochemistry **39,** 9909-9916.

Ziehler, W.A., Morris, J., Scott, F.H., Millikin, C., and Engelke, D.R. (2001). An essential protein-binding domain of nuclear RNase P RNA. RNA **7,** 565-575.

Bernard GUTMANN

Étude d'un nouveau type de RNase P spécifique des eucaryotes chez *Arabidopsis thaliana*

Résumé

La RNase P est impliquée dans la maturation des ARNt en libérant l'extrémité 5' leader des précurseurs d'ARNt. Jusqu'à récemment, il était admis que cette enzyme soit universellement conservée en tant que complexe ribonucléoprotéique. La rupture avec le modèle établi est venue avec l'identification d'une RNase P uniquement protéique dans les mitochondries humaines et chez les plantes. Ces protéines, nommées PRORP (PROteinaceous RNase P), présentent trois paralogues chez *Arabidopsis*, qui sont localisés soit dans les organelles (PRORP1), soit dans le noyau (PRORP2 et 3). Des tests d'activité *in vitro* montrent que les protéines PRORP possèdent seules une activité RNase P. L'étude de lignées ADN-T indique que la fonction des protéines PRORP est essentielle et la fonction de PRORP2 et 3 est redondante. L'analyse des lignées de dérégulation montre que les protéines PRORP possèdent une diversité de substrat et que la RNase MRP, une autre ribonucléoprotéine, n'est pas impliquée dans la maturation des ARNt. Ainsi les protéines PRORP seraient bien les seules enzymes impliquées dans la maturation de l'extrémité 5' des ARNt chez *Arabidopsis*.

Mots clés : plante, RNase P, pentatricopeptide repeat, domaine NYN, maturation des ARNt, endonucléase

Résumé en anglais

RNase P is involved in the maturation of tRNA precursors by cleaving their 5' leader sequences. Until recently this enzyme was considered to be universally occurring as a ribonucleoprotein complex. The breakthrough from the existing model came with the identification of protein-only RNase P in human mitochondria as well as in plants. These proteins that we called PRORP (PROteinaceous RNase P) have three paralogs in *Arabidopsis*, which are localised in organelles (PRORP1) and nuclei (PRORP2 and 3). We have shown that PRORP proteins have RNase P activity *in vitro* as single proteins. *In vivo* the functions of PRORP proteins are essential and the function of PROPR2 and 3 are redundant. PRORP down-regulation mutants, show that PRORP proteins have a variety of other substrates and RNase MRP, another ribonucleoprotein, is not involved in the tRNA maturation. Results show that PRORP proteins would be the only enzymes responsible for RNase P activity in *Arabidopsis*.