

Université de Montréal

**Evolution structurale et fonctionnelle de la  
composante ARN de la RNase P mitochondriale.**

par  
Elias Seif

Département de biochimie  
Faculté de médecine

Thèse présentée à la Faculté des études supérieures  
en vue de l'obtention du grade de doctorat  
en Biochimie

Avril, 2005

©Elias Seif, 2005



W

4

U58

2005

V.115

**Direction des bibliothèques**

**AVIS**

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

**NOTICE**

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal  
Faculté des études supérieures

Cette thèse intitulée :

**Evolution structurale et fonctionnelle de la  
composante ARN de la RNase P mitochondriale.**

présentée par :

Elias Seif

a été évaluée par un jury composé des personnes suivantes :

Dr. Serguei Chteinberg, président-rapporteur

Dr. B. Franz Lang, directeur de recherche

Dr. Pascal Chartrand, membre du jury

Dr. John Burke, examinateur externe

Dr. Pierre Belhumeur, représentant du doyen de la FES

## Résumé

La ribonucléase P (RNase P) est une endonucléase responsable de la maturation en 5' des ARNt chez les eubactéries, les archéobactéries, les eucaryotes, et les organelles. La RNase P est une ribonucléoprotéine dans les trois domaines de la vie, exception faite du chloroplaste de l'épinard et de la mitochondrie humaine où on suspecte l'absence de la sous-unité ARN (ARN-P). Selon les études comparatives, les ARN-P des trois règnes partagent un ancêtre commun. Les ARN-P eubactériens constituent le site catalytique de la RNase P et sont capables de catalyser la maturation des ARNt *in vitro*, en absence de la composante protéique (protéine-P).

Chez les eucaryotes, la RNase P mitochondriale (RNase Pmt) et nucléaire sont distinctes, il a été proposé que l'ARN-P mitochondriale (ARN-Pmt) découle d'un ancêtre commun proche des  $\alpha$ -protéobactéries. La RNase Pmt la mieux étudiée est celle de *S. cerevisiae*, sa sous-unité ARN est codée par le gène *rnpB* dans l'ADN mitochondrial, alors que la composante protéique est codée par un gène nucléaire. La présence de *rnpB* est irrégulière dans les génomes mitochondriaux. Il n'a été identifié que chez le jakobide *R. americana*, le prasinophyte *Nephroselmis olivacea*, et certains ascomycètes.

Mon projet a été établi dans le but d'éclaircir l'évolution de la sous-unité ARN-Pmt avec les objectifs suivants: (i) identifier le plus grand nombre possible de nouveaux gènes *rnpB*, (ii) établir des modèles de structure secondaire pour les ARN-Pmt identifiés, (iii) retracer l'évolution structurale cet ARN, et (iv) déterminer l'impact de cette évolution sur l'activité catalytique. Pour se faire, nous avons effectué des recherches *in silico* pour le gène *rnpB* dans tous les génomes mitochondriaux séquencés. Notre recherche n'était fructueuse que chez les ascomycètes, les zygomycètes et les jakobides. Chez les zygomycètes, notre collection de gènes *rnpB* a été complétée en identifiant quatre nouveaux candidats par PCR. Par la suite, nous

avons établi des modèles de structures secondaires pour tous les nouveaux ARN-Pmt par comparaison phylogénétique, en se basant sur la structure de *Reclinomonas americana*, qui contient tous les éléments structuraux présents dans le modèle consensus minimal eubactérien. Nous avons confirmé l'expression et la maturation de six ARN-Pmt *in vivo* comme prévu par nos modèles, en cartographiant expérimentalement leurs extrémités. En comparant les modèles de structure secondaire, nous avons noté une perte graduelle des hélices ancestrales, un phénomène qui s'accélère chez les champignons. Nous avons testé l'activité catalytique des six ARN-Pmt les mieux conservés (jakobides, prasinophyte, et *Rhizopus*), qui se sont révélés inactifs *in vitro*. Cependant, nous avons produit des ARN-P chimériques catalytiques en fusionnant les séquences de l'ARN-P d'*E. coli* avec celles des ARN-Pmt des jakobides et de *N. olivacea*. Les études de mutagenèse sur ces chimères ont permis d'identifier des éléments structuraux potentiellement responsable de l'inactivité catalytique des ARN-Pmt. Nous suggérons que la réduction structurale des ARN-Pmt est compensée *in vivo* par la composante protéique. Nous discutons aussi des scénarios possibles pour expliquer l'absence du gène *rnpB* des génomes mitochondriaux de la majorité des organismes étudiés.

**Mots-clés :** ribozyme, ribonucléoprotéine, jakobides, ascomycètes, zygomycètes, ARNt, structure secondaire, comparaison phylogénétique, motif GNRA, endonucléase.

## Abstract

Ribonuclease P (RNase P) is an endonuclease that cleaves 5' leader sequences from tRNA precursors in Eubacteria, Archaea, Eukarya and organelles. In most studied cases, RNase P is a ribonucleoprotein containing one RNA component (P-RNA), and one or several protein subunits (P-protein), with possible exceptions in spinach chloroplasts and human mitochondria, where protein-only ribonucleases are suspected. Comparative studies suggest that P-RNAs from all domains of life share a common ancestor. Eubacterial P-RNAs have been shown to carry the catalytic centre of RNase P, and are active in the absence of the P-protein.

Mitochondrial RNase P (mtRNase P) is distinct from its nuclear counterpart, and the mitochondrial RNA subunit (mtP-RNA) is thought to derive from the  $\alpha$ -proteobacterial P-RNA. MtRNase P of *S. cerevisiae* has been intensely investigated. Its RNA subunit is encoded by *rnpB* in the mitochondrial DNA, whereas the P-protein is encoded by a nuclear gene. The presence of *rnpB* in mitochondria of other eukaryotes is patchy; it has been identified in the jakobid *Reclinomonas americana*, the prasinophyte *Nephroselmis olivacea* and few ascomycetes.

My project aims to elucidate the structural and functional evolution of mtP-RNA. The objectives are (i) to identify *rnpB* in mitochondrial genomes, (ii) to establish secondary structure models for mtP-RNAs from several phylogenetic groups, (iii) to trace back the structural evolution of the RNA subunit, and (iv) to determine the impact of structural changes on the catalytic activity. Using *in silico* procedures, we searched all published mitochondrial genomes for *rnpB* and identified new genes in ascomycetes, zygomycetes, and jakobids. In zygomycetes we identified four additional *rnpB* genes by PCR amplification. The extremities of six mtP-RNAs were determined experimentally; they all match the ends as predicted in secondary

structure models. Secondary structure models were established for all newly identified mtP-RNAs based on comparisons with the *R. americana* RNA which shares all structural elements with the minimum bacterial consensus. Structural comparisons reveal the loss of various elements in mtP-RNAs of fungi.

Catalytic activity assays on the six best conserved (three jakobids, a prasinophyte and two zygomycetes) mtP-RNAs reveal that they are inactive *in vitro*. However, fusions of *E. coli* and mitochondrial P-RNA sequences from jakobids and *N. olivacea* produce active chimerical P-RNAs. Based on the analysis of mutations introduced into chimerical P-RNAs, we are able to pinpoint reasons for mtP-RNAs failure to cleave tRNA precursors *in vitro*. We suggest that the mitochondrial P-protein compensates for the reduction of secondary structure of mtP-RNAs. We will also discuss possible scenarios for the absence of *rnpB* from most mitochondrial genomes.

**Keywords :** ribozyme, ribonucleoprotein, jakobids, ascomycetes, zygomycetes, tRNA, Secondary structure, phylogenetic comparison, GNRA motif, endonuclease.



## Table des matières

Résumé .....	ii
Abstract .....	iv
Table des matières .....	vi
Liste des tableaux.....	xi
Liste des codes des acides aminés .....	xvi
Remerciements .....	xviii
Chapitre I : Introduction	
I.1. Homologie de la RNase P dans les trois domaines de la vie.....	3
I.2. Fonctions cellulaires de la RNase P.....	7
I.2.1 Fonctions cellulaires de la RNase P chez les eubactéries .....	7
I.2.1.1 Régulation du cycle lytique du bactériophage P4.....	7
I.2.1.2 Maturation de l'ARN 4,5S.....	8
I.2.1.3 Régulation post-traductionnelle par la RNase P .....	9
I.2.1.3.1 Stabilisation de l'ARNm de l'opéron <i>his</i> chez <i>S. typhimurium</i> .....	9

I.2.1.3.2 La RNase P précipite la dégradation des ARNm chez <i>E. coli</i> .....	10
I.2.1.3.3 La RNase P est impliquée dans le transport membranaire .....	11
I.2.1.4 Maturation du ARNtm (ARN 10Sa) .....	14
I.2.1.5 Maturation des ARNr chez les eubactéries .....	17
I.2.2 Fonctions cellulaire de la RNase P chez les eucaryotes.....	17
I.2.2.1 Rôle de la RNase P eucaryote dans la maturation des ARNr.....	18
I.2.2.2 Rôle de la RNase P eucaryote chez les virus.....	22
I.2.2.2.1 Rôle de la RNase P dans la réplication chez les rétrovirus .....	22
I.2.2.2.2 La RNase P clive les virus des plantes en 3' .....	25
I.3. Reconnaissance du substrat.....	27
I.3.1 Les interactions entre le pré-ARNt et les composantes de la RNase P.....	27
I.3.2 Contribution modulaire à la reconnaissance du pré-ARNt .....	30
I.3.3 Reconnaissance des substrats autres que les pré-ARNt.....	30
I.3.4 Reconnaissance du substrat chez les eucaryotes et les archéobactéries .....	31
I.4. Un pseudo-nœud universel chez la RNase P.....	33
I.4.1 Site d'attache des ions Mg <sup>2+</sup> .....	33

	viii
I.4.2 La réaction de catalyse .....	36
I.5. La RNase P mitochondriale.....	39
I.5.1 Origine de la mitochondrie.....	39
I.5.2 La RNase P mitochondriale chez les champignons .....	43
I.5.3 L'impact de la génomique.....	46
I.6. Définition du projet.....	49
Chapitre II : Articles .....	50
Article 1 .....	51
Article 2.....	63
Article 3.....	75
Chapitre III : Discussion .....	119
D.1 Une nouvelle approche pour identifier les gènes <i>rnpB</i> .....	120
D.2 Evolution de la structure secondaire des ARN-Pmt.....	121
D.2.1 Modélisation par comparaison phylogénétique .....	121
D.2.2 Les structures secondaires des ARN-Pmt des isolats des jakobides sont les mieux conservées.....	121

D.2.3 Evolution convergente vers une structure minimale chez les champignons ...	125
D.2.4 Les ARN-Pmt cibles d'invasion des éléments à double hélices «DHE » .....	126
D.2.5 La perte des éléments structuraux ne suit pas un ordre défini .....	129
D.2.6 Des motifs GNRA et YNMG dans les tétra-boucles du P3, P9, P12 et P18 ...	134
D.3 Effets de la réduction structurale sur l'activité catalytique des ARN-Pmt.....	137
D.3.1 Les ARN-Pmt ne sont pas catalytiques <i>in vitro</i> .....	137
D.3.2 Le domaine P6, 16-17 ne suffit pas pour rendre l'ARN-Pmt de <i>R. americana</i> NZ actif .....	140
D.3.3 La fusion de l'ARN-P d' <i>E. coli</i> avec les ARN-Pmt produit des ARN catalytiques.....	140
D.3.4 Raison multifactorielle pour l'absence d'activité catalytique chez les ARN-Pmt .....	143
D.3.5 Relation entre la conservation de la structure secondaire et l'activité catalytique des chimères .....	144
D.4 La composante protéique compense-t-elle la réduction structurale? .....	146
D.5 Scénarios possibles pour expliquer la distribution discontinue du gène <i>rnpB</i> ...	148
D.5.1 Les gènes <i>rnpB</i> sont dérivés et non-reconnaissables avec notre approche.....	148
D.5.2 Le gène <i>rnpB</i> codé par le génome nucléaire.....	149

D.5.3 Perte de la composante ARN .....	149
D.5.3.1 La composante ARN est absente chez la RNase P du chloroplaste des épinards .....	149
D.5.3.2 La RNase P mitochondriale chez les humains : protéique ou nucléaire? .....	150
D.5.3.2.1 RNase P mitochondriale humaine protéique.....	150
D.5.3.2.2 L'ARN-Pmt humaine est identique à l'ARN-P de la RNase P nucléaire..	151
Conclusions .....	153
Bibliographie .....	clv

## Liste des tableaux

Table D.1: Comparaison entre les tailles des ARN-Pmt et la réduction de leur structure secondaire.....	128
Table D.2 : Fréquences de présence des hélices d'origine eubactérienne chez les ARN-Pmt.....	133

## Liste des figures

Figure I.1: Modèles de structure secondaire représentant la structure de l'ARN-P chez les trois domaines de vie.....	5
Figure I.2: Schéma résumant les fonctions cellulaires de la RNase P chez <i>E. coli</i> et du lien de ces fonctions au transport membranaire.....	12
Figure I.3: Modèle d'action de l'ARNtm.....	15
Figure I.4: Comparaison des pré-ARNr d' <i>E. coli</i> et de <i>Saccharomyces cerevisiae</i> . ...	20
Figure I.5: Modèle du clivage de l'ARNtMet initiateur et de la formation de l'amorce pour la réplication des particules ressemblant aux rétrovirus chez la drosophile (Retrovirus-like particles, RVLP). .....	23
Figure I.6 : Structure globale du complexe RNaseP/pré-ARNt d' <i>E. coli</i> .....	28
Figure I.7: Région d'interaction des ions métalliques avec la région du P1-P4.....	34
Figure I.8: Mécanisme de la catalyse par la RNase P.....	37
Figure I.9: Phylogénie des mitochondries et des $\alpha$ -protéobactéries .....	41
Figure I.10: Modèle de structure secondaire des ARN-P mitochondriaux de <i>S. cerevisiae</i> , <i>S. fibuligera</i> , <i>Candida glabrata</i> , et <i>A. nidulans</i> .....	44
Figure I.11: Distribution du gène <i>rnpB</i> chez les eucaryotes .....	47
Figure D.1: Modèle de structure secondaire de l'ARN-P de <i>Rickettsia prowazekii</i> et de l'ARN-Pmt de <i>R. americana</i> NZ.....	123

Figure D.2: Scénario hypothétique décrivant la perte des éléments structuraux chez les ARN-Pmt.....	131
Figure D.3: Modèle du consensus minimal eubactérien et du « Micro ARN-P ».....	138
Figure D.4: Chimère d'ARN-P de <i>R. americana</i> NZ après l'ajout des P5,6,15,16,17. .....	141



## Liste des abréviations

A	Adénine
ADN	Acide déoxyribonucléique
ADNc	ADN complémentaire
ADNmt	ADN mitochondrial
ARN	Acide ribonucléique
ARN M1	Sous-unité d'ARN de la ribonucléase P d' <i>E. coli</i>
ARN-Pmt	Sous-unité d'ARN de la ribonucléase P mitochondriale
ARNr	ARN ribosomique
ARNt	ARN de transfert
C	Cytosine
C5	Sous-unité protéique de la ribonucléase P d' <i>E. coli</i>
<i>E. coli</i>	<i>Escherichia coli</i>
G	Guanine
kb	Kilo paires de bases (ADN) ou kilo bases (ARN)
kDa	Kilodaltons
ARN-P	Sous-unité d'ARN de la ribonucléase P
Protéine-P	Sous-unité protéique de la ribonucléase P

pré-ARNt	Précurseur de l'ARN de transfert
ARN M1	Sous-unité d'ARN de la ribonucléase P d' <i>E. coli</i>
ARN-Pmt	Sous-unité d'ARN de la ribonucléase P mitochondriale
RNase III	Ribonucléase III
RNase E	Ribonucléase E
RNase MRP	Ribonucléase "Mitochondrial RNA Processing"
RNase P	Ribonucléase P
<i>rnpB</i>	Gène codant la sous-unité ARN de la RNase P
S	Unité de constante de sédimentation, Svedberg
<i>trn</i>	Gène codant un ARNt
T	Thymine
U	Uracyle

## Liste des codes des acides aminés

A	Ala	Alanine
B	Asx	Asparagine ou acide aspartique
C	Cys	Cystéine
D	Asp	Acide aspartique
E	Glu	Acide glutamique
F	Phe	Phénylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Méthionine
N	Asp	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Sérine

T	Thr	Thréonine
V	Val	Valine
W	Trp	Tryptophane
Y	Tyr	Tyrosine
Z	Glx	Glutamine ou acide glutamique

## Remerciements

Je tiens à remercier avec beaucoup de gratitude mon directeur de recherche, le Dr. Franz Lang, pour sa disponibilité, sa patience, sa compréhension, et pour m'avoir guidé tout au long de ce projet par ses nombreuses suggestions et commentaires.

Je veux remercier très sincèrement Lise Forget pour les améliorations suggérées pour cette thèse, ainsi que pour l'aide et les conseils techniques qu'elle m'a offerts depuis mon arrivée dans le laboratoire. Également, je tiens à la remercier pour son amitié et les discussions animées que nous avons partagées.

Ma reconnaissance va également à Jean-Francois Bouffard, Charles Bullerwell, Alexandre Cadieux, Eric Delage, Yannick Jacob, Liisa Koski, Marie-Josée Laforest, Julien Lagarde, Jessica Leigh, Naiara Rodriguez-Ezpeleta, et Zhang Wang pour leurs conseils techniques, les protocoles, les suggestions fructueuses qu'ils m'ont fournis, et leur compagnie qui m'est très chère.

Je veux remercier aussi les membres du groupe de l'OGMP, dirigé par Dr Gertraud Burger, pour leur soutien technique, pour l'accès à l'équipement de séquençage et aux ordinateurs qui furent essentiels pour les travaux de ce projet de recherche.

Finalement, je remercie le Conseil de Recherches Médicales du Canada pour le soutien financier qui a permis de réaliser ce projet.

*Je dédis ce mémoire à mes parents, mon frère et ma soeur, qui m'ont encouragé constamment avec patience, et amour pour passer à travers l'épreuve du doctorat.*

*«L'évolution procède comme un bricoleur qui pendant des millions et des millions d'années, remanierait lentement son oeuvre, la retouchant sans cesse, coupant ici, allongeant là, saisissant toutes les occasions d'ajuster, de transformer, de créer.»*

François JACOB

## **Chapitre I :Introduction**



La découverte de la structure de l'ADN, et du rôle de ce dernier dans la conservation et la transmission de l'information génétique, ainsi que la découverte du rôle des ARNm, et du système de traduction ont permis à Francis Crick en 1958 de formuler ce qu'il appela le « dogme central de la biologie moléculaire »: l'information génétique est transmise des acides nucléiques aux protéines et l'ADN n'est que le support moléculaire d'une information qui s'exprime à travers les protéines, notamment enzymatiques. Ce dogme limitait le rôle de l'ARN à celui de messenger, d'ARNt, ou d'ARNr. Toutefois, les résultats obtenus par H. Fraenkel-Conrat et G. S. Schramm sur le virus de la mosaïque du tabac prouvèrent que l'ARN peut lui aussi servir de dépositaire de l'information génétique. Ce dogme fut encore transgressé, et le rôle de l'ARN revue suite à la découverte de deux ribozymes, l'intron auto-épissable de l'ARN ribosomique de *Tetrahymena thermophila* par le groupe de Thomas R. Cech (Kruger et al., 1982) et l'activité catalytique de l'ARN (ARN M1) de la ribonucléase P (RNase P) d'*Escherichia coli* par le groupe de Sidney Altman (Guerrier-Takada et al., 1983). Dorénavant, l'ARN sera reconnu comme ayant des propriétés catalytiques. En 1989, l'importance de ces découvertes a été soulignée par l'attribution du prix Nobel en chimie aux deux chercheurs, conjointement. Depuis, la famille des ribozymes naturels a grandi par l'ajout de nouveaux introns du groupe I et II (Peebles et al., 1986), et des ribozymes «en tête de marteau» (Hammerhead; (Forster et Symons, 1987), «en épingle à cheveux» (Hairpin; (Feldstein et al., 1989; Hampel et Tritz, 1989; Haseloff et Gerlach, 1989), du virus de l'hépatite delta (Wu et al., 1989), «Varkud satellite» de *Neurospora* (Saville et al., 1990), et des « riboswitch » qui clivent les ARNm (Barrick et al., 2004; Mandal et al., 2003; Nahvi et al., 2004); (Winkler et al., 2004).

L'hypothèse originale qui écartait le rôle catalytique de l'ARN était fondée sur l'observation que ce dernier est peu diversifié en terme de groupements fonctionnels par rapport aux protéines. A pH neutre, il ne possède aucun groupe fonctionnel donneur ou accepteur de protons, et aucun groupement chargé positivement, alors que les protéines contiennent les lysines et arginines qui sont chargées positivement à

pH7. Chez les ARN catalytiques, cette limite est détournée par l'utilisation des cations qui peuvent agir comme des acides selon la définition de Lewis, des accepteurs d'électrons, d'où la dépendance de l'activité catalytique des ribozymes aux ions métalliques. Le cation divalent le plus fréquemment utilisé est le magnésium, qui peut être remplacé par le manganèse très souvent. D'autres cations sont aussi utilisés comme le calcium, le plomb, ou le cobalt. En plus de leur rôle dans le catalyse, les cations divalents ( $Mg^{++}$ ,  $Mn^{++}$  ...), et les cations monovalents ( $Na^+$ ,  $Li^+$  ...) neutralisent les charges négatives très fortes présentes sur les groupes phosphates des ARN (Saenger, 1984), et par conséquent peuvent stabiliser le repliement des conformations actives des ribozymes .

Aussitôt découverts, les ribozymes ont constitué un argument pour avancer la théorie du «monde ARN», selon laquelle la vie sur terre a débuté avec des molécules simples probablement des ARN qui, à la fois, encodent l'information génétique, et catalysent les réactions chimiques nécessaires pour le maintien d'une forme primitive de vie. Dans ce contexte, l'ARN de la RNase P, l'ARN-P, fut considéré comme une relique de ce monde. Sa présence dans les trois domaines de la vie, les eubactéries, les archéobactéries, et les eucaryotes, suggère que l'apparition de cet ARN est très ancienne, et qu'elle précède possiblement la divergence des trois groupes.

Dans l'introduction de cette thèse, je discuterai des fondements d'un consensus parmi les scientifiques qui propose une origine commune pour la composante ARN de la RNase P dans les trois domaines de la vie, pour ensuite décrire brièvement les connaissances actuelles sur le rôle cellulaire de la RNase P, son mode de reconnaissance du substrat et son mécanisme de catalyse. Par la suite, je discuterai l'origine de la RNase P mitochondriale, et de l'état des connaissances sur ce sujet avant le début de mon projet de doctorat.

## I.1. Homologie de la RNase P dans les trois domaines de la vie

Il existe un consensus parmi les scientifiques selon lequel les ARN-P présents dans les trois domaines de la vie partagent le même ancêtre. Ce consensus est supporté par la fonction commune que ces ARNs accomplissent, et par des études comparatives au niveau de la séquence primaire et de la structure secondaire.

Jusqu'à ce jour, la RNase P est la seule ribonucléase connue qui coupe spécifiquement la région guide en 5' des précurseurs des ARNt (5'leader sequence), et ceci dans les trois domaines de la vie ainsi que dans les organelles des eucaryotes. Il s'agit d'une fonction essentielle, puisque, chez *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae*, et *E. coli*, l'inactivation de la RNase P nucléaire par l'élimination du gène codant la composante ARN condamne la cellule à la mort (Cherayil et al., 1987; Lee et al., 1991; Waugh et Pace, 1990). A l'exception possible dans quelques organelles (voir article 1), il a été montré que l'activité catalytique de la RNase P dépend de la sous-unité ARN chez les eucaryotes et les procaryotes, et ceci en abolissant l'activité catalytique de la RNase P par un traitement à la ribonucléase de *Micrococcus* (Franklin et al., 1995; Stark et al., 1978), en éliminant le gène codant l'ARN-P (voir plus haut), ou en montrant que les ARN-P eubactériens et quelques ARN-P archéobactériens sont actifs *in vitro* (Gardiner et al., 1985; Guerrier-Takada et al., 1983b; Pannucci et al., 1999; Wagner et al., 2001).

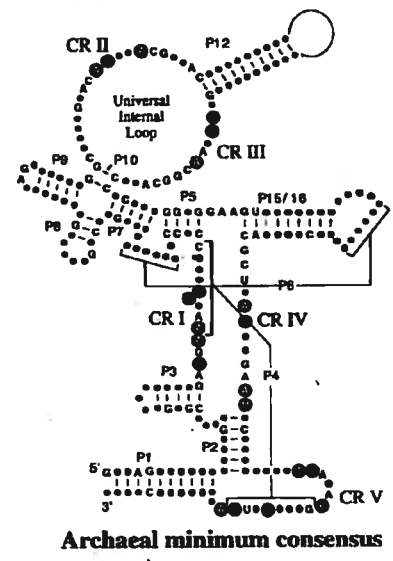
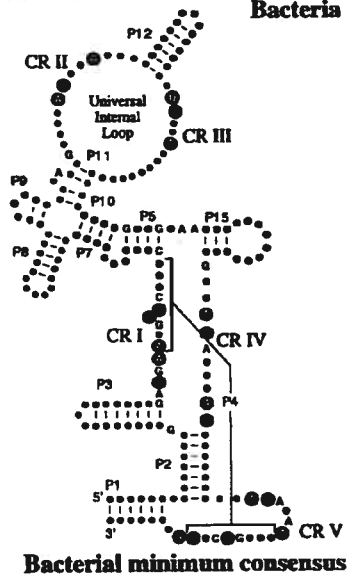
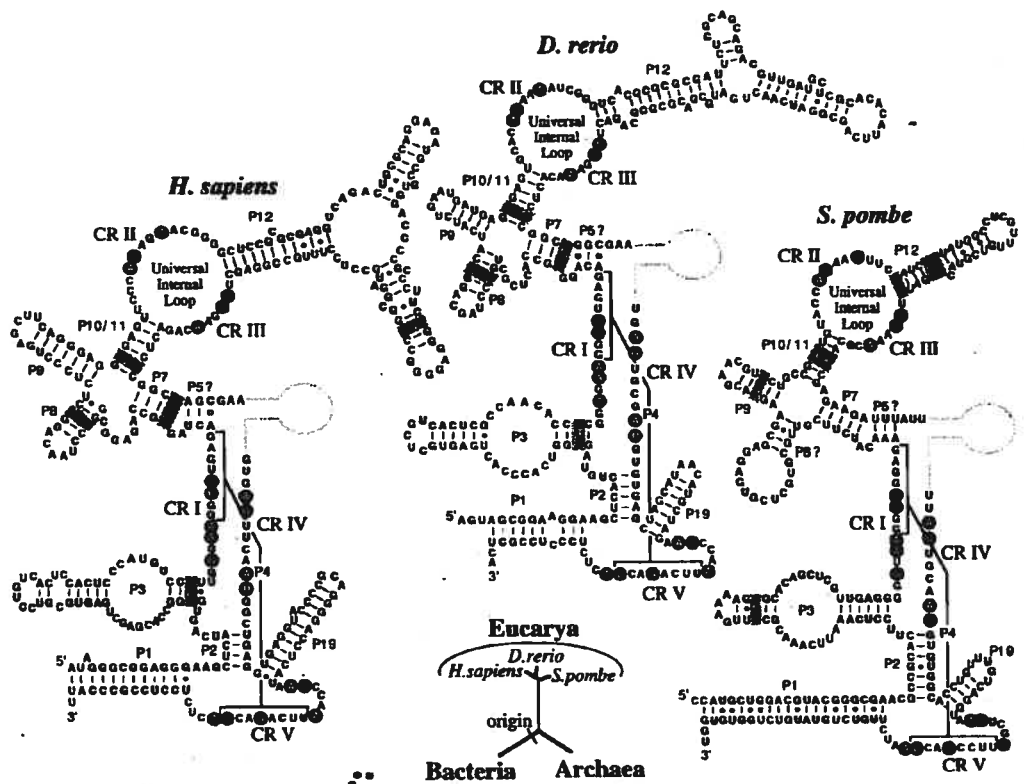
L'origine commune des ARN-P provenant des différents règnes de vie est supportée par les études comparatives. Suite à l'alignement de dix séquences d'ARN-P, provenant des eubactéries, archéobactéries, eucaryotes et une séquence mitochondriale, cinq motifs universellement conservés ont été identifiés au niveau de la séquence primaire (CRI-V)(Chen et Pace, 1997). Ces régions ont été identifiées aussi chez les ARN-P des plastides (de la Cruz et Vioque, 2003), organelles responsables de la photosynthèse. Cette homologie est encore plus marquée au niveau

de la structure secondaire, puisque comme chez la plupart des ARN, la conservation au niveau de la structure secondaire des ARN-P est plus forte que celle observée au niveau de la séquence primaire. Les études de modélisation des ARN-P par comparaison phylogénétique ont mené à des structures secondaires consensus chez les eubactéries (Siegel et al., 1996) et les archéobactéries (Haas et al., 1996), similaires entre elles et aux modèles proposés pour les ARN-P des eucaryotes *S. pombe*, *Homo sapiens* et *Dino rerio* (Frank et al., 2000; Tranguch et Engelke, 1993). Tous ces modèles partagent un noyau conservé qui inclut une structure en forme de cage maintenue par un pseudo-nœud (Chen et Pace, 1997; Forster et Altman, 1990b), une boucle interne entre les hélices P11-14, et quatre hélices en forme de croix (Figure I.1) (Chen et Pace, 1997).

En se basant sur les données décrites ci-dessus, nous pouvons conclure qu'une ARN-P était présente chez l'ancêtre commun des trois domaines de la vie, il y a 3 à 4 milliards d'années (Feng et al., 1997), une estimation qui restent très controversée. Il est fort probable que l'apparition de cet ARN date de la même époque que l'apparition des ARNt et des ribosomes.

Figure I.1: Modèles de structure secondaire représentant la structure de l'ARN-P chez les trois domaines de vie.

Les modèles du consensus minimal des eubactéries (Siegel et al., 1996) et des archéobactéries (Haas et al., 1996), dans lesquels les nucléotides invariables sont indiqués par des lettres. Les nucléotides qui sont universellement conservés sont indiqués par des cercles pleins. Pour les structures des ARN-P nucléaires de *H. sapiens*, *Danio rerio* et *S. pombe*: le nom d'hélice contenant un point d'interrogation, marque une incertitude de l'annotation de cette structure à cause de la variation de séquence primaire; les lignes grises indiquent la position du P15 par rapport aux modèles bactériens; dans les boîtes, on indique les covariations qui appuient la présence de ces structures (Frank et al., 2000; Tranguch et Engelke, 1993). Cette figure a été tirée de Chen et Pace, (1997).



## **I.2. Fonctions cellulaires de la RNase P**

Comme mentionné plus haut, la RNase P dans les trois règnes de vie est responsable de la maturation des précurseurs des ARNt (pré-ARNt). Cependant, la RNase P peut couper plusieurs autres substrats qui ressemblent structurellement à des ARNt.

### **I.2.1 Fonctions cellulaires de la RNase P chez les eubactéries**

Les rôles accomplis par la RNase P chez les eubactéries sont les mieux connus. Elle est responsable de la maturation de l'ARN CI, l'ARN 4,5S, certains ARNm, l'ARNtm, et les ARNr.

#### **I.2.1.1 Régulation du cycle lytique du bactériophage P4**

L'état de lysogénie du bactériophage P4 est sous le contrôle d'un petit ARN, appelé ARN CI, qui cause la terminaison précoce de la transcription de l'opéron responsable de la lyse cellulaire (Briani et al., 2001). Cet ARN confère une immunité pour la cellule hôte contre le bactériophage P4 en s'attachant à des séquences complémentaires à sa propre séquence, située en amont du premier gène de l'opéron ciblé; par la suite, la transcription sera terminée par un mécanisme dépendant du facteur Rho (Briani et al., 1996). Il a été montré *in vivo*, que le rôle de la RNase P dans ce processus est au niveau de la maturation en 5' de l'ARN CI (Forti et al., 1995). Selon le modèle prédit de la structure secondaire, les extrémités de cet ARN peuvent se replier pour former une hélice, un CCA protubérant en 3', et une région simple brin en 5' (Forti et al., 2002), ce qui constitue un substrat idéal pour la RNase P (Hansen et al., 2001).

### I.2.1.2 Maturation de l'ARN 4,5S

L'ARN 4,5S a été baptisé ainsi en rapport avec sa mobilité électrophorétique sur les gels de polyacrylamide. Il est très stable; *in vivo*, il est présent à une concentration comparable à celle d'un ARNt. L'ARN 4,5S se retrouve chez plusieurs procaryotes et le gène qui le code, le *ffs*, est essentiel pour la survie d'*E. coli*. Cet ARN est l'homologue structural de l'ARN 7S chez les archéobactéries et les eucaryotes. Ce dernier fait partie de la particule SRP (Signal Recognition Particle), une particule formée de six protéines et d'une molécule d'ARN, qui est essentielle pour la translocation des protéines à travers la membrane du réticulum endoplasmique (Poritz et al., 1988). Les études de complémentation ont montré que le gène codant l'ARN 7S peut remplacer le *ffs* chez *E. coli* (Brown, 1991). *In vivo*, l'ARN 4,5 S lie la protéine Ffh qui a une taille de 48kD. La Ffh est l'homologue de la protéine eucaryote SRP54, et fait partie d'une particule SRP eubactérienne impliquée dans le trafic des protéines de sécrétion chez les eubactéries (Nagai et al., 2003; Ribes et al., 1990). Plusieurs études ont montré que l'ARN 4,5S est aussi impliqué dans la traduction, où il facilite l'étape de translocation durant la synthèse des polypeptides sur le ribosome en interagissant avec le facteur d'élongation EF-G (Nakamura et al., 2001; Sagar et al., 2004).

La RNase P est responsable de la maturation de l'ARN 4,5S, comme suggéré par l'accumulation du précurseur de l'ARN 4,5S *in vivo* dans les mutants de *rnpB* thermosensibles (Bothwell et al., 1976). Ce rôle a été confirmé par des études *in vitro*, qui ont montré que la RNase P coupe d'une façon endonucléolytique l'extrémité 5' du précurseur pour générer deux fragments, l'ARN 4,5S mature et la région en 5' de cet ARN (Bothwell et al., 1976; Guerrier-Takada et Altman, 1984).



### **I.2.1.3 Régulation post-traductionnelle par la RNase P**

Des études effectuées chez *Salmonella typhimurium* et *E. coli* ont montré que la RNase P est impliquée dans la maturation des ARNm de plusieurs opérons (voir plus bas). Cette maturation affecte la stabilité de ces ARNm ce qui constitue une régulation post-traductionnelle de l'expression de ces opérons.

#### **I.2.1.3.1 Stabilisation de l'ARNm de l'opéron *his* chez *S. typhimurium***

Les gènes responsables de la biosynthèse de l'acide aminé histidine sont codés par l'opéron *his*. Chez *S. typhimurium*, les études de protection contre la digestion par la nucléase S1, des mesures de la stabilité de l'ARN, des expériences de retardement sur gel et des tests de transcription *in vitro* ont montré que les 5 cistrons les plus distaux (en 3') de l'ARNm de l'opéron *his* sont coupés du reste du message, et forment un ARN de 3900 nucléotides. Ce nouvel ARN est plus stable que son précurseur (Alifano et al., 1992). Les résultats d'une étude utilisant des mutants thermosensibles pour la RNase P et la RNase E ont montré que les deux ribonucléases ont un rôle dans la maturation de l'ARN de 3900 nucléotides (Alifano et al., 1994). Les investigations plus poussées ont montré que la RNase E coupe le transcrit primaire de l'opéron *his* à une distance de 620 nucléotides en amont de l'extrémité 5' de l'ARN 3900 nucléotides mature, formant ainsi un précurseur de cet ARN. Par la suite, ce précurseur sera coupé par la RNase P à un site situé à 76 nucléotides en amont du codon d'initiation de la traduction du cistron *hisB*. La coupure par la RNase P est située en 5' d'une région qui peut être repliée en une structure tige boucle suivie de la séquence NCCA (Alifano et al., 1994). Il a été suggéré que la stabilité de l'ARN de 3900 nucléotides est conférée par cette structure qui le protège contre les nucléases (Alifano et al., 1996; Bouvet et Belasco, 1992). L'efficacité de la coupure du précurseur par la RNase P est modulée par la traduction du message; cette coupure requiert que le ribosome soit attaché au cistron *hisB* (Alifano et al., 1992; Alifano et al., 1994). Il y a deux hypothèses pour expliquer comment les ribosomes

attachés à l'ARN du *hisB* augmentent l'efficacité de la coupure par la RNase P, soit en favorisant la formation de la structure tige-boucle qui peut être reconnue par la RNase P, soit en stabilisant temporairement le précurseur en le protégeant contre la dégradation 5' → 3', ce qui augmente le nombre des précurseurs intacts disponibles (Alifano et al., 1996). Ce système de stabilisation du messenger permet de coupler la stabilité du messenger à la disponibilité des ribosomes.

#### **I.2.1.3.2 La RNase P précipite la dégradation des ARNm chez *E. coli***

En utilisant des mutants d'*E. coli* thermosensibles pour la RNase P (cellules A49)(Baer et al., 1989), il a été montré que cette dernière est impliquée dans la maturation des ARNm des opérons *tna* (dégradation du tryptophane), *his* (biosynthèse de l'histidine), *rbs* (transport du D-ribose), *secG* (excrétion de protéines), *lac* (lactose), ainsi que des ORFs b0669 et b0671 (Li et Altman, 2003; Li et Altman, 2004). Les auteurs ont cloné et transcrit *in vitro* certaines régions intergéniques de ces messagers, et ils ont montré que la RNase P coupe à l'intérieur de ces régions, et entre le début du transcrit et le codon d'initiation de la traduction de l'opéron *secG*. Chez les mutants thermosensibles pour la RNase P, les auteurs ont observé une augmentation de la quantité d'ARNm du *tna*, et une disparition des intermédiaires de ce messenger qui sont les produits de la maturation par la RNase P. Dans ces mutants, l'activité catalytique du *tnaA* est plus élevée que dans les cellules normales, ce qui est compatible avec l'augmentation de la quantité d'ARNm observée. En ce basant sur ces données, et sur le fait que les produits de la RNase P sont des bons substrats pour la RNase E (Coburn et Mackie, 1999), les auteurs ont testé la sensibilité des messagers de *lacYA* avant et après la coupure par la RNase P à la digestion par la RNase E. Ils ont remarqué que la RNase E dégrade cet ARNm d'une façon très efficace après la coupure par la RNase P contrairement à ce qu'on observe pour le transcrit original. Suite à ces observations, les auteurs ont proposé un modèle selon lequel la RNase P coupe les transcrits de certains opérons, ce qui les rend disponible à la dégradation par la RNase E. Ce modèle a été confirmé *in vivo* en

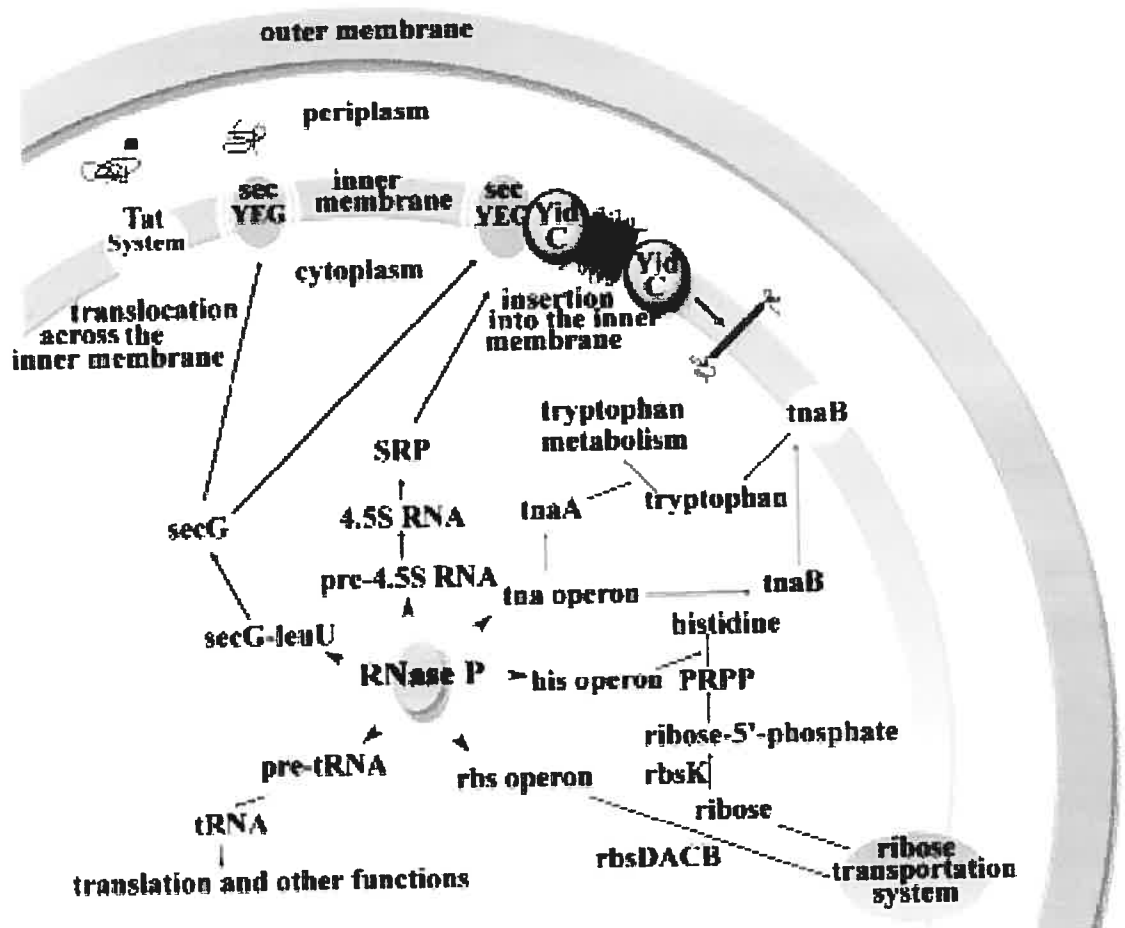
clonant la région intergénique de l'opéron *lac* en amont d'un gène rapporteur, et en transformant des cellules sauvages et mutants thermosensibles avec cette construction (Li et Altman, 2004).

#### **I.2.1.3.3 La RNase P est impliquée dans le transport membranaire**

En examinant les fonctions déjà décrites plus haut, on remarque que la RNase P est un élément clé dans le transport transmembranaire chez *E. coli*. Elle est impliquée dans quatre fonctions en rapport avec ce transport ; (i) la maturation de l'ARN de la SRP (Signal Recognition Particle) eubactérienne (ARN 4,5S) (ii) et de la région 5' non traduite du *secG*, ainsi que (iii) du contrôle de l'expression de la perméase du tryptophane et (iv) des gènes impliqués dans le transport du ribose (voir section I.2.1.3.2)(Figure I.2)(Li et Altman, 2003). Ce nouveau rôle n'avait pas été suspecté précédemment, et dans l'avenir, il serait important d'étudier la régulation de l'expression de la RNase P afin de mieux comprendre comment cette dernière fait le lien entre la traduction, en étant impliquée dans la maturation des ARNt et des ARNr (voir la section I.2.1.4), et le transport membranaire au niveau de la cellule.

Figure I.2 : Schéma résumant les fonctions cellulaires de la RNase P chez *E. coli* et du lien de ces fonctions au transport membranaire.

Cette figure a été tirée de (Li et Altman, 2003).



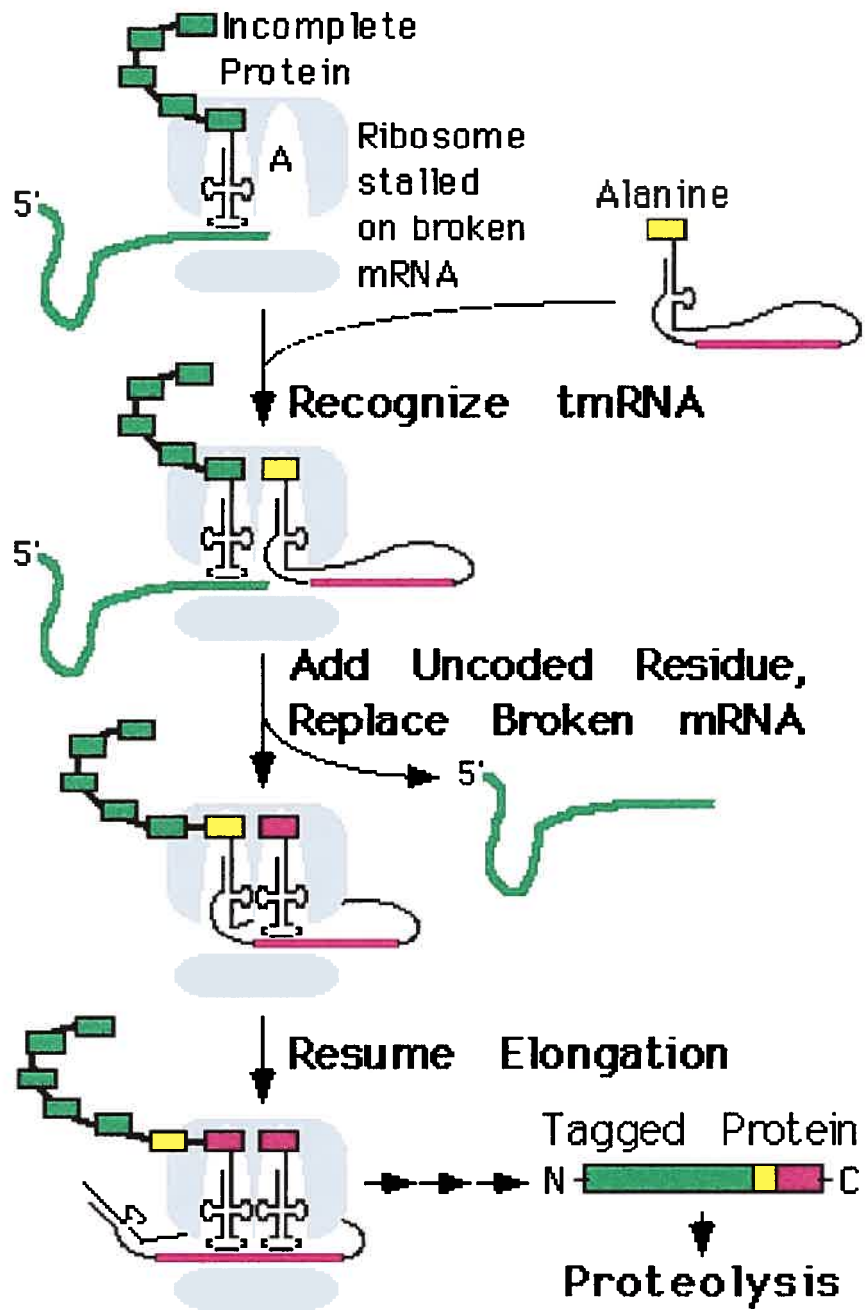
#### I.2.1.4 Maturation du ARN<sub>tm</sub> (ARN 10Sa)

Initialement, l'ARN<sub>tm</sub> était connu comme l'ARN 10Sa. Les études de modélisation comparative de ce gène chez *Mycobacterium tuberculosis*, *E. coli* et *Bacillus subtilis* ont révélé une similarité entre cet ARN et la moitié de la structure de l'ARN<sup>Ala</sup> qui correspond à la tige acceptrice et l'hélice T (Komine et al., 1994; Tyagi et Kinger, 1992; Ushida et al., 1994). La tige acceptrice de l'ARN 10Sa accepte l'acide aminé alanine *in vivo* et *in vitro* (Komine et al., 1994). L'autre partie de cet ARN encode une protéine "tag", et peut donc servir d'ARN<sub>m</sub> d'où son nouveau nom ARN<sub>tm</sub> (Gueneau de Novoa et Williams, 2004). Le rôle de l'ARN<sub>tm</sub> est de s'attacher aux ribosomes bloqués par des ARN<sub>m</sub> partiellement dégradés; une fois l'ARN<sub>tm</sub> attaché, sa partie codante sera traduite, et la protéine encodée s'ajoute à la protéine incomplète produite par l'ARN<sub>m</sub> brisé. Par la suite, la partie "tag" sera reconnue par une protéase associée à la membrane, la protéase tsp qui dégradera la protéine "tag" ainsi que la protéine partiellement traduite (Figure I.3)(Gueneau de Novoa et Williams, 2004). Par ce mécanisme, l'ARN<sub>tm</sub> débloque les ribosomes et permet de dégrader les protéines incomplètes.

Il a été montré que la RNase P coupe l'extrémité 5' du précurseur de l'ARN<sub>tm</sub> en cartographiant l'extrémité 5' d'un ARN<sub>tm</sub> provenant d'un mutant thermosensible pour la RNase P, et ceci par extension d'amorce (Komine et al., 1994). Dans le cadre de mes travaux de doctorat, j'ai participé à des expériences et des analyses dans lesquelles nous avons identifié pour la première fois des ARN<sub>tm</sub> mitochondriaux potentiels chez les jakobides, et nous avons montré que l'ARN-P d'*E. coli* (ARN M1) peut couper *in vitro* les précurseurs de ces ARN<sub>tm</sub> mitochondriaux putatifs (Jacob et al., 2004).

Figure I.3: Modèle d'action de l'ARNtm.

La première partie de la figure montre un ribosome bloqué par un ARNm incomplet ou partiellement dégradé. Ce ribosome reconnaît l'ARNtm, qui rentre dans le site A du ribosome, remplace l'ARN messenger et débloque le ribosome. Les résidus encodés par l'ARNtm seront ajoutés à la protéine partiellement synthétisée à partir de l'ARNm incomplet. Ces résidus ajoutés vont servir de signal pour la tsp qui va dégrader la protéine plus tard. La figure a été tirée du site de l'ARNtm (<http://www.indiana.edu/~tmRNA/>)(Gueneau de Novoa et Williams, 2004).





### **I.2.1.5 Maturation des ARNr chez les eubactéries**

Chez les eubactéries, la RNase P est impliquée dans la maturation des ARNr au niveau du transcrit primaire. Elle est responsable de la maturation en 5' des pré-ARNt présents dans la région intergénique situées entre l'ARN 30S et 16S et qui peut inclure un ou plusieurs ARNt (Figure I.4). Dans des cellules mutées pour la RNase III, la RNase P peut compenser cette déféctuosité en libérant les précurseurs des ARNs 16S et 23S ; ces molécules sont alors utilisables pour former des ribosomes fonctionnels (King et al., 1986; Morrissey et Tollervey, 1995). Par conséquent, il a été suggéré que la RNase P est une alternative à la RNase III dans la maturation des ARNr. Il est commun de trouver de la redondance dans les mécanismes qui produisent des molécules essentielles pour la survie de la cellule (Morrissey et Tollervey, 1995).

### **I.2.2 Fonctions cellulaire de la RNase P chez les eucaryotes**

La RNase P eucaryote est plus complexe que la RNase P eubactérienne. Chez les humains, elle contient aux moins dix sous-unités protéiques en plus de la sous-unité ARN (Guerrier-Takada et al., 2002), comparé à une seule protéine chez les eubactéries. Cependant, à part la maturation des ARNt, les rôles documentés de la RNase P eucaryote se limitent à seulement deux fonctions possibles et mal connues, la maturation des ARNr et la réplication des virus. Le plus grand nombre de protéines-P reflète la plus grande complexité des cellules eucaryotes par rapport aux eubactéries; alors que le petit nombre de fonctions accomplies découle de deux raisons possibles. Premièrement, il est fort probable que plusieurs fonctions restent inconnues, et deuxièmement, la complexité des cellules eucaryotes amène une spécialisation dans les fonctions; par exemple, la RNase P semble avoir perdu son rôle dans la maturation des ARNr au profit d'une nouvelle ribonucléase, la RNase MRP.

### I.2.2.1 Rôle de la RNase P eucaryote dans la maturation des ARNr

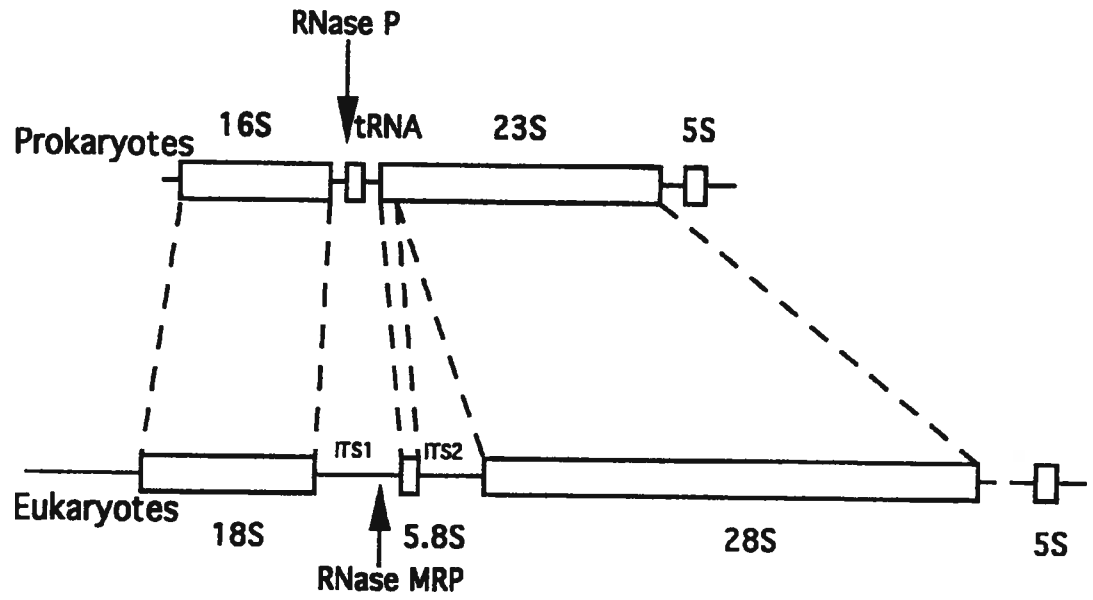
La ribonucléase "Mitochondrial RNA Processing" (RNase MRP) est essentiellement connue pour son rôle dans la réplication de l'ADN mitochondrial, (Morrissey et Tollervey, 1995). Cependant, il existe un parallèle entre le rôle de la RNase P chez les eubactéries, et celui de la RNase MRP chez les eucaryotes en ce qui concerne la maturation des ARNr. Les ARNt présents dans l'espaceur de l'ARNr eubactérien sont remplacés par une région d'espaceur interne transcrite (ITS1) chez les eucaryotes, et l'extrémité 5' de ce dernier est coupée par la RNase MRP (Figure I.4). Cette ribonucléase est une ribonucléoprotéine qui partage plusieurs caractéristiques avec la RNase P. Certains chercheurs proposent que la RNase MRP a évolué à partir de la RNase P suite à une duplication, après laquelle, la RNase MRP a acquis la fonction accomplie par la RNase P au niveau de la maturation des ARNr, celle de couper la région correspondant à l'espaceur. Cette hypothèse est supportée par la similarité entre les ARN de la RNase MRP et la RNase P au niveau de la structure primaire et secondaire (Forster et Altman, 1990b), et par le fait qu'elles partagent plusieurs de leurs composantes protéiques (Welting et al., 2004). La RNase MRP et la RNase P sont toutes les deux présentes dans le nucléole des eucaryotes, et selon les résultats des hybridations *in situ*, la RNase P est même associée à la RNase MRP. Cependant, il n'est pas clair si ces deux ribonucléoprotéines sont associées directement ou à travers un complexe macromoléculaire comme les pré-ribosomes (Lee et al., 1996a). Cette observation a amené les chercheurs à étudier le rôle de la RNase P dans la maturation des ARNr chez les eucaryotes.

Des mutations dans l'ARN-P nucléaire de *S. cerevisiae* causent l'accumulation d'un précurseur de l'ARNr 5,8S avec une extension de 35 nucléotides à l'extrémité 3', alors qu'une mutation dans la sous-unité ARN de la RNase MRP cause l'accumulation d'un précurseur avec une extension à l'extrémité 5' (Chamberlain et al., 1996). Le rôle de la RNase MRP à ce niveau est bien connu et documenté, cependant ce n'est pas le cas pour celui de la RNase P. Dans des tests *in vitro*, la RNase P n'est pas capable de couper

l'extrémité 3' du précurseur de l'ARNr 5,8S. Cette observation indique que le précurseur non mature isolé dans l'étude utilisant une ARN-P mutée n'est pas le produit direct de la mutation du gène *rnpB*. Dans ce cas, de quoi découle cette anomalie dans la maturation de l'ARN 5,8S ? Plusieurs hypothèses ont été émises. Il a été proposé que la RNase P est impliquée dans des voies de maturation de l'ARNr inconnues jusqu'à maintenant, ou que la mutation dans la RNase P affecte la structure des particules ribonucléoprotéiques dont elle fait partie, et par suite elle cause une

Figure I.4: Comparaison des pré-ARNr d'*E. coli* et de *Saccharomyces cerevisiae*.

L'arrangement de base est conservé, mais on note trois différences majeures. Premièrement, l'équivalent de la grande sous-unité ribosomique bactérienne (ARNr 23S) est divisée en deux par l'insertion de la région ITS2 (intergenic transcribed spacer) chez la majorité des eucaryotes. Deuxièmement, l'ARNr 5S est synthétisé d'une façon indépendante chez les eucaryotes. Troisièmement, les eubactéries possèdent des ARNt dans l'espaceur entre les ARNr 16S et 23S qui sont différents selon la bactérie. Les sites de coupure par les RNase P et MRP sont indiqués par des flèches. Cette figure a été tirée de (Lee et al., 1996a).



anomalie dans les voies de maturation des ARNr dans lesquelles ces particules sont impliquées; cependant ces hypothèses restent à démontrer (Chamberlain et al., 1998; Chamberlain et al., 1996).

### **I.2.2.2 Rôle de la RNase P eucaryote chez les virus**

Chez les virus, la RNase P peut cliver des éléments ressemblant à des ARNt dans les génomes ARN du virus de l'hépatite C et de plusieurs virus chez les plantes (voir plus bas). Cependant, le rôle le mieux connu de la RNase P eucaryote découle du fait qu'elle est responsable de la maturation des ARNt, et que ces derniers sont utilisés comme amorces pour l'initiation de la reverse transcription des génomes ARN des rétrovirus et des rétrotransposons. Chez le rétrotransposon de la famille *copia*, on note même que la RNase P est capable de couper pas seulement au 5' de l'ARNt, mais à l'intérieur de l'ARNt afin de produire l'amorce requis (voir plus bas). Cette réaction est nommée « hyperprocessing ».

#### **I.2.2.2.1 Rôle de la RNase P dans la réplication chez les rétrovirus**

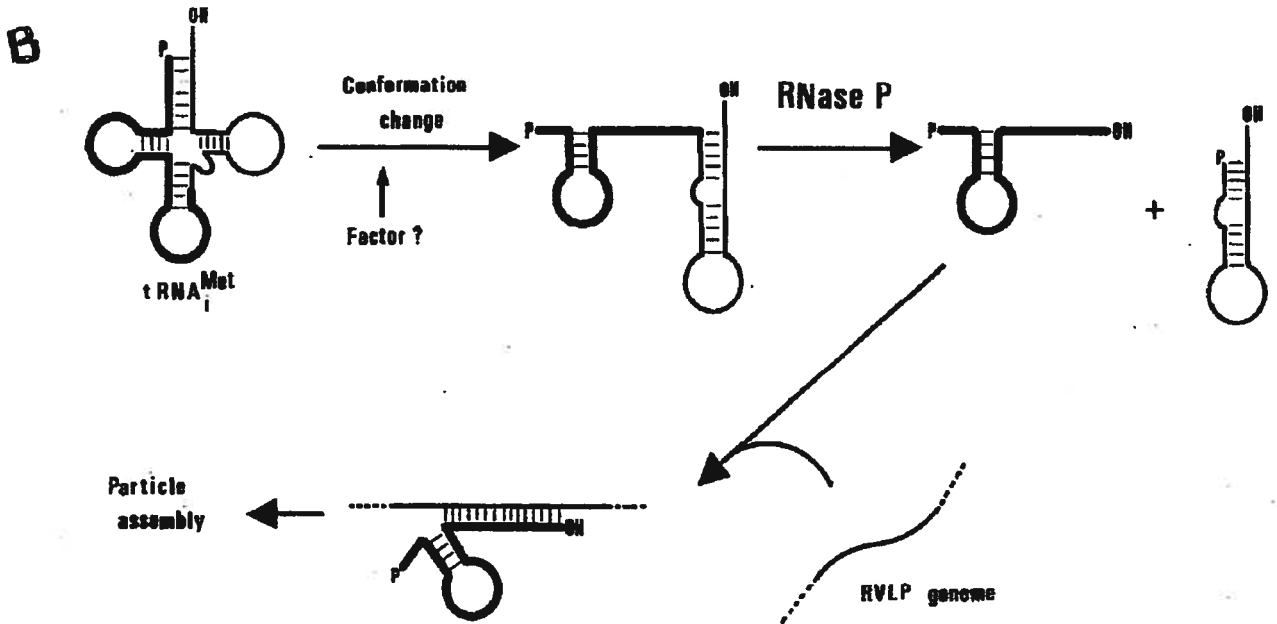
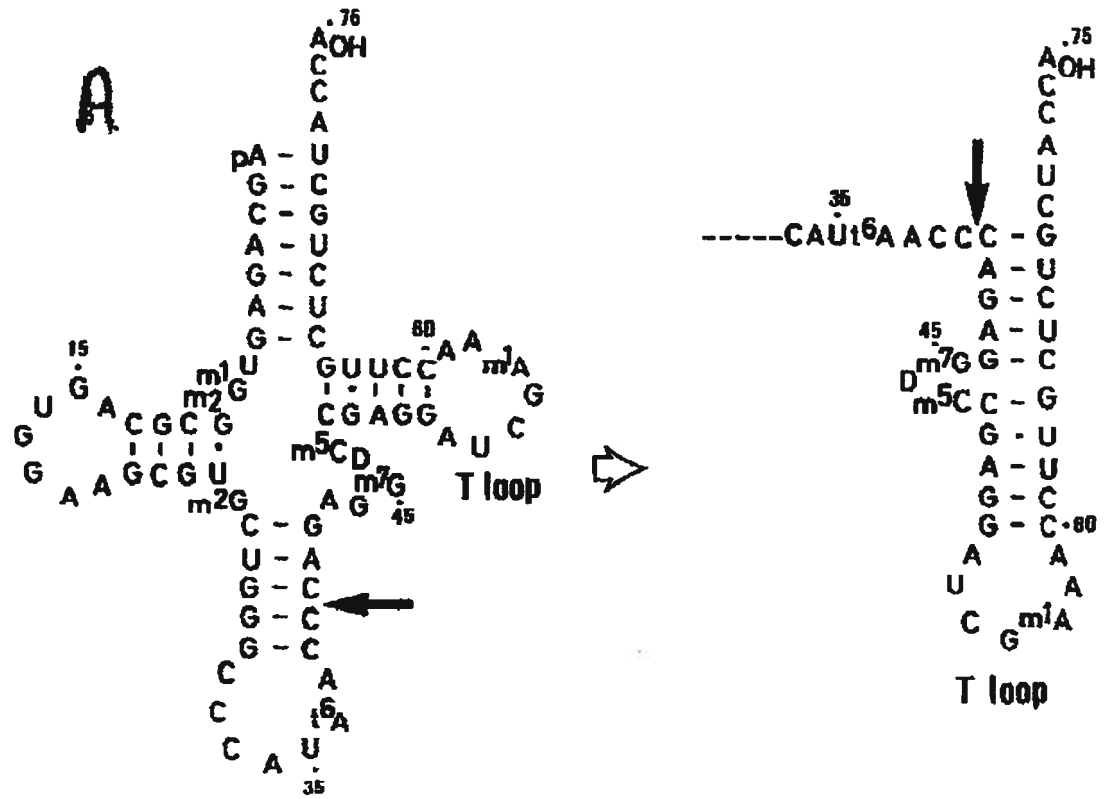
Chez les drosophiles, on retrouve des particules qui ressemblent à des rétrovirus (RVLP) et qui contiennent une molécule d'ARN qui a une taille de 5 kb et certains petits ARN incluant des ARNt. Une activité de transcription inverse a été détectée dans ces particules qu'on pense être des intermédiaires de la transposition d'un élément appartenant à la famille de rétrotransposon *copia* (Kikuchi et al., 1986). En général, la réplication des génomes ARN des rétrovirus débute par le site d'attache de l'amorce où les derniers 18 nucléotides de l'extrémité 3' d'une molécule d'ARNt s'apparient et servent d'amorce pour la reverse transcriptase (Mak et Kleiman, 1997). Chez les particules RVLP des drosophiles (décrites plus haut), une séquence central, située autour de la boucle de l'anticodon de l'ARNt<sup>Met</sup> d'initiation est complémentaire au site d'attache

Figure I.5 : Modèle du clivage de l'ARNtMet initiateur et de la formation de l'amorce pour la réplication des particules ressemblant aux rétrovirus chez la drosophile (Retrovirus-like particles, RVLP).

A. Le changement de conformation proposé pour l'ARNtMet d'initiation. Les flèches en gras indique le site du clivage par la RNase P afin de produire l'amorce requise pour la reverse transcription du rétrotransposon *copia*.

B. Modèle proposé pour la formation de l'amorce requise pour la reverse transcription du rétrotransposon *copia*. La région qui correspond à cette amorce est montrée en gras.

Cette figure a été tirée de l'article (Kikuchi et al., 1990).





de l'amorce de la reverse transcription et non pas l'extrémité 3' de l'ARNt mature. Il a été montré que dans ce cas, l'ARNt en question peut assumer une conformation alternative et peut être coupé par la RNase P pour générer l'amorce requise pour la réplication (Kikuchi, 1995; Kikuchi et al., 1990) (Figure I.5).

Le mécanisme que l'IRES utilise pour initier la synthèse protéique reste inconnu, mais certains éléments structuraux sont bien conservés dans ce site, entre autre la similarité structurale aux ARNt. Cette ressemblance pourrait être l'élément clé pour la reconnaissance des IRES par les ribosomes et le déclenchement de la synthèse protéique (Lyons et Robertson, 2003). En effectuant des tests *in vitro*, dans le but de trouver des cibles potentielles pour la thérapie génique utilisant la RNase P humaine, des chercheurs ont montré que cette dernière peut reconnaître et cliver le génome du virus de l'hépatite C à deux endroits (Nadal et al., 2002). Le premier est au niveau de l'IRES, et le deuxième est situé à l'intérieur du génome, entre la région qui code les gènes structuraux et les gènes non structuraux. Cependant, il n'existe aucun argument qui supporte un rôle pour ces clivages dans la physiologie des virus, puisque pour l'instant, il n'a pas été montré que ces clivage ont lieu *in vivo*, ni que le virus de l'hépatite C accède au noyau (Bartenschlager et Lohmann, 2000), là où la majorité de la RNase P nucléaire réside, donc il se peut que ce clivage sois le résultat du simple fait que la région ressemble structurellement à un ARNt, une propriété qui serait utilisée pour accéder aux ribosomes.

#### **I.2.2.2.2 La RNase P clive les virus des plantes en 3'**

Chez les plantes, la plupart des génomes ARN simple brin des virus contiennent des structures en forme d'hélices et de pseudo-noeuds proches de leur extrémité 3' (Pleij et al., 1987), qui forment une structure en forme de « L » similaire à celle des ARNt. Ces

éléments structuraux peuvent être aminoacylés par l'ajout de la valine pour les tymovirus et les furovirus, la tyrosine pour les bromovirus, cucumovirus et hordeivirus, et par l'histidine pour les tobamovirus (voir références dans (Wientges et al., 2000)). En plus, ces structures peuvent être clivés par la RNase P *in vitro* (Guerrier-Takada et al., 1988), servir de substrat pour l'ajout du CCA par la nucleotidyl-transférase, et lier les facteurs d'élongation de la traduction EF-Tu et EF-1 $\alpha$  (Wientges et al., 2000). Initialement, il a été proposé que ces structures sont des reliques du monde ARN, où ils servaient dans la réplication des virus ARN (Weiner et Maizels, 1987). Récemment, il a été montré que ces éléments agissent comme des « enhancer » dans la traduction (Matsuda et Dreher, 2004).

### **I.3. Reconnaissance du substrat**

Comme décrit précédemment, plusieurs substrats sont reconnus par la RNase P. Ceci nous mène à se poser la question comment la RNase P reconnaît-elle ces substrats? Les études expérimentales, essentiellement menées chez les eubactéries, ont montré que les deux composantes, protéique et ARN, de la ribonucléoprotéine sont impliquées dans la reconnaissance du substrat. Dans ce qui suit, je résumerai ce qui est connu quant aux interactions entre un pré-ARNt et chacune de ces composantes, pour ensuite décrire un mécanisme de reconnaissance qui englobe le rôle de chacune de celles-ci.

#### **I.3.1 Les interactions entre le pré-ARNt et les composantes de la RNase P**

Chez *E. coli*, les études de pontage, mutagenèse, et interférence chimique, effectuées sur l'ARN-P ont montré que le domaine « S », appelé domaine de spécificité, et qui est constitué de la région s'étendant le P7 au P12 (en bleu dans la Figure I.6) (Harris et Christian, 2003), rentre en contact avec la tige et la boucle T du pré-ARNt (Knap et al., 1990; LaGrandeur et al., 1994; Nolan et al., 1993) (Loria et Pan, 1997; Pan et al., 1995).

Le domaine « C », appelé domaine catalytique, qui est formé des hélices P1-P5, et P15 (en vert dans la Figure I.6), interagit avec le motif NCCA en 3' des pré-ARNt et avec le nucléotide en position -1 (en amont) du site de coupure (Kirsebom et Svard, 1994; LaGrandeur et al., 1994; Oh et al., 1998).

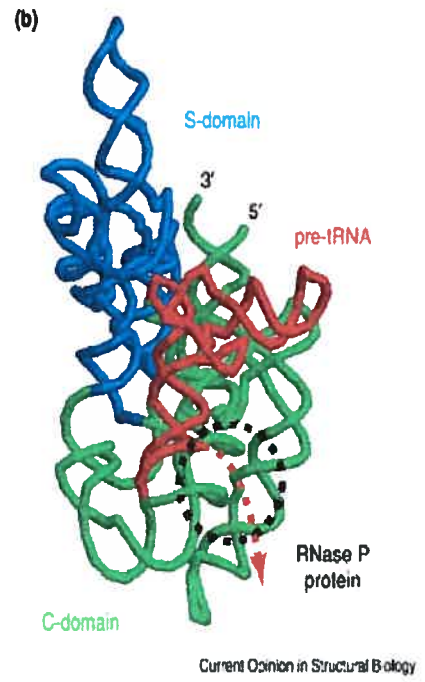
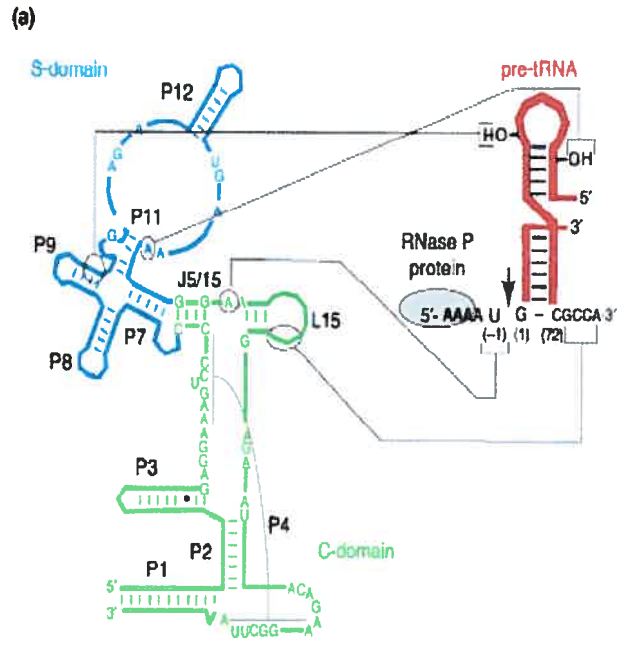
Quant à la protéine-P, la comparaison des constantes cinétiques et d'affinité au substrat de l'holoenzyme et de l'ARN-P seul, indique qu'elle interagit avec les pré-ARNt, au niveau de la région guide en 5' des ARNt (5' leader sequence) (Crary et al., 1998; Kurz et al., 1998). Les études de pontage ont montré que cette dernière s'attache à une fente sur la surface de la composante protéique (Niranjanakumari et al., 1998).

Figure I.6 : Structure globale du complexe RNaseP/pré-ARNt d'*E. coli*.

(a) Consensus de la structure secondaire des ARN-P eubactériens (Hall et Brown, 2001). La lettre P est utilisée pour désigner une hélice, J pour la région en simple brin qui joint deux hélices. Le domaine « S » de spécificité est en bleu et le domaine « C » de catalyse en vert. La tige acceptrice ainsi que la boucle et l'hélice T sont en rouge. La région guide en 5' des ARNt (5' leader sequence) est en noir, et l'interaction de cette région avec la protéine-P est montrée en forme d'ovale gris. La flèche indique le site de coupure par la RNase P. Les lignes indiquent une interaction entre les sites aux extrémités de ces lignes. Les lettres en majuscule sur le modèle consensus représentent les nucléotides conservés universellement.

(b) Modèle de la structure tertiaire du complexe RNaseP/pré-ARNt (Chen et al., 1998; Massire et al., 1998). Le code de couleur est le même que dans (a). Les positions probables de la protéine-P et de la région guide en 5' de l'ARNt sont marquées respectivement par des lignes coupées noire et rouge.

Cette figure a été tirée de (Harris et Christian, 2003).



### **I.3.2 Contribution modulaire à la reconnaissance du pré-ARNt**

Les expériences de mutagenèse dirigée ont permis d'élucider la contribution de chaque module à la reconnaissance du substrat. Elles ont dévoilé qu'en absence du domaine « S », le domaine « C » de l'ARN-P et la composante protéique peuvent interagir, lier le substrat et accomplir la réaction catalytique, aussi efficacement pour un pré-ARNt que pour un substrat qui consiste en une région simple brin suivie d'une hélice courte (Loria et Pan, 2001). Ces expériences, ainsi que les études enzymologiques et structurales supportent le modèle mécanistique de reconnaissance du substrat selon lequel la protéine-P interagit avec la région guide en 5' de l'ARNt (Christian et al., 2002b), et qu'elle augmente le taux de catalyse en favorisant l'attache du pré-ARNt (le substrat) à la ribonucléoprotéine et la relâche de l'ARNt mature (le produit) (Harris et Christian, 2003). Quant au domaine « C », il contient le site de catalyse et rentre en contact avec le pré-ARNt au niveau du site de coupure sans montrer aucune spécificité; cette dernière sera octroyée par le domaine de spécificité « S », qui lie la boucle et l'hélice T, et qui est responsable de la différenciation entre un ARNt et les autres substrats.

### **I.3.3 Reconnaissance des substrats autres que les pré-ARNt**

Le modèle précédent explique comment un pré-ARNt aussi bien qu'un pré-ARNtm (précurseur de l'ARNtm) est reconnu par la RNase P, puisque la structure des ARNtm a conservé la tige acceptrice, le motif RCCA, ainsi que la boucle et l'hélice T (Williams et Bartel, 1996). Mais comment les autres substrats sont-ils reconnus? Présentement, l'état des connaissances ne permet pas de répondre à cette question d'une façon détaillée, mais deux éléments clés sont à signaler. Premièrement, la RNase P n'a pas besoin des boucle et tige T pour reconnaître un substrat. Les expériences de mutagenèse ont montré que le substrat minimal pour la RNase P d'*E. coli* peut se limiter à un ARN qui contient une hélice de 4 à 16 pb qui mime la tige acceptrice, et d'une région simple brin en 5' formée d'au moins un nucléotide (Forster et Altman, 1990a; Kirsebom et Vioque, 1995; Liu et Altman, 1996); pour une coupure efficace par l'ARN-

P seul *in vitro*, le motif CCA est nécessaire. Deuxièmement, la reconnaissance des substrats se fait au niveau de la structure tertiaire, et à ce niveau, il a été montré que la protéine-P permet à l'ARN d'accommoder de nouveaux substrats plus efficacement, probablement en adaptant la structure générale du ribozyme (Kirsebom et Vioque, 1995; Liu et Altman, 1994; Vioque et al., 1988).

#### **1.3.4 Reconnaissance du substrat chez les eucaryotes et les archéobactéries**

L'information présentée dans la section précédente provient des eubactéries, et essentiellement d'*E. coli* et *B. subtilis*; et elle ne décrit pas nécessairement le mode de reconnaissance du substrat dans les autres règnes. Chez les eucaryotes, nous observons certaines différences à ce niveau. Premièrement, le CCA en 3' terminal des ARNt n'est pas encodé dans le génome des eucaryotes comme chez *E. coli*, mais plutôt rajouté après la maturation des pré-ARNt. Deuxièmement, chez les humains, le substrat minimal de la RNase P doit obligatoirement contenir la tige et la boucle T (Yuan et Altman, 1995); en plus de la tige acceptrice et de la région simple brin en 5'. La nécessité de la présence de l'hélice T pour le clivage par la RNase P eucaryote la rend moins versatile que la RNase P eubactérienne; biologiquement, ceci se traduit par l'observation que son seul substrat connu est le pré-ARNt. Très peu d'information existe sur la reconnaissance du substrat par la RNase P eucaryote à part ce qui a été présenté plus haut. Ceci est la conséquence du fait que les ARN-P eucaryotes seuls ne sont pas actifs, et que l'association des composantes protéiques et ARN *in vitro* n'a pas réussi à reproduire une activité catalytique.

Les archéobactéries ont une composante ARN qui ressemble à celle des eubactéries, mais une composante protéique qui ressemble à celle des eucaryotes (Hall et Brown, 2002). Pour l'instant aucune étude ne s'est attardée à la reconnaissance des substrats chez les archéobactéries, en partie, parce que les premiers ARN-P archéobactériens testés n'étaient pas actifs *in vitro*. Mais récemment, les ARN-P des méthanobactéries, thermococcis, et halobactériades se sont révélés actifs *in vitro* en absence de la composante protéique (Pannucci et al., 1999). De plus, un groupe de

chercheurs a reconstitué une activité catalytique, en associant quatre composantes de la sous-unité protéique et l'ARN-P de l'archéobactérie hyper thermophile *Pyrococcus horikoshii* (Kouzuma et al., 2003), ce qui ouvre la porte à de nouvelles études dans l'avenir sur la reconnaissance du substrat par cette RNase P.



## **I.4. Un pseudo-nœud universel chez la RNase P**

Comme décrit dans la section 1 de cette introduction, ce qui rassemble les RNase P de tous les règnes c'est l'ARN-P et plus précisément, le site catalytique formé par l'hélice P4 qui se replie en un pseudo-nœud (Chen et Pace, 1997; Forster et Altman, 1990b). Cette structure a été identifiée chez tous les ARN-P connus sans exception, et son importance dans la catalyse a été documentée. Dans ce qui suit, je résumerai ce qui est connu sur l'implication de cette région dans la réaction de catalyse.

### **I.4.1 Site d'attache des ions $Mg^{2+}$**

La conservation universelle du pseudo-nœud formé du P4 a attiré l'attention des chercheurs, puisqu'elle laissait suspecter un rôle biologique central. Cette observation a été renforcée par le fait que les mutants qui manquent la région P1-P4 sont inactifs, alors que la délétion des autres hélices est moins critique (Schlegl et al., 1994).

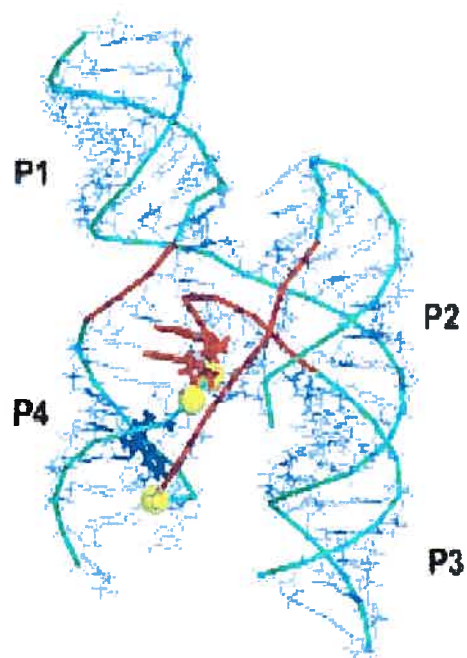
Les études de pontage ont montré que les nucléotides du P4 et de la région simple brin avoisinante sont très proche ( $<3 \text{ \AA}$ ) du site de coupure du substrat (Christian et Harris, 1999; Christian et al., 1998; Kufel et Kirsebom, 1996). La majorité des groupes fonctionnels importants pour la réaction de catalyse s'y trouvent (Harris et Pace, 1995b; Kaye et al., 2002a; Kazantsev et Pace, 1998), notamment les atomes d'oxygène des groupes phosphates des nucléotides A67, G68, et A352 chez l'ARN-P d'*E. coli*. Le remplacement de ces derniers par des atomes de soufre baisse le taux de catalyse de 1000 à 10 000 fois (Christian et al., 2000; Harris et Pace, 1995b)(Figure I.7). L'effet de ces changements peut être renversé en utilisant des ions métalliques thiophiliques ( $Mn^{2+}$  ou  $Cd^{2+}$ )(Christian et al., 2002a; Hardt et al., 1995; Harris et Pace, 1995b), ce qui montre que l'interaction des ions  $Mg^{2+}$  avec le P4 est cruciale pour la catalyse. La nature des nucléotides dans le P4 influence la spécificité des ions impliqués dans la catalyse; par exemple, il a été montré qu'échanger le C70 par un U cause une augmentation de la capacité du ribozyme à utiliser le  $Ca^{2+}$  (Frank et Pace, 1997). Dans le P4, on retrouve aussi un bulge universellement conservé (U69) (Figure I.7), les études de mutagenèse ont montré que ce bulge est important pour le maintien d'une géométrie favorable à l'attache des ions métalliques (Kaye et al., 2002b; Schmitz et Tinoco, 2000).

Figure I.7 : Région d'interaction des ions métalliques avec la région du P1-P4

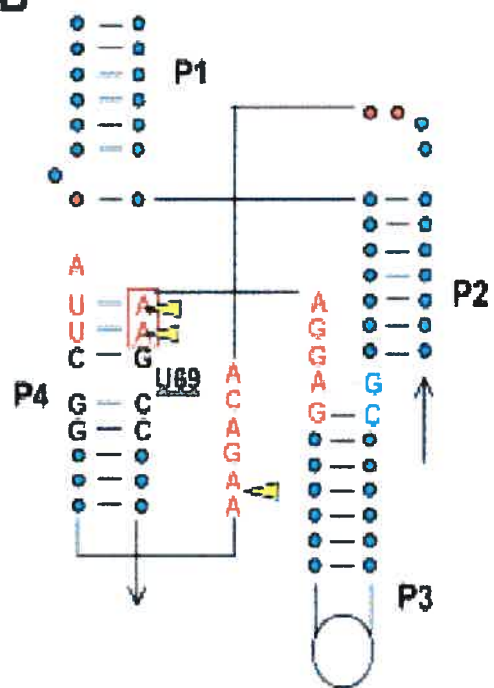
- (A) Modèle de structure tertiaire de la région P1-P4 (Chen et al., 1998; Massire et al., 1998). Les nucléotides U69, A67 et A66 sont en bleu foncé et rouge respectivement. Les groupes fonctionnels sensibles au changement des atomes d'oxygène en soufre sont en forme de sphères jaunes.
- (B) Structure secondaire de la jonction des hélices P1-P4. Ce modèle montre l'empilement (stacking) des quatre hélices. Les nucléotides conservés universellement sont en lettres majuscules. Les sites sensibles aux changements en phosphothiorate sont marqués par des flèches jaunes.

Cette figure a été tirée de (Kaye et al., 2002b).

A



B



### I.4.2 La réaction de catalyse

L'attache des ions métalliques au P4 et à la région avoisinante est la clé de la réaction de catalyse, car ces ions vont catalyser la réaction de clivage pour produire un ARNt mature avec un groupe phosphate en 5', et un groupe hydroxyle en 3' de la région guide en 5' des ARNt (Guerrier-Takada et al., 1986). Le cation divalent le plus efficace pour accomplir cette réaction est le magnésium suivi du manganèse, et du calcium (Smith et al., 1992). Ces cations ionisent une molécule d'eau pour produire un ion hydroxyle qui s'attaque au lien phosphodiester du substrat par un mécanisme de substitution nucléophile concertée (SN2) (Cassano et al., 2004; Smith et Pace, 1993) (Figure I.8). Les autres rôles de ces cations sont d'induire la catalyse en interagissant simultanément avec le phosphomonoester et le nucléophile durant l'état de transition, en stabilisant la charge négative sur le groupe partant (Cassano et al., 2004), et en orientant le groupe hydroxyle d'une manière à faciliter la réaction chimique (Chen et al., 1997) (Figure I.8).

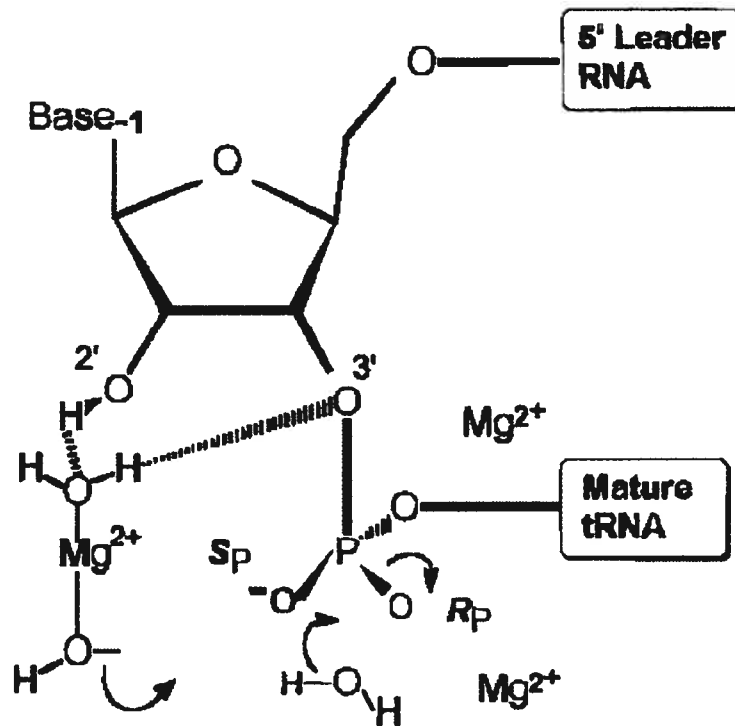
La conservation universelle du P4, l'observation que l'ARN-P soit nécessaire pour l'activité de la RNase P chez les eucaryotes et les archéobactéries (Cherayil et al., 1987; Franklin et al., 1995; Lee et al., 1991; Waugh et Pace, 1990), et le fait que le magnésium est requis pour l'activité catalytique de la RNase P dans tous les règnes suggèrent que cette région assure le même rôle chez tous les ARN-P, et que fort probablement le mécanisme de catalyse par ces derniers ressemble à ce qui a été présenté plus haut (Figure I.8).

Figure I.8 : Mécanisme de la catalyse par la RNase P basé sur les modèles proposés dans (Guerrier-Takada et al., 1986) et (Smith et Pace, 1993).

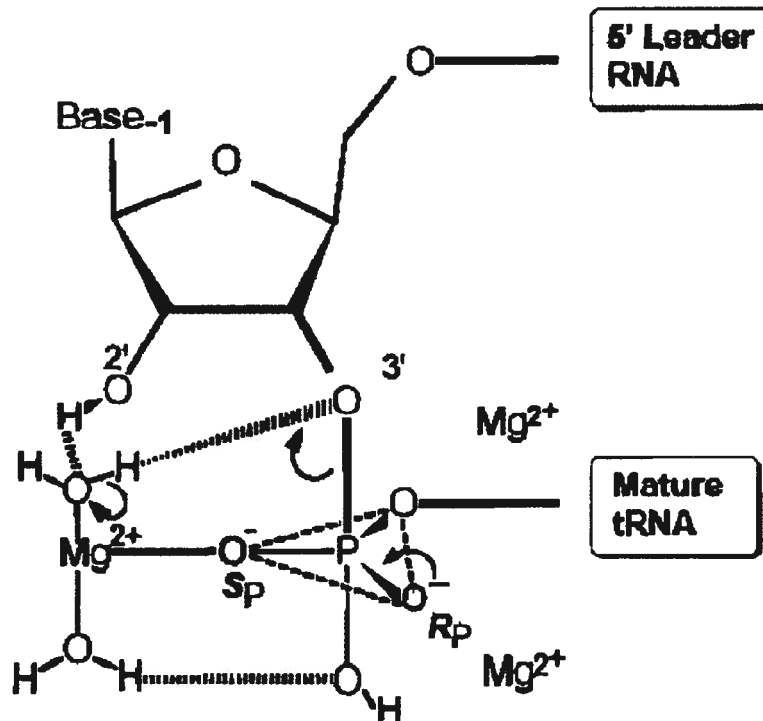
- (A) État de pré-transition. Le groupe  $[\text{Mg}(\text{OH})]^+$ , fait une coordination indirecte avec le 2' OH en amont du site de coupure, et ionise une molécule d'eau.
- (B) État de transition. L'oxygène pro-S<sub>P</sub> est en coordination avec l'hydrate de magnésium. Deux ions  $\text{Mg}^{2+}$  stabilisent l'oxygène pro-R<sub>P</sub> et O3' pendant le clivage du lien O3'-P.

Cette figure a été tirée de (Chen et al., 1997).

## A. Pre-Transition State



## B. Transition State



## I.5. La RNase P mitochondriale

Le pseudo-nœud formé du P4 est reconnu comme une structure de base de l'ARN-P, et dans certains ARN-P mitochondriaux (ARN-Pmt) c'est le seul motif conservé (voir plus bas). Les ARN-Pmt sont parmi les ARN-P les plus divergents. Ils sont très hétérogènes en taille et en structure secondaire, ce qui rend l'étude de leur évolution d'un intérêt particulier. Dans la section qui suit, je résumerai ce qui est connu concernant l'origine de la RNase P mitochondriale, sa composition, la distribution du gène qui encode la sous-unité ARN, et l'impact de la génomique sur ce domaine.

### I.5.1 Origine de la mitochondrie

Au début des années 70, Margulis a fait renaître une hypothèse qui existait depuis plus d'un siècle (voir les références dans (Lang et al., 1999)), et qui propose que la mitochondrie soit le résultat d'une endosymbiose entre une cellule eucaryote, et une eubactérie (Margulis, 1970). Les facteurs qui ont amorcé cette endosymbiose (échange de métabolites,...)(Martin et Müller, 1998), la nature de la cellule hôte (eucaryote primitif, ou fusion d'une eubactérie avec une archéobactérie)(Moreira et Lopez-Garcia, 1998), et les conditions dans lesquelles cet événement a eu lieu sont encore débattus (Gray et al., 1999), mais le support pour l'origine  $\alpha$ -protéobactérienne de la mitochondrie ne fait que grandir au fur et à mesure que les séquences des génomes mitochondriaux et eubactériens deviennent disponibles (Gray et Spencer, 1996; Gray, 1989). Cette conclusion est basée sur l'analyse des gènes codés dans les génomes mitochondriaux (Gray et al., 1999). Il y a même eu des tentatives pour chercher le sous-groupe à l'intérieur des  $\alpha$ -protéobactéries auquel s'apparente l'ancêtre de la mitochondrie. Ces tentatives montrent que ce dernier peut être aussi proche du groupe des *Rickettsia* que celui des *Rhizobium* ou des *Rhodospirillum* (Figure I.9), cependant le nombre de génomes  $\alpha$ -protéobactériens et mitochondriaux complètement séquencés présentement ne permet pas de clore cette question (Lang et al., 1999).

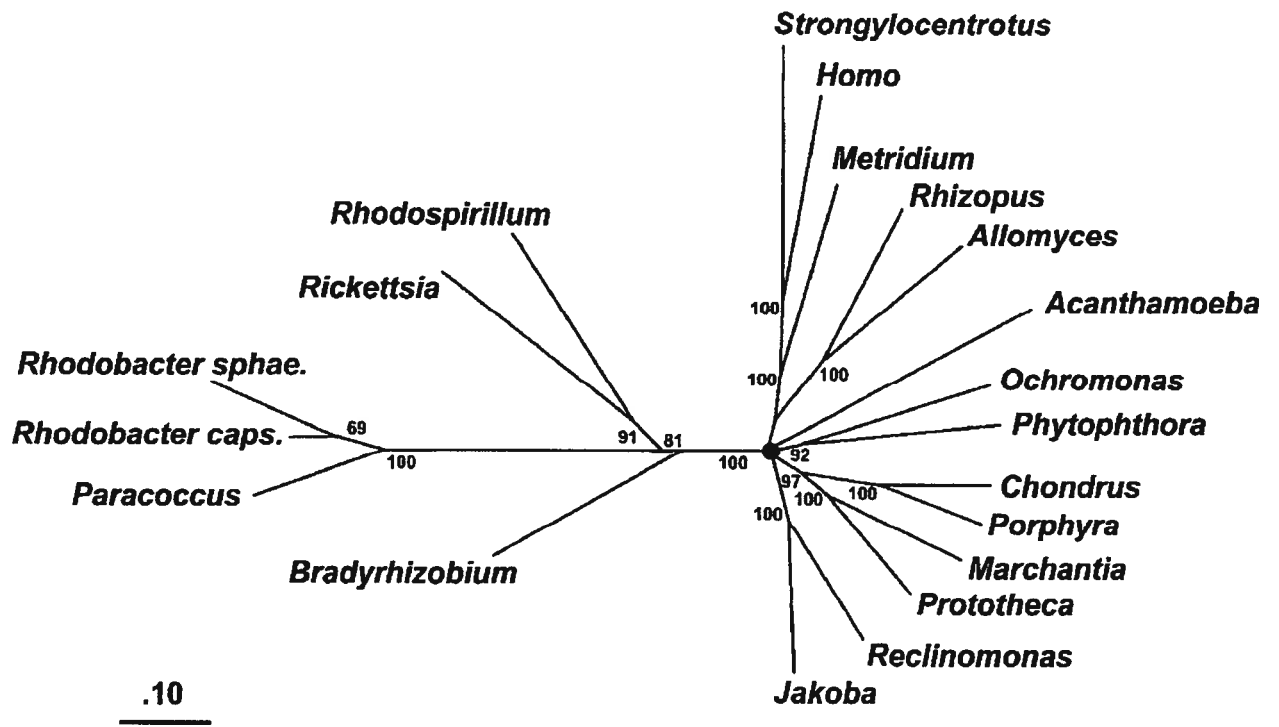
Dans ce contexte, l'origine de la RNase P mitochondriale devrait être la même que celle de la mitochondrie, donc eubactérienne; à moins que les composantes de l'activité de la RNase P mitochondriale proviennent du noyau de la cellule hôte.



Figure I.9 : Phylogénie des mitochondries et des  $\alpha$ -protéobactéries

L'analyse a été faite avec les séquences protéiques concaténées des gènes suivants : apocytochrome *b* (Cob), et cytochrome oxidase, sous-unités 1 à 3 (Cox1-3). Les taxons inclus représentent les groupes phylogénétiques majeurs des eucaryotes et toutes les séquences disponibles des  $\alpha$ -protéobactéries. Les analyses phylogénétiques ont été faites en utilisant PROTDIST/FITCH, qui permet une correction Jin/Nei pour les sites avec des taux de changements variables. Le coefficient de variation utilisé est 0.5. Le cercle noir indique des branches courtes fusionnées qui ne montrent pas de résolution (support par boot strap  $\leq 60\%$ ). L'échelle montre la moyenne de nombre de substitutions par site. La topologie présentée est supportée par une analyse de maximum de vraisemblance. Les organismes et séquences (numéro d'accèsion dans GenBank entre parenthèses) sont *Strongylocentrotus purpuratus* (oursin de mer; X12631), *Paracoccus denitrificans* (X05829, M17522, X05934, X05828), *R. prowazekii* (AJ235270 to AJ235273), *Rhodobacter sphaeroides* [X56157, X62645, M57680, C45164 (proteine)], *Rhodospirillum rubrum* (X55387), *Rhodobacter capsulatus* (X05630), *Bradyrhizobium japonicum* (J03176, X68547); le reste des séquences sont disponibles dans (Lang et al., 1999).

Cette figure a été tirée de (Lang et al., 1999).



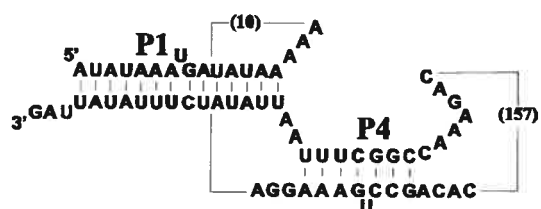
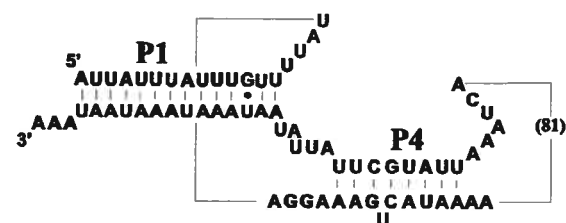
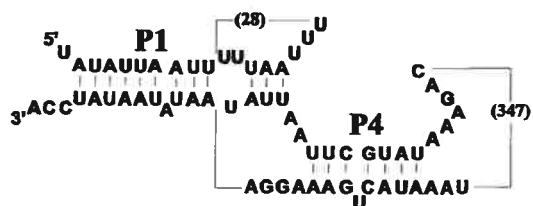
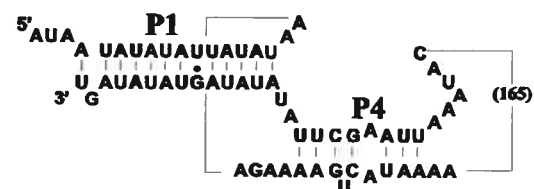
### I.5.2 La RNase P mitochondriale chez les champignons

La première RNase P mitochondriale qui a été identifiée est celle de *S. cerevisiae* (Martin et al., 1983; Underbrink-Lyon et al., 1983). Dans ce cas, l'ARN-Pmt est codée dans le génome mitochondrial. Sa séquence primaire et sa structure secondaire ne ressemblent pas à celles de l'ARN-P nucléaire, et l'activité mitochondriale est indépendante de celle du noyau, ce qui laisse croire que l'origine de l'ARN-Pmt chez cette levure est fort probablement eubactérienne et non pas nucléaire. La composante protéique est codée par le génome nucléaire et n'est similaire à aucune protéine-P eubactérienne ou nucléaire connue; elle est dix fois plus grande que la protéine-P d'*E.coli* (Dang et Martin, 1993; Morales et al., 1992). Les études comparatives de séquences ont permis d'identifier plusieurs nouveaux gènes *rnpB* codés dans les mitochondries de levures phylogénétiquement voisines de *S. cerevisiae* (voir article 1). L'activité de la RNase P mitochondriale a été étudiée aussi chez l'ascomycète filamenteux *Aspergillus nidulans*, où une ARN-Pmt a été identifiée et purifiée (Lee et al., 1996b). Quant à la composante protéique, elle reste toujours inconnue mais sept protéines se retrouvent purifiées avec l'activité catalytique. Alors une question se pose, est-ce qu'il s'agit de sept sous-unités protéiques ? Est-ce que certaines de ces protéines sont simplement des impuretés ? Pour l'instant aucune réponse n'a été donnée.

Contrairement à ce qu'on observe chez les eubactéries, les ARN-Pmt de *S. cerevisiae* et d'*Aspergillus nidulans* ne sont pas catalytiques *in vitro*, mais ils sont tous les deux essentiels pour l'activité catalytique de la RNase P mitochondriale (Lee et al., 1996c; Underbrink-Lyon et al., 1983). Les ARN-Pmt chez les champignons se caractérisent par des séquences riches en A et U, et une taille qui varie entre 140 nts chez *Saccharomycopsis fibuligera* et 423 nts chez *S. cerevisiae* (Stribinskis et al., 1996), ce qui rend leur repliement en structure secondaire par comparaison phylogénétique une tâche très ardue. Les modèles de structure secondaire qui existaient pour ces ARN avant le début de mon projet se limitaient aux P1 et P4 (Figure I.10).

Figure I.10 : Modèle de structure secondaire des ARN-P mitochondriaux de *S. cerevisiae*, *S. fibuligera*, *Candida glabrata*, et *A. nidulans*.

Ces modèles ont été établis par comparaison. Cette figure a été tirée de (Lee et al., 1996c).

*Aspergillus nidulans**Saccharomycopsis fibuligera**Saccharomyces cerevisiae**Candida glabrata*

Lee, Y.C., Lee, B.J. et Kang, H.S. (1996) Eur. J. Biochem. Vol. 235:297-303.

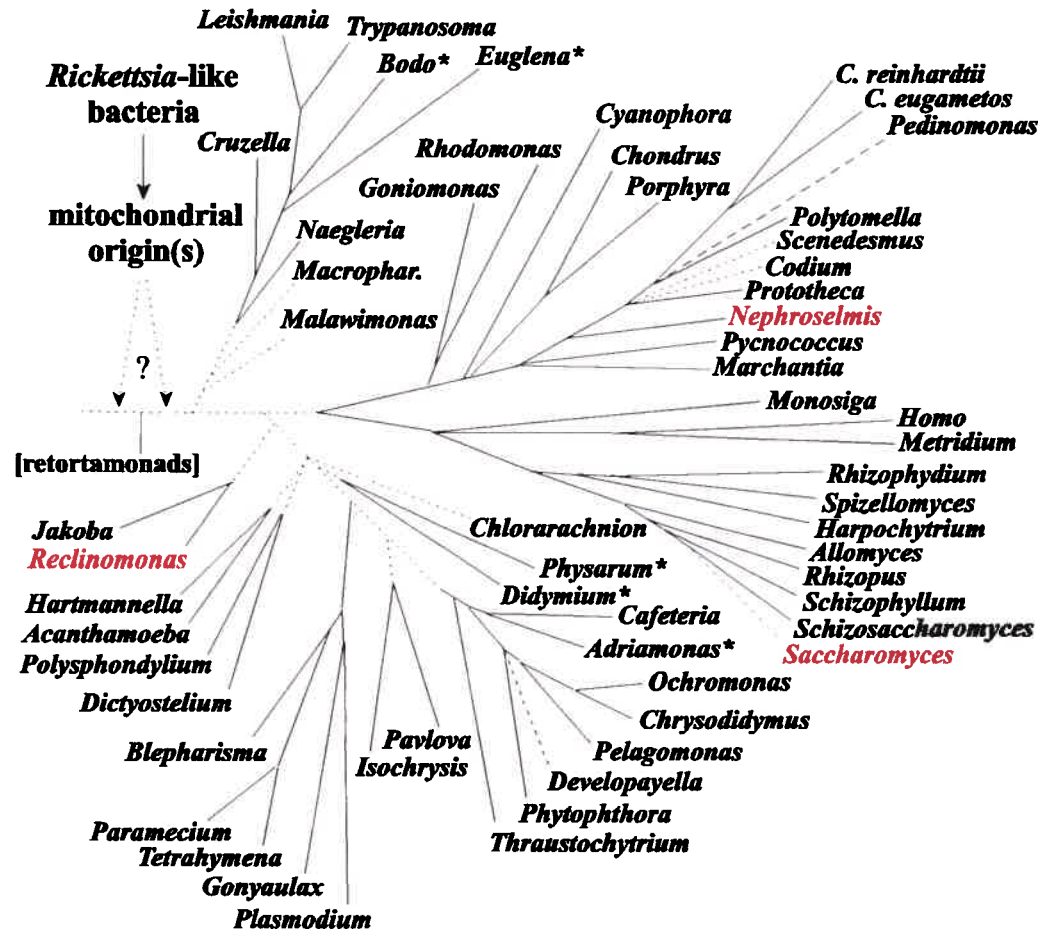
### I.5.3 L'impact de la génomique

Dans les dernières années, les techniques de séquençage et d'analyse de séquences ont évolué d'une manière révolutionnaire et ont permis d'obtenir la séquence complète d'un grand nombre de génomes mitochondriaux. Dans le domaine de la génomique mitochondrial, l'impact de cette percée technologique s'est traduit par la découverte d'un génome mitochondrial conservé chez le jakobide *Reclinomonas americana* ressemblant à un génome eubactérien en miniature (Lang et al., 1997). Ce génome contient le plus grand nombre de gènes codés par une mitochondrie, dont plusieurs n'ont jamais été retrouvés chez une mitochondrie précédemment comme les gènes encodant les sous-unités d'une ARN-polymérase eubactérienne (*rpoA-D*), plusieurs gènes de protéines ribosomiques (*rpl1, 10, 18, 19, 27, 31, 32*), de protéines impliquées dans le transport membranaire (*secY, yejW*), facteur d'élongation (*tufA*) et de composantes de la chaîne respiratoire (*atp3, cox11*). En plus, il encode une ARN-Pmt qui contient tous les éléments structuraux présents dans le modèle du consensus minimal de la structure secondaire des ARN-P eubactériens (voir Figure 2A de l'article 1). Cette observation indique que l'ARN-Pmt chez les jakobides a été hérité de l'endosymbiote eubactérien. Cette trouvaille a été suivie par l'identification d'un ARN-Pmt codé par le génome mitochondrial de l'algue verte *Nephroselmis olivacea*, qui lui aussi ressemble aux ARN-P eubactériens (Turmel et al., 1999).

Cette percée technologique a aussi dévoilé que la plupart des génomes mitochondriaux séquencés n'encodent pas de gène *rnpB* (Figure I.11). En fait, seuls les jakobides et l'algue verte *N. olivacea* encodent un, à l'extérieur du règne des champignons. Et même à l'intérieur de ce règne, la distribution de ce gène est très sporadique, aucun basidiomycète ou chytridiomycète ne l'encode. Dans les génomes mitochondriaux complètement séquencés des ascomycètes, seul *A. nidulans* et quelques levures encodent un gène *rnpB* alors qu'il est absent chez la majorité.

Figure I.11: Distribution du gène *rnpB* chez les eucaryotes

Phylogénie hypothétique des eucaryotes, basée sur des observations morphologiques et des analyses phylogénétiques. Les branches pointillées de cet arbre sont basées sur des hypothèses déduites de données morphologiques, alors que les lignes pleines sont basées sur des analyses phylogénétiques appliquées aux séquences protéiques de gènes mitochondriaux. L'origine de la mitochondrie est montrée par deux flèches. Les espèces qui contiennent les gènes *rnpB* dans leur génome mitochondrial sont en rouge. Cette figure est une version modifiée de l'arbre présenté dans l'article (Gray et al., 1998).





## I.6. Définition du projet

Mon projet a débuté afin de mieux comprendre l'évolution de l'ARN-Pmt. Mes buts étaient (i) de trouver de nouveaux gènes *rnpB* dans les génomes mitochondriaux déjà séquencés en se basant sur les régions consensus identifiées chez l'ARN-Pmt de *R. americana*, (ii) établir des modèles de structure secondaire pour tous les ARN-Pmt par comparaison phylogénétique d'une façon systématique afin de mieux comprendre l'évolution de cette structure, (iii) tester l'activité catalytique *in vitro* de l'ARN-Pmt de *R. americana* et (iv) de faire des liens entre la structure secondaire et l'activité catalytique de ces ARN.

Dans le chapitre suivant, je présenterai mes résultats concernant la distribution du gène encodant la sous-unité ARN de la RNase P mitochondriale (*rnpB*)(article1), l'évolution de la structure secondaire de la composante ARN chez les champignons (articles 1 et 2), et les jakobides (article 3), ainsi que l'effet de cette évolution sur l'activité catalytique des ARN-P (article 3).

## **Chapitre II : Articles**

## Article 1

### **Mitochondrial RNase P RNAs in ascomycete fungi: lineage-specific variations in RNA secondary structure.**

**Elias R. Seif, Lise Forget, Nancy C. Martin<sup>1</sup> and B. Franz Lang\***

Program in Evolutionary Biology, Canadian Institute for Advanced Research, Département de Biochimie, Université de Montréal, 2900 Boul. Edouard-Montpetit, Montréal, Québec, H3T 1J4, Canada.

<sup>1</sup> University of Louisville, Department of Biochemistry and Molecular Biology, School of Medicine, University of Louisville, Louisville, KY 40292-0001, USA.

*RNA*, 2003, vol.9, no. 9, pp. 1073-1083.

#### Contribution of each author :

Elias Seif : did most of the experimental work, results analysis, and preparation of the manuscript and figures.

Lise Forget : helped in the experimental work.

Nancy C. Martin : participated in manuscript preparation, and supplied experimental data.

B. Franz Lang : participated in manuscript and figure preparation.

# Mitochondrial RNase P RNAs in ascomycete fungi: Lineage-specific variations in RNA secondary structure

ELIAS R. SEIF,<sup>1</sup> LISE FORGET,<sup>1</sup> NANCY C. MARTIN,<sup>2</sup> and B. FRANZ LANG<sup>1</sup>

<sup>1</sup>Program in Evolutionary Biology, Canadian Institute for Advanced Research, Département de Biochimie, Université de Montréal, Montréal, Québec H3T 1J4, Canada

<sup>2</sup>University of Louisville, Department of Biochemistry and Molecular Biology, School of Medicine, University of Louisville, Louisville, Kentucky 40292-0001, USA

## ABSTRACT

The RNA subunit of mitochondrial RNase P (mtP-RNA) is encoded by a mitochondrial gene (*rnpB*) in several ascomycete fungi and in the protists *Reclinomonas americana* and *Nephroselmis olivacea*. By searching for universally conserved structural elements, we have identified previously unknown *rnpB* genes in the mitochondrial DNAs (mtDNAs) of two fission yeasts, *Schizosaccharomyces pombe* and *Schizosaccharomyces octosporus*; in the budding yeast *Pichia canadensis*; and in the archi-ascomycete *Taphrina deformans*. The expression of mtP-RNAs of the predicted size was experimentally confirmed in the two fission yeasts, and their precise 5' and 3' ends were determined by sequencing of cDNAs generated from circularized mtP-RNAs. Comparative RNA secondary structure modeling shows that in contrast to mtP-RNAs of the two protists *R. americana* and *N. olivacea*, those of ascomycete fungi all have highly reduced secondary structures. In certain budding yeasts, such as *Saccharomycopsis fibuligera*, we find only the two most conserved pairings, P1 and P4. A P18 pairing is conserved in *Saccharomyces cerevisiae* and its close relatives, whereas nearly half of the minimum bacterial consensus structure is retained in the RNAs of fission yeasts, *Aspergillus nidulans* and *Taphrina deformans*. The evolutionary implications of the reduction of mtP-RNA structures in ascomycetes will be discussed.

**Keywords:** Evolution; phylogenetic modeling; RNA processing; organelle; *rnpB*; yeasts

## INTRODUCTION

RNase P is a ribonucleoprotein (for exceptions, see below) that is universally present in eubacteria, archaeobacteria, and eukaryotes, as well as in mitochondria and chloroplasts. It participates in the processing of tRNAs, 4.5 S RNAs, and other small RNAs, by endonucleolytic removal of 5' leader sequences from RNA precursors (Peck-Miller and Altman 1991).

The RNA subunits of eubacterial RNase P (P-RNA) have been intensely studied and are currently best understood. *Escherichia coli* P-RNA is essential for the enzymatic activity in vivo and is the only subunit necessary for activity in vitro (Stark et al. 1978; Kole and Altman 1981; Guerrier-Takada et al. 1983). Comparisons of all known P-RNAs have revealed significant primary and secondary structure similarities, indicating that these molecules evolved from a com-

mon ancestral RNA. The in vitro activity of P-RNAs in the absence of a protein component has been demonstrated in numerous instances (Guerrier-Takada et al. 1983; Gardiner et al. 1985; Wagner et al. 2001). Nevertheless, the single eubacterial protein component of RNase P (P-protein) is essential for the enzymatic activity in vivo. It has been suggested that the protein participates in formation of an active-site architecture, by interacting with the 5' leader of tRNAs (Crary et al. 1998), and that it increases the catalytic activity of RNase P by acting as an electrostatic shield between the negatively charged P-RNA and tRNAs (Guerrier-Takada et al. 1983; Gardiner et al. 1985; Christian et al. 2002). The P-protein also broadens the substrate specificity of RNase P, by enhancing its affinity for non-tRNA substrates. For example, 4.5 S RNA and C4 RNA precursors are processed more efficiently by a P-RNA/P-protein complex than by P-RNA alone (Peck-Miller and Altman 1991; Hartmann et al. 1995).

The P-RNA secondary structure model of Archaeobacteria is strikingly similar to that of eubacteria, except for the lack of the P13, P14, and P18 helices (Brown and Haas 1995; Brown 1999). Archaeal P-RNAs were long thought to be

**Reprint requests to:** B. Franz Lang, Département de Biochimie, Université de Montréal, 2900 Boul. Edouard-Montpetit, Montréal, Québec, H3T 1J4, Canada; [redacted] fax: (514) 343-2210.

Article and publication are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.5880403>.

inactive without their protein partner (see Brown and Haas 1995). However, the P-RNAs of Methanobacteria, Thermococci, and Halobacteria do have catalytic activity under extreme ionic conditions (Pannucci et al. 1999). In fact, deletion mutants show that the structural elements missing in archaeal P-RNAs are also dispensable for *in vitro* catalysis by the *E. coli* RNA molecule (Darr et al. 1992; Haas et al. 1994). Although archaeobacterial P-RNA structures resemble those of eubacteria, the archaeal holoenzyme is much larger. It has at least four protein subunits that appear to be homologs of the eukaryotic P-proteins (Hall and Brown 2002).

Eukaryotic nuclear P-RNAs have been investigated in detail in ascomycete fungi and in animals. They have no catalytic activity *in vitro* and, as in archaeobacteria, consist of an RNA subunit and several proteins, at least nine in yeast and 10 in human (Gopalan et al. 2002; Guerrier-Takada et al. 2002). A secondary structure model of the eukaryotic P-RNA conforms convincingly to the bacterial consensus structure with only minor deviations (Chen and Pace 1997; Frank et al. 2000). In plants, RNase P has been purified, and a P-RNA component has been suggested based on the sensitivity of the carrot RNase P activity to micrococcal nuclease treatment (Franklin et al. 1995). Although the nuclear RNase P from wheat is resistant to nuclease treatment, its density and its low isoelectric point also indicate the presence of a P-RNA subunit (Arends and Schön 1997). In fact, RNase P from rat liver (Jayanthi and Van Tuyle 1992) and the archaeobacterium *Sulfolobus solfataricus* (Darr et al. 1990) resist micrococcal nuclease treatment but do contain a P-RNA subunit (Altman et al. 1993; Harris et al. 2001). This indicates that the RNA subunit is protected from digestion by P-proteins.

Mitochondria and chloroplasts contain distinct, organelle-specific RNase P activities. The analysis of organellar P-RNAs has been complicated by the patchy occurrence of the *rnpB* gene, both in chloroplast and mitochondrial DNAs (mtDNAs). The chloroplast-encoded P-RNA of *Cyanophora paradoxa* folds into a Cyanobacteria-like secondary structure (in agreement with the cyanobacterial origin of chloroplasts) and is essential for RNase P activity (Baum et al. 1996; Pascual and Vioque 1999). Chloroplast DNA (cpDNA)-encoded *rnpB* genes have been found only in the green alga *Nephroselmis olivacea* (Turmel et al. 1999), the red alga *Porphyra purpurea* (Reith and Munholland 1995), and *Cyanidium caldarium*. We identified the latter, previously unrecognized *rnpB* gene, at positions 99,959–100,305 of GenBank record AF022186, (Glöckner et al. 2000). All other known cpDNAs do not seem to encode a P-RNA. A recent prediction of a maize chloroplast *rnpB* gene (Collins et al. 2000) is based on rather weak secondary structure similarities to known homologs, does not consider otherwise highly conserved primary sequence motifs, and has not been confirmed experimentally. Inferences that do not take primary sequence conservation into account are compromised because A+T-rich chloroplast sequences can fit almost any

given consensus structure and are, therefore, of little predictive value. Moreover, several biochemical studies indicate that chloroplast RNase P (cpRNase P) may not contain an organelle coded P-RNA. For instance, spinach cpRNase P activity resists micrococcal nuclease treatment and has physical properties consistent with the presence of a protein-only enzyme. These characteristics indicate that spinach cpRNase P is indeed a protein-only enzyme (Thomas et al. 2000).

Mitochondrial RNase P activities have been studied in various yeasts (see below), the ascomycete fungus *Aspergillus nidulans* (Lee et al. 1996a), human (Doersen et al. 1985; Rossmanith and Karwan 1998; Puranam and Attardi 2001; Rossmanith and Potuschak 2001), *Trypanosoma brucei* (Salavati et al. 2001), potato (Marchfelder and Brennicke 1994), wheat (Hanic-Joyce and Gray 1990), and carrot (Franklin et al. 1995). The most detailed information on the biochemical and genetic properties of mitochondrial RNase P (mtP-RNA) is available for *Saccharomyces cerevisiae*. Its RNA subunit was first identified as a mitochondrially encoded molecule by analyzing yeast mitochondrial mutants deficient in mitochondrial tRNA processing and protein synthesis (Miller and Martin 1983; Underbrink-Lyon et al. 1983). The unusually large protein subunit has been shown to be nucleus-encoded (Morales et al. 1992; Dang and Martin 1993).

Additional *rnpB* genes have been identified by sequence similarity in mtDNAs of numerous budding yeasts (*T. glabrata*, Clark-Walker et al. 1985; *Saccharomycopsis fibuligera*, Wise and Martin 1991a; *Kluyveromyces lactis*, Wilson et al. 1989; *Saccharomyces exiguus*, Wise and Martin 1991b; *Saccharomyces douglasii*; Ragnini et al. 1991; *Saccharomyces chevalieri*, *Saccharomyces ellipsoideus*, and *Saccharomyces diastaticus*, Sbisa et al. 1996; and *Saccharomyces castellii*, Petersen et al. 2002), the protist *Reclinomonas americana* (Lang et al. 1997), and the prasinophyte green alga *N. olivacea* (Turmel et al. 1999), but not in the mtDNAs of plants, animals, a great number of protists, or nonascomycete fungi (Lang et al. 1999). From an evolutionary standpoint, it is puzzling that the occurrence of mitochondrially encoded *rnpB* genes is so patchy. Have these genes been lost from the mtDNAs of plants, animals, most fungi, and protists, or do we fail to identify them because they are so extremely derived? In fact, highly derived, extremely A+U-rich mtP-RNAs are characteristic in yeast. Consequently, predictions of yeast mtP-RNA secondary structures are difficult, despite the availability of comparative data. In addition, the drastic size variations of these RNA molecules obscure the identification of RNA secondary structure elements (Wise and Martin 1991a). For example, the respective lengths of predicted mtP-RNAs are 423 nt for *S. cerevisiae* (Stribinskis et al. 1996), 227 nt for *Torulopsis glabrata* (Shu et al. 1991), and as short as 140 nt for *Saccharomycopsis fibuligera* (Wise and Martin 1991a).

The mitochondrially encoded mtP-RNA of the protist *Reclinomonas americana* was the first identified to contain

all structural elements defined in the eubacterial P-RNA consensus structure (Lang et al. 1997). Its features served as a hallmark for reanalyzing the mtP-RNA secondary structure of *A. nidulans*. This analysis reveals a secondary structure that is substantially more similar to both the *Reclinomonas* mitochondrial and to the eubacterial consensus (Martin and Lang 1997) than to the previously published model (Lee et al. 1996b). In this article, we present the results of our ongoing efforts to identify ascomycete *rnpB* genes, and to infer their RNA secondary structures by comparative phylogenetic modeling. We provide evidence that the mitochondrial *rnpB* genes have been lost several times independently in fungi.

RESULTS

Search for ascomycete mitochondrial *rnpB* genes

To identify previously unrecognized, possibly highly derived *rnpB* genes in fungal mtDNAs, we used two equally productive in silico procedures. The more straightforward of the two was based on sequence pattern recognition (PERL regular expression matching), including the most conserved *rnpB* sequence motifs CRI and CRV (GGAAAGTC...50–500 nt...ACANAANNNGCTTAT; Fig. 1; Chen and Pace 1997). Potential deviations from the consensus sequence were explored by gradually changing the primary sequence motifs, and subsequently, all matches were validated by folding of the highly conserved P4 helix (Fig. 2A,B). Only those candidate *rnpB* sequences that fall into a non-coding region of a given mtDNA were further investigated. The second approach used the dedicated pattern matching tool, RNAMOT (Laferrière et al. 1994), which searches for both sequence similarity and RNA secondary structure folding. Because this program does not allow errors,

possible deviations from the consensus sequence were also explored by systematic changes of the sequence descriptor. All publicly available fungal mtDNAs were analyzed accordingly.

Significant CRI and CRV matches turned up for the three known *rnpB* genes in *S. cerevisiae* (Miller and Martin 1983), *S. castellii* (Petersen et al. 2002), and *A. nidulans* (Lee et al. 1996a), and four new candidates emerged in *Pichia canadensis* (Sekito et al. 1995), *Schizosaccharomyces pombe* (Lang et al. 1983), *Schizosaccharomyces octosporus* (Bullerwell et al. 2003b), and *Taphrina deformans*, a fungus classified within Archiascomycota (or Taphrinomycetales, according to the National Center for Biotechnology Information [NCBI] nomenclature). Sequence alignments are shown in Figure 1. However, we failed to detect mitochondrial *rnpB* genes in six other ascomycetes and in all basidiomycete and chytridiomycete fungi for which complete, or near-complete, mtDNA sequences are available (Table 1), even at search pattern stringency close to background levels.

Modeling of fungal mtP-RNA secondary structures

Fungal mitochondrial *rnpB* genes, including those newly identified, vary considerably in size and primary sequence. In addition, most are exceedingly rich in A and T, which makes RNA modeling an intricate task. It requires substantial comparative sequence information to determine a biologically meaningful solution among the many alternative structure predictions. To deal with the dilemma that in the absence of primary sequence conservation, A+T-rich sequences can be fitted to any given consensus structure, we have based our predictions on three criteria, namely, similarity to (1) the bacterial minimum consensus structure, (2) the universally conserved motifs in bacterial P-RNAs, and (3) the secondary structure models of *R. americana* and *A. nidulans* mtP-RNA.

From studies of eubacterial P-RNAs, it is known that the P1 and P4 helices are the functionally most important regions of this RNA molecule (i.e., most sensitive to mutations; see the bacterial consensus structure in Fig. 2A; Schlegel et al. 1994). This is reflected by the presence of the highly conserved mitochondrial CRI and CRV motifs in all mitochondrial *rnpB* sequences (Fig. 1), which together fold into a P4 helix. In addition, here we demonstrate for the first time that P1 is also conserved in all known mtP-RNAs (Fig. 2A,B), with the possible exception of *K. lactis*, as discussed below. However, several other helical regions that characterize the bacterial consensus structure are lost in mtP-RNAs, as outlined below.

Species	CR I	CR IV	CR V
<i>E. coli</i>	GAGGAAAGGCCGGC (250)	CUAGAUAAUAGACU (6)	GACAGAA-CCCGGCUUAG
<i>R. prowazekii</i>	GAGGAAAGGCCGGAC (201)	CUAGAUAAAUAAACU (18)	UACAGAA-UCCGGCUUAG
<i>R. americana</i>	UAGGAAAGGCCGGAC (193)	UCAGAGAAAUAGAC (28)	GACAAAA-UCCAGCUUAG
<i>N. olivacea</i>	AAGGAAAGGCCGGGC (218)	UCAGAUAAAUAGAC (29)	GACAAAA-CCCGGCUUAG
<i>A. nidulans</i>	AAGGAAAGGCCGGACA (134)	UUAGAAUUAAUAAA (8)	AACAGAA-ACCGGCUUAA
<i>T. deFormans</i>	AAGGAAAGGCCGGGA (211)	UAGAAACAAGUGUA (12)	AACAUAA-UCCGGCUUAG
<i>S. pombe</i>	UAGGAAAGGCCGGAU (135)	UUUUAGDAAUUCCA (5)	AACAGAA-UCCAGCUUAG
<i>S. octosporus</i>	AAGGAAAGGCCGGAA (91)	UUUAAGUAUUUCCA (8)	UACAUAA-UCCAGCUUAG
<i>S. cerevisiae</i>	UAGGAAAGGCCAUAAA (339)	AUAGUUAAAUUUU (1)	UACAGAA-AUAGGCUUAA
<i>S. castellii</i>	AAGCAAADCAUAAA (379)	UAAGUUAAAUAAA (3)	AACAAA-CUAGGCUUAA
<i>S. exiguus</i>	AAGGAAAGGCCAUAAU (207)	AUAGUUAAAUUUU (1)	AACAAAA-AUAGGCUUAG
<i>T. glabrata</i>	AAGAAAGGCCAUAAA (150)	AUAGUUAAAUAAA (0)	UACAUAA-AUAGGCUUAG
<i>S. fibuligera</i>	UAGGAAAGGCCAUAAA	no match	(81) AACAAAA-UUAGGCUUAG
<i>K. lactis</i>	UAGG-AAAGCCAUUU (100)	AUAGAAAUAAU—	(0) AACAUAA-AUAGGCUUAG
<i>P. canadensis</i>	AAGGAAAGGCCAUAAU (72)	CUAGUUAAAUAAA (0)	AACAAAAAUAGGCUUAG

FIGURE 1. Alignment of the universally conserved regions (CRs) I, IV, and V. CRs of the universally conserved nucleotides are in bold, following the analysis by Chen and Pace 1997. Shaded nucleotides are conserved at >75% in *E. coli*, *R. prowazekii*, and mitochondrial P-RNAs. The numbers between brackets indicate the number of nucleotides that separate the conserved regions.

Budding yeast mtP-RNAs

To define a consensus RNA secondary structure model for budding yeast mtP-RNAs, we performed a comparative analysis of all available sequences, except those from *S. chevalieri*, *S. douglasii*, *S. ellipsoideus*, and *S. diastaticus*,

because they are almost identical to the *S. cerevisiae* sequence.

In addition to the CRI and CRV motifs, we found a potential, previously unrecognized CRIV motif (AGNNNNAU; Figs. 1, 2A). Identification of this motif was based on the assumption that a P2 helix would be absent in

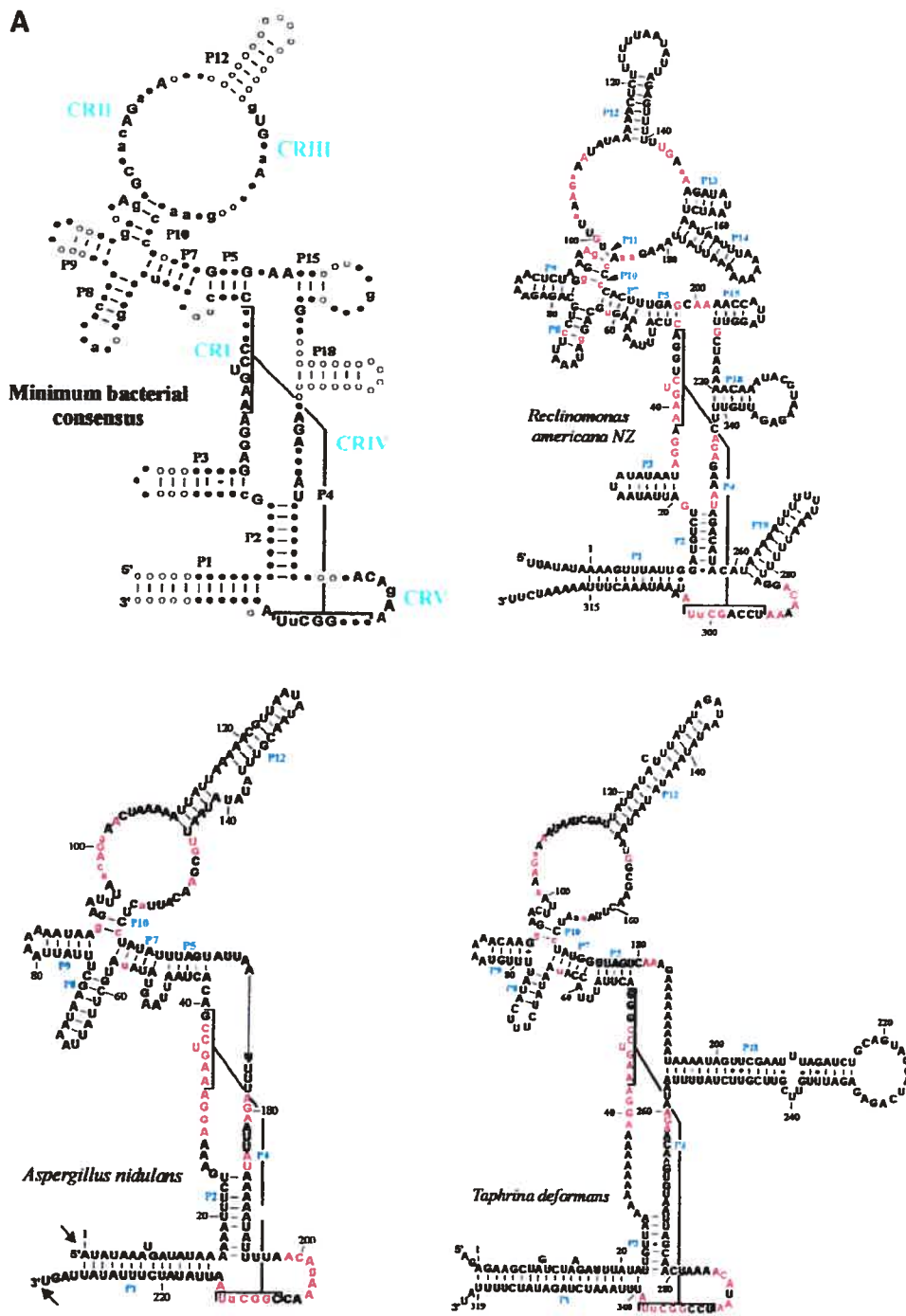
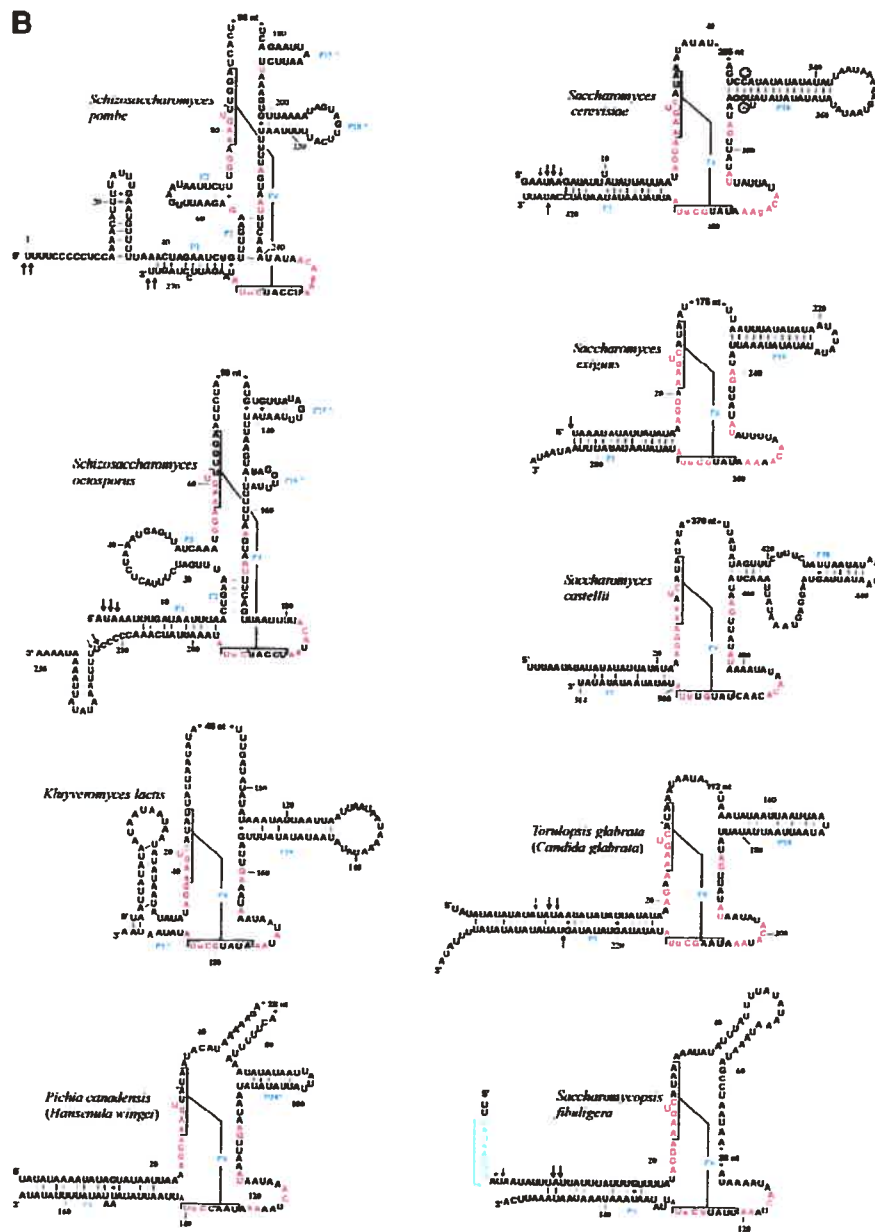


FIGURE 2. (Continued on next page)

all yeast mtP-RNAs, such that it would be positioned directly adjacent to the CRV motif. By using these criteria, CRIV motifs were found in all but the *S. fibuligera* and *K. lactis* *rnpB* sequences (Figs. 1, 2B). In addition, in *K. lactis* the P1 helix is formed by only few base pairs (labeled "P1?")

in Fig. 2B), the CRI motif has a deletion of a universally conserved A, and CRV has a U in the third position that is otherwise a C (Fig. 1). We confirmed the unconventional deletion of an A in the *K. lactis* CRI motif, by resequencing the clone containing the *rnpB* gene. Although a transcript of



**FIGURE 2.** Minimum bacterial P-RNA consensus and mtP RNA secondary structure models of *R. americana* and ascomycete fungi. (A) Positions in red are invariant in the minimum bacterial consensus; uppercase letters in the mtP RNAs indicate 100%, lowercase letters at least 90%, conservation of the minimum bacterial consensus sequence. The conserved motifs CRI through CRV are indicated in blue. The arrows indicate experimentally determined RNA extremities; arrow thickness is proportional to the percentage of molecules ending at this position. (B) mtP-RNA secondary structures from nine additional ascomycetes. In the *S. cerevisiae* structure, nucleotides that co-vary with the *Saccharomyces douglasii* sequence are in red circled letters. A putative promoter in *S. fibuligera* is highlighted in blue. Asterisks next to nucleotides in the P4 of *S. pombe*, *S. octosporus*, *C. glabrata*, and *P. canadensis* indicate mispairs in these helices. Note that the *Candida* structure P1 helix can be further extended by 9 to 12 bp from its mature ends. It is possible that the noncanonical A-G pair adjacent to the RNA processing sites in this mtP-RNA serves as a signal for RNA maturation.



TABLE 1. Occurrence and features of *rnpB* in complete fungal mtDNA sequences

Organismal group	mtDNA- encoded <i>rnpB</i> <sup>a</sup>	P2	P3	P15	P18	P5,7,8, 9,10,12	Accession #
Ascomycota (Euascomycota)							
<i>Aspergillus nidulans</i>	■	■	□	□	□	■	<sup>b</sup>
<i>Hypocrea jecorina</i>	□						NC_003388
<i>Neurospora crassa</i>	□						<sup>c</sup>
<i>Podospira anserina</i>	□						NC_001320
<i>Verticillium lecanii</i>	□						AF487277
Ascomycota (Hemiascomycota)							
<i>Candida albicans</i>	□						NC_002653
<i>Candida glabrata</i>	■	□	□	□	■	□	AJ511533
<i>Pichia canadensis</i>	■	□	□	□	■	□	NC_001762
<i>Saccharomyces castellii</i>	■	□	□	□	■	□	AF437291
<i>Saccharomyces cerevisiae</i>	■	□	□	□	■	□	NC_001224
<i>Yarrowia lipolytica</i>	□						NC_002659
Ascomycota (others)							
<i>Schizosaccharomyces japonicus</i>	□						NC_004332
<i>Schizosaccharomyces octosporus</i>	■	■	■	■	■	□	AF275271
<i>Schizosaccharomyces pombe</i>	■	■	■	■	■	□	NC_001326
<i>Taphrina deformans</i>	■	■	□	□	■	■	AY262107 <sup>d</sup>
Basidiomycota <sup>e</sup>							
	□						
Chytridiomycota <sup>e</sup>							
	□						

Filled and empty squares stand for presence or absence of the gene or helices, respectively.

<sup>a</sup>All mtP-RNAs contain P1 and P4.

<sup>b</sup>The *A. nidulans* sequence is almost complete: ODAS1, CAA33481, AAA99207, AAA31737, CAA25707, AAA31736, CAA23994, X15442, P15956, CAA23995, CAA33116, X00790, X15441, X06960, J01387, and X01507.

<sup>c</sup>Completed at Whitehead Institute, currently available at <http://pages.slu.edu/faculty/kennel/j/genbank.html>.

<sup>d</sup>The GenBank record documents the *rnpB* sequence, the complete sequence is unpublished.

<sup>e</sup>Mitochondrial *rnpB* genes have not been identified in basidiomycete (*Cantharellus cibarius*, unpubl.; *Cryptococcus neoformans*, NC00\_4336; *Microbotryum violaceum*, unpubl.; *Schizophyllum commune*, NC\_003049) or chytridiomycete fungi (*Allomyces macrogynus*, NC\_001715; *Harpochytrium94*, AY182005; *Harpochytrium105*, AY182006; *Hyaloraphidium curvatum*, NC\_003048).

the expected size is found in *K. lactis* mitochondria (Wilson et al. 1989), the high incidence of noncanonical features raises the question whether or not *rnpB* codes for a functional mtP-RNA.

Further comparative analysis revealed another conserved base pairing, P18 (Fig. 2A), in the five phylogenetically closely related organisms *S. cerevisiae*, *S. castellii*, *T. glabrata*, *S. exiguus*, and *K. lactis* (Fig. 2B). These helical regions start at the same position as P18 of *R. americana* and as P18 in the bacterial minimum consensus model, 2 nt upstream of the CRIV motif (Fig. 1B). Additional support for this P18 helix comes from a compensatory exchange of a base pair (G C to A-U) between the *S. cerevisiae* and *S. douglasii* mtP-RNAs (see the circled nucleotides of the *S. cerevisiae* structure in Fig. 2B). In *P. canadensis*, a P18 pairing is possible (labeled "P18?" in Fig. 2B); however, it is not located at the same position relative to the CRIV motif. Finally, we did not detect a CRIV motif or a P18 pairing in the *S. fibuligera* mtP-RNA.

#### Fission yeasts

The *S. pombe* *rnpB* gene was only identified recently, although the complete sequence of mtDNA has been available

for more than a decade. Differences in both the CRI and CRV motifs, relative to the consensus (Fig. 1), made it difficult to recognize this gene. After identifying potential CRI, CRIV, and CRV motifs in an intergenic region of the *S. pombe* mtDNA, we confirmed the presence of an *rnpB* homolog in *S. octosporus* at the same location in the mtDNA, flanked by the same tRNA genes, *trnA* and *trnG*.

The comparative structural modeling of the two *Schizosaccharomyces* sequences reveals not only the presence of P1 and P4, as in budding yeasts, but also P2 and P3 pairings that distinguish them from most other fungal RNA structures (Fig. 2B). Moreover, we identified potential P15 and P18 pairings. However, only the *S. pombe* P18 is located at the same distance from CRIV as in other species, and the potential *S. octosporus* P18 helix would be only 3 bp long (Fig. 2B). The central sequence of *rnpB* contains three motifs that are identical in the fission yeasts (4, 6, and 9 nt long; Fig. 2B), but without similarity to the bacterial consensus.

*T. deformans* has been grouped together with *Schizosaccharomyces* species in phylogenetic analyses (termed archiascomyetes, Nishida and Sugiyama 1993; for an alternative view, see Bullerwell et al. 2003b; Leigh et al. 2003), but the mtP-RNA secondary structure of *T. deformans* exhibits

striking similarity only to the euascomycete *A. nidulans* (Fig. 2B). The only principal structural difference is the extra P18 pairing in *T. deformans*.

### Mapping of mtP-RNAs

As indicated in the bacterial consensus structure (Fig. 2A), the 5' and 3' extremities of most mapped P-RNA molecules are directly adjacent, or at least close to, the P1 helix of P-RNAs. We show here by primer extension that the putative *S. fibuligera* mtP-RNA is processed at its 5' end to an RNA molecule with a perfect terminal P1 helix (Figs. 2B, 3). In *S. pombe* and *S. octosporus*, the 5' ends were initially mapped by primer extension experiments (Fig. 3) and confirmed by RT-PCR of the circularized RNAs and sequencing of the PCR product. The sequencing results confirmed the primer extension results shown in Figure 2B and revealed that there is only a marginal heterogeneity of mtP-RNA extremities.

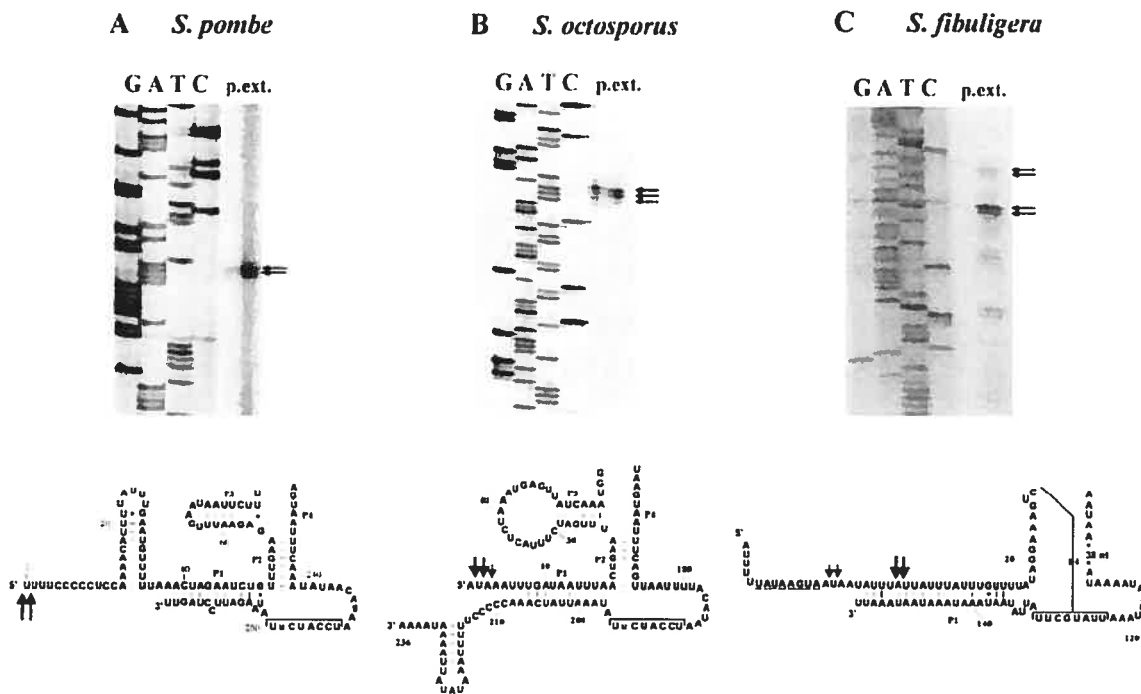
Results further show that the 3' end of *S. pombe* and the 5' end of *S. octosporus* mtP-RNAs correspond well to the predicted RNA structure model (Fig. 2B). The 5' end of *S. pombe* and the 3' end of *S. octosporus* are somewhat longer

than expected, at sequence motifs rich in cytidines and uridines. Similar motifs are present at most 3' ends of mitochondrial protein-coding genes in *S. pombe*, *S. octosporus* (Bullerwell et al. 2003b), and several zygomycete and basidiomycete mitochondrial genomes (B.F. Lang, unpubl.). These sequence motifs may be RNA processing signals (Lang et al. 1983; Trinkl et al. 1989), and protect the RNA extremities against further exonuclease digestion.

### DISCUSSION

#### Highly reduced mtP-RNA structures in fungi

We report here the identification of four previously undetected mitochondrial *rnpB* genes in ascomycetes. Comparative RNA secondary structure modeling of previously and newly discovered sequences allowed us to establish a minimum consensus secondary structure for fungal mtP-RNAs (for a complete collection of mtP-RNA structures, see <http://megasun.bch.umontreal.ca/People/lang/rnpB/>). The fungal mitochondrial consensus includes only two out of the five conserved motifs (CRI and CRV) and two out of 11 conserved helices (P1 and P4; Fig. 2A) found in the mini-



**FIGURE 3.** RNA mapping of *S. pombe*, *S. octosporus*, and *S. fibuligera* mtP-RNAs. (A) *S. pombe* primer extension experiment. The bottom part of the figure shows the partial, corresponding region of the mtP RNA structure. The main bands are consistent with sequencing of RT-PCR product of in vitro circularized P-RNA. The thickness of the arrows is proportional to the proportion of RNA extremities of the RT-PCR experiment. (B) *S. octosporus* primer extension experiment. The main bands are consistent with sequencing of RT-PCR product of in vitro circularized P-RNA. The thickness of the arrows is proportional to the proportion of RNA extremities of the RT-PCR experiment. (C) Primer extension for *S. fibuligera* mtP-RNA. The weak upper signal matches the transcription initiation site at the potential promoter sequence (underlined). The strongest signal corresponds to the mature end of the mtP-RNA.

mum bacterial consensus structure. The presence of additional conserved structural elements appears to tightly correlate with the phylogenetic relatedness. For example, P2, P3, P15, and P18 are found in the two *Schizosaccharomyces* species, but only P18 is found in most (but not all) budding yeasts (Table 1). Accordingly, the occurrence of P2, P5, P7–10, and P12 in both *T. deformans* and *A. nidulans* possibly indicates that *T. deformans* is an ancestral fungus that branches at the base of the ascomycetes. Although it has been proposed to be related to *Schizosaccharomyces* species based on rRNA phylogenies (Nishida and Sugiyama 1993), it does not appear to be related based on either mtP-RNA structure or from mitochondrial protein sequences (Leigh et al. 2003).

The reduction of the structural complexity of mtP-RNAs is more extensive in fungal mitochondria than in nuclear (Frank et al. 2000) and archaeal P-RNAs (Harris et al. 2001). It is likely compensated by increase in size and/or number of P-proteins (e.g., the large size of the mitochondrial P-protein in *S. cerevisiae*; Morales et al. 1992; Dang and Martin 1993) and the possibility of multiple mitochondrial P-proteins in *A. nidulans* (Lee et al. 1996b). An investigation of fungal mitochondrial P-proteins and of native mtP-RNA enzymes could be useful in shedding light on the principles related to the replacement of RNA by protein structure. Despite the substantial loss of structural complexity, the fungal mtP-RNAs are likely to be the catalytic component of RNase P. Deletion analysis shows that bacterial P1 and P4 elements are intimately involved in catalytic function (Guerrier-Takada and Altman 1992; Schlegel et al. 1994). This hypothesis is also supported by genetic and biochemical studies in *S. cerevisiae* and *A. nidulans*, showing that the RNA subunit is essential for mtP-RNA activity (Miller and Martin 1983; Underbrink-Lyon et al. 1983; Lee et al. 1996a).

### Processing of fungal mtP-RNAs

The processing of precursor molecules of structural RNAs usually depends on highly specific biochemical activities that recognize RNA structure. This prompted the question whether this also applies to mtP-RNAs. The RNA mapping data presented here indicate two types of mtP-RNA processing. First, in *S. pombe*, *S. octosporus*, and several budding yeasts (including *S. cerevisiae* and *C. glabrata*; Shu and Martin 1991), *rnpB* genes are flanked by tRNA genes. The removal of tRNA sequences from the RNA precursor directly liberates mature mtP-RNA molecules in some instances (e.g., as shown here in *S. pombe*, Fig. 2B; and as predicted in *P. canadensis*, *S. exiguus*, and *K. lactis* from the positioning of the postulated processing sites). It follows the tRNA "punctuation model" first described as the key principle of human mitochondrial RNA processing (Ojala et al. 1980), as well as in *S. pombe* and a few other fungi (Paquin et al. 1997). Note that RNA processing by tRNA punctua-

tion is not a general principle in budding yeast mitochondria, although applied in maturation of the mtP-RNA. In fact, there is a high overall incidence of mitochondrial *rnpB* genes that are flanked on both sides, or at least one side (and then most frequently at the 3' end), by tRNA genes. This implies that in numerous instances, RNase P itself has a defining role in 3' processing of its RNA subunit. Second, in some cases, 3' and 5' end processing occurs subsequent to tRNA removal. For example, processing of 3' ends of transcripts containing cytidine-uridine-rich motifs occurs in *S. octosporus* (Fig. 2B). The primary transcript in *S. cerevisiae* includes tRNA<sup>Metf</sup>, mt-PRNA (also called Rpm1r), and tRNA<sup>Pro</sup>. The tRNAs are first separated from the mtP-RNA precursor, and secondary maturation occurs at both P-RNA ends (Stribinskis et al. 1996), with the possible involvement of the mitochondrial P-protein (Stribinskis et al. 2001). The *rnpB* gene of *S. fibuligera* is flanked only downstream by tRNA<sup>Pro</sup>, but upstream by a protein coding gene (*cox2*). The sequence immediately upstream of *rnpB* contains a conserved octanucleotide promoter motif (TATAAGTA; Fig. 2) that serves as a promoter in *S. cerevisiae*. Our primer extension analysis in *S. fibuligera* indicates that *rnpB* is indeed transcribed from this putative promoter (Fig. 3), so that starting transcription close to the 5' end is yet another potential mechanism leading to the definition of mtP-RNA extremities. The same mechanism potentially applies to *S. cerevisiae*, as an alternative pathway to maturation by tRNA punctuation (Biswas 1996). Finally, in *A. nidulans* and *T. deformans* none of the previously discussed mechanisms seem to apply.

### Mitochondrial *rnpB* genes in other fungi, plants, or protists?

We failed to detect mitochondrial *rnpB* genes in 18 out of 26 complete fungal mtDNAs. These negative matches included all studied basidiomycete and chytridiomycete fungi, four euascomycete fungi, two budding yeasts, and one fission yeast (Table 1). In these species, a mitochondrial *rnpB* either migrated to the nucleus, or had its function replaced by a nuclear encoded gene, or, less likely, diverged to a degree that it is no longer recognizable. There is evidence to indicate that *rnpB* is indeed absent from most of the above-listed mtDNAs. In the chytridiomycete fungi *Harpochytrium94* and *Harpochytrium105* (Bullerwell et al. 2003a), intragenic regions are either too short or filled with repeats so there is no space to fit an *rnpB* gene. The chytridiomycete fungus, *A. macrogynus*, has the most ancestral (undervived) mtDNA sequence among all known fungi, with respect to both gene content and gene sequence similarity. Therefore, although one might expect that *A. macrogynus* would have an easily identifiable, eubacterial-like mitochondrial *rnpB*, we were unable to locate even a highly derived version of this gene. *S. pombe* and *S. octosporus* mtDNAs contain *rnpB* sequences that are highly similar and easily identified by sequence comparisons. Yet, there is no significant match

with a third fission yeast mtDNA, that of *S. japonicus*. *P. anserina*, *H. jecorina*, and *Verticillium lecanii* are closely related to *A. nidulans*. *A. nidulans* is the only one of the four that has a eubacteria-like *rnpB* sequence. Otherwise, their mitochondrial gene content is identical, which indicates that none of these four species has undergone accelerated evolution. Evidently, sequences from more deeply diverging members of the euascomycetes need to be investigated to support the idea that loss of mitochondrial *rnpB* is typical for this fungal group.

This widespread loss raises the question of how the RNase P function is provided in the absence of a mitochondrial encoded *rnpB* gene. One possibility is that a protein-only enzyme, coded by nuclear gene(s) and imported into the organelle, is responsible for 5' processing of precursor tRNAs, a situation postulated for spinach chloroplast (Thomas et al. 2000). An alternate possibility is that an RNA is required for catalysis but that it is encoded in the nucleus and imported into mitochondria to function. Biochemical characterization of the RNase P activity isolated from mitochondria lacking an endogenous *rnpB* will be necessary to differentiate between these possibilities. Such studies have been undertaken for human mtP-RNA, but there is an unresolved debate whether or not the absence of mitochondrially encoded *rnpB* is compensated by the nuclear P-RNA, or by a protein-only enzyme (Rossmannith and Karwan 1998; Puranam and Attardi 2001; Rossmannith and Potuschak 2001).

## MATERIALS AND METHODS

### DNA and RNA extraction, PCR, and sequencing

*S. pombe* strain *ade7-50h<sup>-</sup>* (kindly provided by U. Leupold, University of Bern, Switzerland) *S. octosporus* (ATCC 2479), and *T. deformans* (NRRL T-857) were grown in 1% yeast extract plus 3% glycerol liquid medium. For mitochondrial DNA and RNA extractions, cells were broken mechanically, and a mitochondrial fraction was isolated by differential centrifugation (Lang et al. 1977). This fraction was lysed in the presence of 1% SDS and 100 µg/mL proteinase K for 1 h at 37°C, and after phenol-chloroform extraction, the nucleic acids were precipitated with ethanol. For RNA extractions of *Schizosaccharomyces* strains, the cell walls were digested with lytic enzymes from *Trichoderma harzianum* (Sigma L-2265), and mitochondria were purified by differential centrifugation. The mitochondrial fraction was solubilized in the presence of 1% SDS and 100 µg/mL proteinase K for 1 h at 37°C. After a phenol-chloroform extraction, the high-molecular-weight RNA fraction was precipitated with 2M LiCl, redissolved in RNase-free water, and ethanol-precipitated.

### Primer extension

Primer extensions were performed with primers 5'-CCCTCTGGGTTTCITTTTTTA-3' (*S. pombe*) and 5'-GATGGATTATGTAAAATTAAGT-3' (*S. octosporus*). The primers were labeled at

the 5' end with  $\gamma$ -P32 ATP and T4 polynucleotide kinase (Boehringer). The mitochondrial RNA was incubated for 30 min at 37°C in the presence of the respective primers, dNTPs, AMV buffer (Boehringer), and AMV reverse transcriptase (Boehringer 1495062). The resulting product was ethanol-precipitated and loaded on a high resolution sequencing gel (see above). The DNA sequencing ladder that served as a size marker was produced with the same primer from a DNA template containing the *rnpB* region from the organism being examined.

### RNA ligation and RT-PCR

RNA ligation of mtP-RNAs, followed by RT-PCR amplification, was performed to determine precise 5' and 3' end(s). The protocol was essentially the same used to circularize tRNAs (Yokobori and Pääbo 1995), except that ~10 µg of RNA was used as starting material. The primers used for RT-PCR were 5'-CCCTCTTG GTTTCITTTTTTA 3' and 5'-TTTTAGTAAATTTCAATATATA CAG 3' for *S. pombe*, and 5'-TCCAAACTTTCCATTTGATAAC-3' and 5'-CAGTTAATTTTACATAATCCATC-3' for *S. octosporus*.

## ACKNOWLEDGMENTS

We thank G. Burger and C. Bullerwell for comments on the manuscript, Wang Zhang for participation in DNA sequencing, Marlene Steffen for technical assistance, and H. Fukuhara (Orsay, France) for sharing information on *K. lactis rnpB* and providing a clone carrying its sequence. This project was supported by operating grants from the Canadian Institute for Health Research (CIHR) to B.F.L. and the National Institutes of Health (NCM). Salary and interaction support (to B.F.L.) from the Canadian Institute for Advanced Research is gratefully acknowledged.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Received May 15, 2003; accepted June 18, 2003.

## REFERENCES

- Altman, S., Wesolowski, D., and Puranam, P.S. 1993. Nucleotide sequences of the RNA subunit of RNase P from several mammals. *Genomics* 18: 418–422.
- Arends, S. and Schön, A. 1997. Partial purification and characterization of nuclear ribonuclease P from wheat. *Eur. J. Biochem.* 244: 635–645.
- Baum, M., Cordier, A., and Schön, A. 1996. RNase P from a photosynthetic organelle contains an RNA homologous to the cyanobacterial counterpart. *J. Mol. Biol.* 257: 43–52.
- Biswas, T.K. 1996. Expression of the mitochondrial RNase P RNA subunit-encoding gene from a variant promoter sequence in *Saccharomyces cerevisiae*. *Gene* 170: 23–30.
- Brown, J.W. 1999. The Ribonuclease P Database. *Nucleic Acids Res.* 27: 314
- Brown, J.W. and Haas, E.S. 1995. Ribonuclease P structure and function in Archaea. *Mol. Biol. Rep.* 22: 131–134.
- Bullerwell, C.E., Forget, L., and Lang, B.F. 2003a. Evolution of monopharidalean fungi based on complete mitochondrial genome sequences. *Nucleic Acids Res.* 31: 1614–1623.

- Bullerwell, C.E., Leigh, J., Forget, L., and Lang, B.F. 2003b. A comparison of three fission yeast mitochondrial genomes. *Nucleic Acids Res.* 31: 759–768.
- Chen, J.L. and Pace, N.R. 1997. Identification of the universally conserved core of ribonuclease P RNA. *RNA* 3: 557–560.
- Christian, E.L., Zahler, N.H., Kaye, N.M., and Harris, M.E. 2002. Analysis of substrate recognition by the ribonucleoprotein endonuclease RNase P. *Methods* 28: 307–322.
- Clark Walker, G.D., McArthur, C.R., and Sriprakash, K.S., 1985. Location of transcriptional control signals and transfer RNA sequences in *Torulopsis glabrata* mitochondrial DNA. *EMBO J.* 4: 465–473.
- Collins, L.J., Moulton, V., and Penny, D. 2000. Use of RNA secondary structure for studying the evolution of RNase P and RNase MRP. *J. Mol. Evol.* 51: 194–204.
- 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<sup>Asp</sup>. *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.* 269: 19791–19796.
- Darr, S.C., Pace, B., and Pace, N.R. 1990. Characterization of ribonuclease P from the archaeobacterium *Sulfolobus solfataricus*. *J. Biol. Chem.* 265: 12927–12932.
- Darr, S.C., Zito, K., Smith, D., and Pace, N.R. 1992. Contributions of phylogenetically variable structural elements to the function of the ribozyme ribonuclease P. *Biochemistry* 31: 328–333.
- 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.
- Frank, D.N., Adamidi, C., Ehringer, M.A., Pitulle, C., and Pace, N.R. 2000. Phylogenetic-comparative analysis of the eukaryal ribonuclease P RNA. *RNA* 6: 1895–1904.
- Franklin, S.E., Zwick, M.G., and Johnson, J.D. 1995. Characterization and partial purification of two pre-tRNA 5' processing activities from *Daucus carota* (carrot) suspension cells. *Plant J.* 7: 553–563.
- Gardiner, K.J., Marsh, T.L., and Pace, N.R. 1985. Ion dependence of the *Bacillus subtilis* RNase P reaction. *J. Biol. Chem.* 260: 5415–5419.
- Glöckner, G., Rosenthal, A., and Valentin, K. 2000. The structure and gene repertoire of an ancient red algal plastid genome. *J. Mol. Evol.* 51: 382–390.
- Gopalan, V., Vioque, A., and Altman, S. 2002. RNase P: Variations and uses. *J. Biol. Chem.* 277: 6759–6762.
- Guerrier-Takada, C., and Altman, S. 1992. Reconstitution of enzymatic activity from fragments of M1 RNA. *Proc. Natl. Acad. Sci.* 89: 1266–1270.
- 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.
- Guerrier-Takada, C., Eder, P.S., Gopalan, V., and Altman, S. 2002. Purification and characterization of Rpp25, an RNA-binding protein subunit of human ribonuclease P. *RNA* 8: 290–295.
- Haas, E.S., Brown, J.W., Pitulle, C., and Pace, N.R. 1994. Further perspective on the catalytic core and secondary structure of ribonuclease P RNA. *Proc. Natl. Acad. Sci.* 91: 2527–2531.
- 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.
- Hanic-Joyce, P.J. and Gray, M.W. 1990. Processing of transfer RNA precursors in a wheat mitochondrial extract. *J. Biol. Chem.* 265: 13782–13791.
- 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.
- Hartmann, R.K., Heinrich, J., Schlegel, J., and Schuster, H. 1995. Precursor of CA antisense RNA of bacteriophages P1 and P7 is a substrate for RNase P of *Escherichia coli*. *Proc. Natl. Acad. Sci.* 92: 5822–5826.
- Jayanthi, G.P. and Van Tuyle, G.C. 1992. Characterization of ribonuclease P isolated from rat liver cytosol. *Arch. Biochem. Biophys.* 296: 264–270.
- Kole, R. and Altman, S. 1981. Properties of purified ribonuclease P from *Escherichia coli*. *Biochemistry* 20: 1902–1906.
- Laferriere, A., Gautheret, D., and Cedergren, R. 1994. An RNA pattern matching program with enhanced performance and portability. *Comput. Appl. Biosci.* 10: 211–212.
- Lang, B.F., Burger, G., Doxiadis, I., Thomas, D.Y., Bandlow, W., and Kaudewitz, F. 1977. A simple method for the large-scale preparation of mitochondria from microorganisms. *Anal. Biochem.* 77: 110–121.
- Lang, B.F., Ahne, F., Distler, S., Trinkl, H., Kaudewitz, F., and Wolf, K. 1983. Sequence of the mitochondrial DNA, arrangement of genes and processing of their transcripts in *Schizosaccharomyces pombe*. In *Molecular biology of the fission yeast* (eds. A. Nasim, et al.), pp. 3118–3119, Academic Press, San Diego, CA.
- 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.
- Lang, B.F., Gray, M.W., and Burger, G. 1999. Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.* 33: 351–397.
- Lee, Y.C., Lee, B.J., Hwang, D.S., and Kang, H.S. 1996a. Purification and characterization of mitochondrial ribonuclease P from *Aspergillus nidulans*. *Eur. J. Biochem.* 235: 289–296.
- Lee, Y.C., Lee, B.J., and Kang, H.S. 1996b. The RNA component of mitochondrial ribonuclease P from *Aspergillus nidulans*. *Eur. J. Biochem.* 235: 297–303.
- Leigh, J., Seif, E., Rodriguez, N., Jacob, Y., and Lang, B.F. 2003. Fungal evolution meets fungal genomics. In *Handbook of fungal biotechnology*, 2nd ed. (ed. Arora D.). Marcel Dekker Inc., New York (in press).
- Marchfelder, A. and Brennicke, A. 1994. Characterization and partial purification of tRNA processing activities from potato mitochondria. *Plant Physiol.* 105: 1247–1254.
- Martin, C.A. and Lang, B.F. 1997. Mitochondrial RNase P: The RNA family grows. *Nucleic Acids Symp. Ser.* 36: 42–44.
- Miller, D.L. and Martin, N.C. 1983. Characterization of the yeast mitochondrial locus necessary for tRNA biosynthesis: DNA sequence analysis and identification of a new transcript. *Cell* 3: 911–917.
- Morales, M.J., Dang, Y.L., Lou, Y.C., Sulo, P., and Martin, N.C. 1992. A 105 k-Da protein is required for yeast mitochondrial RNase P activity. *Proc. Natl. Acad. Sci.* 89: 9875–9879.
- Nishida, H. and Sugiyama, J. 1993. Phylogenetic relationships among *Taphrina*, *Saitoella*, and other higher fungi. *Mol. Biol. Evol.* 10: 431–436.
- Ojala, D., Merkel, C., Gelfand, R., and Attardi, G. 1980. The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. *Cell* 22: 393–403.
- 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.* 96: 7803–7808.
- Paquin, B., Laforest, M.J., Forget, L., Roewer, I., Wang, Z., Longcore, J., and Lang, B.F. 1997. The fungal mitochondrial genome project: Evolution of fungal mitochondrial genomes and their gene expression. *Curr. Genet.* 31: 380–395.
- Pascual, A. and Vioque, A. 1999. Functional reconstitution of RNase P activity from a plastid RNA subunit and a cyanobacterial protein subunit. *FEBS Lett.* 442: 7–10.
- Peck-Miller, K.A. and Altman, S. 1991. Kinetics of the processing of the precursor to 4.5S RNA, a naturally occurring substrate for RNase P from *Escherichia coli*. *J. Mol. Biol.* 221: 1–5.
- Petersen, R.F., Langkjaer, R.B., Hvidtfeldt, J., Gartner, J., Palmén, W.,

- Ussery, D.W., and Piskur, J. 2002. Inheritance and organisation of the mitochondrial genome differ between two *Saccharomyces* yeasts. *J. Mol. Biol.* 318: 627–636.
- 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.
- Ragnini, A., Grisanti, P., Rinaldi, T., Frontali, L., and Palleschi, C. 1991. Mitochondrial genome of *Saccharomyces douglasii*: Genes coding for components of the protein synthetic apparatus. *Curr. Genet.* 19: 169–174.
- Reith, M. and Munholland, J. 1995. Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. *Plant Mol. Biol. Repr.* 13: 333–335.
- Rossmannith, W. and Karwan, R.M. 1998. Characterization of human mitochondrial RNase P: Novel aspects in tRNA processing. *Biochem. Biophys. Res. Commun.* 247: 234–241.
- Rossmannith, W. and Potushak, T. 2001. Difference between mitochondrial RNase P and nuclear RNase P. *Mol. Cell. Biol.* 21: 8236–8237.
- Salavati, R., Panigrahi, A.K., and Stuart, K.D. 2001. Mitochondrial ribonuclease P activity of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 115: 109–117.
- Sbisa, E., Pesole, G., Tullo, A., and Saccone, C. 1996. The evolution of the RNase P- and RNase MRP- associated RNAs: Phylogenetic analysis and nucleotide substitution rate. *J. Mol. Evol.* 43: 46–57.
- Schlegl, J., Hardt, W.D., Erdmann, V.A., and Hartmann, R.K. 1994. Contribution of structural elements to *Thermus thermophilus* ribonuclease P RNA function. *EMBO J.* 13: 4863–4869.
- Sekito, T., Okamoto, K., Kitano, H., and Yoshida, K. 1995. The complete mitochondrial DNA sequence of *Hansenula wingei* reveals new characteristic of yeast mitochondria. *Curr. Genet.* 28: 39–53.
- Shu, H.H. and Martin, N.C. 1991. RNase P RNA in *Candida glabrata* mitochondria is transcribed with substrate tRNAs. *Nucleic Acids Res.* 19: 6221–6226.
- Shu, H.H., Wise, C.A., Clark Walker, G.D., and Martin, N.C. 1991. A gene required for RNase P activity in *Candida (Torulopsis) glabrata* mitochondria codes for a 227-nucleotide RNA with homology to bacterial RNase P RNA. *Mol. Cell. Biol.* 11: 1662–1667.
- Stark, B.C., Kole, R., Bowman, E.J., and Altman, S. 1978. Ribonuclease P: An enzyme with an essential RNA component. *Proc. Natl. Acad. Sci.* 75: 3717–3721.
- 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., Ellis, S.R., and Martin, N.C. 2001. Rpm2, the protein subunit of mitochondrial RNase P in *Saccharomyces cerevisiae*, also has a role in the translation of mitochondrially encoded subunits of cytochrome c oxidase. *Genetics* 158: 573–585.
- 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.
- Trinkl, H., Lang, B.F., and Wolf, K. 1989. Nucleotide sequence of the gene encoding the small ribosomal RNA in the mitochondrial genome of the fission yeast *Schizosaccharomyces pombe*. *Nucleic Acids Res.* 17: 6730.
- Turmel, M., Lemieux, C., Burger, G., Lang, B.F., Otis, C., Plante, I., and Gray, M.W. 1999. The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*. Two radically different evolutionary patterns within green algae. *Plant Cell* 11: 1717–1730.
- Underbrink-Lyon, K., Miller, D.L., Ross, N.A., Fukuhara, H., and Martin, N.C. 1983. Characterization of a yeast mitochondrial locus necessary for tRNA biosynthesis: Deletion mapping and restriction mapping studies. *Mol. Gen. Genet.* 191: 512–518.
- Wagner, M., Fingerhut, C., Gross, H.J., and Schon, A. 2001. The first phytoplasmal RNase P RNA provides new insights into the sequence requirements of this ribozyme. *Nucleic Acids Res.* 29: 2661–2665.
- Wilson, C., Ragnini, A., and Fukuhara, H. 1989. Analysis of the regions coding for transfer RNAs in *Kluyveromyces lactis* mitochondrial DNA. *Nucleic Acids Res.* 17: 4485–4491.
- Wise, C.A. and Martin, N.C. 1991a. Dramatic size variation of yeast mitochondrial RNAs suggests that RNase P RNAs can be quite small. *J. Biol. Chem.* 266: 19154–19157.
- . 1991b. Sequence analysis of *Saccharomyces exiguus* mitochondrial DNA reveals an RNase P RNA flanked by two tRNA genes. *Nucleic Acids Res.* 19: 4773.
- Yokobori, S. and Paabo, S. 1995. Transfer RNA editing in land snail mitochondria. *Proc. Natl. Acad. Sci.* 92: 10432–10435.

## Article 2

### **Comparative mitochondrial genomics in zygomycetes: bacteria-like RNase P RNAs, mobile elements, and a close source of the group I intron invasion in angiosperms.**

**Elias Seif<sup>1</sup>, Jessica Leigh<sup>2</sup>, Yu Liu<sup>1</sup>, Ingeborg Roewer<sup>3</sup>, Lise Forget<sup>1</sup>, and B. Franz Lang<sup>1</sup>**

*Nucleic Acids Research*, 2005, vol.33, no. 2, pp. 734-744

<sup>1</sup>Program in Evolutionary Biology, Canadian Institute for Advanced Research, Département de Biochimie, Université de Montréal, 2900 Boul. Edouard-Montpetit, Montréal, Québec, H3T 1J4, Canada.

<sup>2</sup> Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax (Nova Scotia), B3H 4H7, Canada.

<sup>3</sup>National Research Council of Canada, Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK, S7N 0W9 Canada.

#### Contribution of each author:

Elias Seif : did most of the experimental work, participated in genome analysis, and phylogenetic reconstruction, and prepared the manuscript and figures.

Jesscia Leigh : participated in phylogenetic reconstruction and preparing the manuscript and figures.

Yu Liu : participated in phylogenetic reconstruction.

Lise Forget and Ingeborg Roewer : participated in the experimental work.

B. Franz Lang : participated in genome anlysis, phylogenetic reconstruction, and preparing the manuscript and figures.

# Comparative mitochondrial genomics in zygomycetes: bacteria-like RNase P RNAs, mobile elements and a close source of the group I intron invasion in angiosperms

Elias Seif<sup>1</sup>, Jessica Leigh<sup>2</sup>, Yu Liu<sup>1</sup>, Ingeborg Roewer<sup>3</sup>, Lise Forget<sup>1</sup> and B. Franz Lang<sup>1,\*</sup>

<sup>1</sup>Program in Evolutionary Biology, Canadian Institute for Advanced Research; Centre Robert Cedergren, Département de Biochimie, Université de Montréal, 2900 Boulevard Edouard-Montpetit, Montréal, Québec, Canada H3T 1J4, <sup>2</sup>Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax (Nova Scotia), Canada B3H 4H7 and <sup>3</sup>National Research Council of Canada, Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK, Canada S7N 0W9

DDBJ/EMBL/GenBank accession nos<sup>†</sup>

## ABSTRACT

To generate data for comparative analyses of zygomycete mitochondrial gene expression, we sequenced mtDNAs of three distantly related zygomycetes, *Rhizopus oryzae*, *Mortierella verticillata* and *Smittium culisetae*. They all contain the standard fungal mitochondrial gene set, plus *rnpB*, the gene encoding the RNA subunit of the mitochondrial RNase P (mtP-RNA) and *rps3*, encoding ribosomal protein S3 (the latter lacking in *R.oryzae*). The mtP-RNAs of *R.oryzae* and of additional zygomycete relatives have the most eubacteria-like RNA structures among fungi. Precise mapping of the 5' and 3' termini of the *R.oryzae* and *M.verticillata* mtP-RNAs confirms their expression and processing at the exact sites predicted by secondary structure modeling. The 3' RNA processing of zygomycete mitochondrial mRNAs, SSU-rRNA and mtP-RNA occurs at the C-rich sequence motifs similar to those identified in fission yeast and basidiomycete mtDNAs. The C-rich motifs are included in the mature transcripts, and are likely generated by exonucleolytic trimming of RNA 3' termini. Zygomycete mtDNAs feature a variety of insertion elements: (i) mtDNAs of *R.oryzae* and *M.verticillata* were subject to invasions by double hairpin elements; (ii) genes of all three species contain numerous mobile group I introns, including one that is closest to an intron that invaded angiosperm mtDNAs; and (iii) at least one additional case of a mobile

element, characterized by a homing endonuclease insertion between partially duplicated genes [Paquin,B., Laforest,M.J., Forget,L., Roewer,I., Wang,Z., Longcore,J. and Lang,B.F. (1997) *Curr. Genet.*, 31, 380–395]. The combined mtDNA-encoded proteins contain insufficient phylogenetic signal to demonstrate monophyly of zygomycetes.

## INTRODUCTION

Fungi constitute a huge group of highly diverse organisms, including some of the most-studied and best-understood eukaryotic model systems: 'baker's yeast' (*Saccharomyces cerevisiae*), fission yeast (*Schizosaccharomyces pombe*), and the filamentous euascomycetes *Neurospora crassa* and *Aspergillus nidulans*. These species all belong to the Ascomycota. Substantially fewer scientific studies have been performed in members of the sister phylum Basidiomycota and very little is known in the remaining two phyla, Zygomycota and Chytridiomycota, often classified as 'lower fungi'. This expression is a taxonomically vague concept borrowed from Aristotle's philosophy, visioning directed evolution from the simple (primitive, low) to the highly complex. The misnomer is most evident in the 'higher' ascomycetes and basidiomycetes, which evolve toward microscopic, unicellular and genetically simplified yeast-like organisms in some lineages, and toward morphologically complex, gene-rich and biochemically versatile multicellular organisms in others.

Although the availability of complete nuclear and mitochondrial sequences of more than a dozen ascomycetes provides a

\*To whom correspondence should be addressed. Tel: +1 514 343 5842; Fax: +1 514 343 2210; [REDACTED]

<sup>†</sup>AY863211, AY863212, AY8632133, AY861439–AY861442

© The Author 2005. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oupjournals.org



strong basis for biochemical investigations, only a few complete mitochondrial sequences are known from chytrids, and none from zygomycetes, a situation that has motivated the work presented here. In fact, the number of zygomycete nuclear gene sequences (mostly rRNA sequences) is so limited that it is impossible to determine with confidence whether or not Zygomycota is a monophyletic taxon (1–3). The lack of resolution in these analyses is consistent with estimates that even the combined LSU- and SSU-rRNA would contain far too little information to resolve many fungal phylogenetic relationships with confidence (4). This dataset is at most sufficient for resolving fungal inferences below the phylum level (5,6).

Sequences of complete mtDNAs from several zygomycetes might be a first step in remediating this situation. Mitochondrial phylogenies can be based on up to 13 protein sequences, and have been shown to resolve deep divergences in the fungal and animal lineages (7–9). For instance, Alexopoulos *et al.* (10) indicate that 'additional study is needed to determine whether the class (Trichomycetes) is a monophyletic group belonging to Zygomycota, or merely a collection of orders grouped together on the basis of a unique shared habitat'. Molecular phylogenies based on rRNA sequences were successful in moving the order Amoebidiales away from zygomycetes, although they were then placed with mesomycetozoa rotists, whose phylogenetic affiliation was unresolved and controversial (11). Only subsequent analysis with multiple mitochondrial proteins placed this group at the base of the animal divergence, with high confidence (9). This example clearly illustrates the requirement of multi-gene datasets with at least several thousand amino acid positions, for resolution of trees at the kingdom level.

Furthermore, zygomycete mtDNAs are of considerable interest for comparative gene expression studies: our preliminary data indicated the presence of a mitochondrial *rnpB* gene, which encodes the RNA subunit of RNase P, the enzymatically active part of an endonuclease (ribonucleoprotein) responsible for tRNA maturation. In mitochondria, the size and sequence of the RNA subunit varies substantially, which has considerably complicated its identification. The gene is apparently absent from all completely sequenced basidiomycete and chytridiomycete mtDNAs, and has a patchy distribution in ascomycetes (12); to date, there are no published data on zygomycete mitochondrial RNase P (mtP-RNA).

To help remediate the lack of data for phylogenetic inferences, and to facilitate biochemical investigations and comparative mitochondrial genome analyses, we have sequenced mtDNAs from three distantly related zygomycetes, the Mucorales *Rhizopus oryzae*, the Mortierellales *Mortierella verticillata* and the Harpellales *Smittium culisetae*. In this article, we compare their mitochondrial genomes (gene content, gene organization, genetic code and widely conserved 3' RNA processing sites). We will then present secondary structure models and expression data for seven newly identified zygomycete mtP-RNAs. Finally, we will test whether or not zygomycetes are monophyletic, and provide evidence that the group I introns invasion of *cox1* gene in angiosperms originated in a zygomycete close to *Rhizopus*.

## MATERIALS AND METHODS

### Strains and culture conditions

The various zygomycete strains were obtained from Kerry O'Donnell (National Center for Agricultural Utilization Research, Peoria, IL; NRRL), R.W. Lichtwardt (Department of Botany, University of Kansas, Lawrence, KS; RWL) and Carolyn Babcock (Canadian Collection of Fungal Cultures, Ottawa; DAOM). All strains, *R.oryzae* (DAOM 148428, previously designated as *Rhizopus stolonifer*), *R.stolonifer* (DAOM 194667), *Rhizopus oligosporus* (NRRL 2710), *M.verticillata* (NRRL 6337), *Radiomyces spectabilis* (NRRL 2753), *Mucor mucedo* (NRRL 3635) and *S.culisetae* (strain 18-3; R.W. Lichtwardt), were grown in YG medium consisting of 0.5% yeast extract and 3% glycerol. Liquid cultures of 500 ml in 2 L Erlenmeyer flasks were grown at room temperature under gentle shaking (~100 r.p.m.).

### DNA and RNA extractions

For mtDNA and RNA extractions, the cells were broken mechanically, and a mitochondrial fraction was isolated by differential centrifugation (8). This fraction was lysed in the presence of 1% SDS and 100 µg/ml proteinase K at 50°C for 1 h, and after phenol–chloroform extraction, the nucleic acids were precipitated with ethanol. For RNA purifications, the high molecular weight RNA fraction was precipitated with 2 M LiCl, redissolved in RNase-free water and ethanol-precipitated. MtDNAs from all zygomycete strains were purified from total cellular nucleic acids by Cesium chloride/bisbenzimidate density gradient centrifugation.

### Cloning and sequencing of complete mtDNAs

Library construction and DNA sequencing followed previously published protocols (8). Briefly, mtDNAs were physically sheared by nebulization (13), and a size fraction of 1300–4000 bp was recovered after agarose gel electrophoresis. The DNA was incubated with a mixture of T7 DNA polymerase and *Escherichia coli* DNA polymerase I (the Klenow fragment) to generate blunt ends, and then cloned into the EcoRV cloning site of the phagemid pBFL6 (B. F. Lang, unpublished data). Recombinant plasmids containing mtDNA inserts were identified by colony hybridization using mtDNA as a probe. Clones contained in the random libraries were sequenced to ~8-fold coverage, and remaining gaps were closed by primer walking or sequencing of PCR-amplified DNA fragments. The expected quality of the sequenced mtDNAs is <1 error in 10 000 bp.

The mtDNA sequences of *R.oryzae*, *M.verticillata* and *S.culisetae* have been deposited in GenBank (accession nos AY863212, AY863211 and AY8632133, respectively).

### PCR amplification of *rnpB* genes

Mitochondrial *rnpB* genes of *R.stolonifer*, *M.mucedo*, *R.spectabilis* and *R.oligosporus* were PCR-amplified from ~100 ng of the respective mtDNAs in a 50 µl reaction mixture [200 µM dNTP, 2.5 mM MgCl<sub>2</sub>, 2 nM of primers, 5 µl of 10× buffer and 3 U of DNA polymerase mixture from the Expand high fidelity kit (Roche Catalog no. 1732650) and degenerate primers]. The annealing temperature of the PCR amplification was 50°C. Sequences of the degenerate primers are

5'-GTAATGGCAGCATACTAGACTCAT-3' and 5'-TTGAACTCCCAAGTTTTATGTATG-3'. The amplification products were cloned and sequenced. The *rnpB* sequences of *R.stolonifer*, *R.oligosporus*, *M.mucedo* and *R.spectabilis* have been deposited in GenBank (accession nos AY861439, AY861440, AY861441 and AY861442, respectively).

#### RNA circularization by ligation and RT-PCR

RNA ligation of mtP-RNAs, followed by RT-PCR amplification, was performed to determine the precise 5' and 3' termini of RNase P RNAs according to the previously published protocols (12). The primers are 5'-CTCTTATAGGATAATA-CAAAGTTG-3' and 5'-GGCCGAAGAATAAAGAGGGA-3' for *M.verticillata*, and 5'-ACCCTAATTTTCATTAGATATTT-3' and 5'-AATCCTTAGTAAGGATAGCTT-3' for *R.oryzae*.

#### RT-PCR of mtP-RNA from *R.oryzae*

Mitochondrial RNA of *R.oryzae* was treated with DNase I and extracted with phenol/chloroform, until no genomic DNA could be amplified by PCR using the mtP-RNA-specific primers 5'-TTCTTAGAGTTAAATAAGCC-3' and 5'-TTGGAGGAAAGTCCGGG-3'. Following these treatments, we amplified the mtP-RNA by RT-PCR as described above using a sample without reverse transcription as negative control. Following amplification and separation on a 0.8% agarose gel, the resulting DNA fragment was cloned and sequenced.

#### S1 mapping

DNA oligonucleotides and a 10 bp DNA ladder (Invitrogen 10821-015) were labeled at their 3' termini with ddATP-<sup>32</sup>P (Amersham PB10235) and terminal deoxynucleotidyl transferase (MBI Fermentas EP0161) according to the manufacturer's recommendations.

A total of 100 000 c.p.m. of gel-purified labeled oligonucleotides were hybridized to ~10 µg of total RNA, as described in the protocol by Hahn and Breeden ([http://www.fhrc.org/labs/hahn/methods/mol\\_bio\\_meth/s1\\_oligo\\_probe.html](http://www.fhrc.org/labs/hahn/methods/mol_bio_meth/s1_oligo_probe.html)), and S1 nuclease digestions were carried out at 37°C for 30 min, after the addition of 20 U of S1 nuclease and the buffer provided by the manufacturer (MBI Fermentas EN0321). The product was then ethanol-precipitated and dissolved in 4 µl of RNase-free water. An aliquot of 2 µl of the product was mixed with 2 µl of the dye solution provided with the 10 bp ladder (Invitrogen 10821-015), denatured at 75°C and loaded on a 7% polyacrylamide denaturing sequencing gel. The following oligonucleotides were used:

SSU-rRNA, 5'-AAATAAAGGGTTTAATATATTGGGAGGGACTTATTGTCCCCCGGTAATAACCATTCCAGC-CACTCGTCCCGAACGGCT-3'

*coxI* mRNA, 5'-CAGATTCTAAGGGGGTGTATATTA-TTAATTAATTAAGATTGAACTGGTAATGAATTTA-CAGTATGGA-3'.

#### Phylogenetic inference

Mitochondrial protein sequences from all completely sequenced zygomycetes, chytridiomycetes and basidiomycetes were included in phylogenetic inference (for species names and GenBank accession nos, see legend of Figure 5). Protein sequences from apocytochrome *b* (Cob), as well as

7 subunits of NADH dehydrogenase (Nad), 3 cytochrome *c* oxidase (Cox) subunits and 2 ATP synthetase (Atp) subunits were aligned with Muscle (14), concatenated and trimmed with Gblocks [using default parameters (15)] to remove ambiguously aligned regions. The resulting alignment contained 2890 aligned positions. Maximum likelihood (ML) phylogenies were inferred from this alignment using both PHYML (16) and IQPNNI (17); ML bootstrap support was determined based on 100 replicates, using both programs. All phylogenies were inferred using the JTT amino acid substitution model and gamma distribution ( $\Gamma$ ) correction for variation of rates across sites.

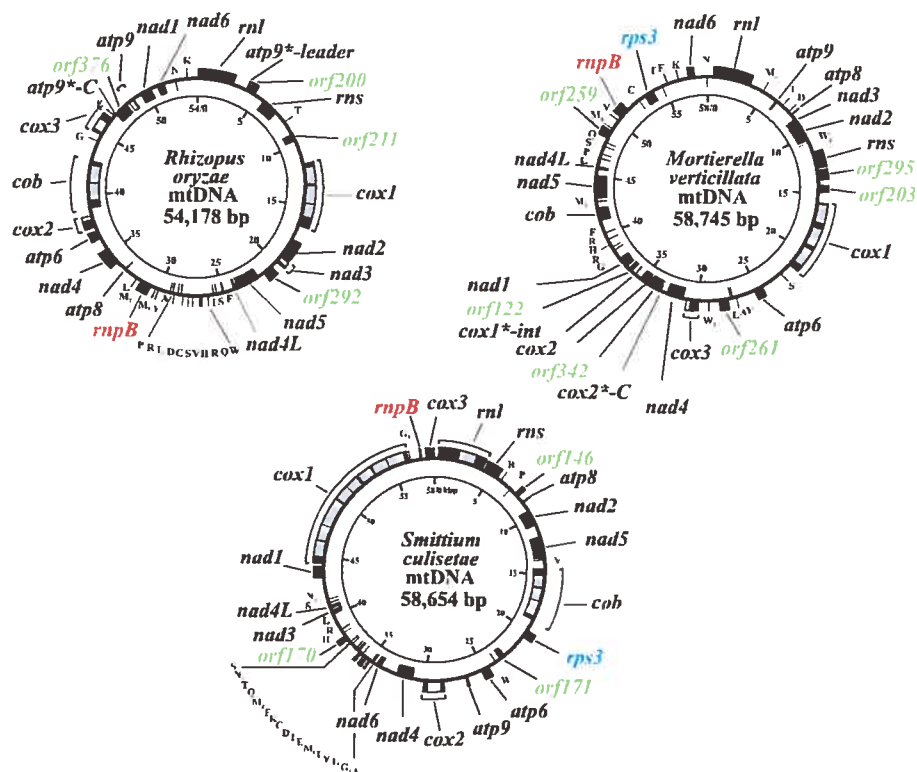
Phylogenetic inference based on concatenated datasets can lead to tree reconstruction artefacts, such as long-branch attraction, resulting from differences in relative evolutionary rates between genes [sometimes referred to as the covarion-like behavior of different genes (18)]. Ideally, during the ML tree search procedure, nuisance parameters, such as branch lengths and the  $\Gamma$  distribution shape parameter ( $\alpha$ ), could be optimized separately for each gene; however, this method results in the estimation of many more parameters, potentially more than can be statistically justified. For this reason, we partitioned the dataset into four functional categories: Atp (253 positions), Cob (361 positions), Cox (876 positions) and Nad (1400 positions). The additional parameters included in this partitioned dataset were justified using the  $\chi^2$  test ( $P < 0.0001$ ). Using this partitioned dataset, a topology was inferred from this alignment using MrBayes, with branch lengths and  $\alpha$  parameter unlinked across partitions (19). In addition, an ML tree was determined from this partitioned dataset using an adaptation of the method of (18). All possible tree topologies were generated, with constraints of groups that received at least 95% bootstrap support under ML, using both IQPNNI and PHYML (Supplementary Figure 1). Log-likelihoods were calculated separately for each partition using PHYML, under each of the topologies, and the sum over all partitions was calculated for each tree. The tree found to maximize the sum log-likelihood of the dataset was taken to be the ML tree (this method is referred to henceforth as separate analysis). Bootstrap support was also calculated using both of these methods, based on 100 replicates. Likelihood ratio tests were performed using genewise optimized site likelihoods, given the 99 tree topologies described in Supplementary Figure 1, and using Tree-Puzzle to generate sitewise likelihoods (20) along with CONSEL (21).

Additionally, phylogenetic analysis of closely related fungal, green algal and plant intronic open reading frames (ORFs) of the same *coxI* intron (for species names and GenBank accession nos, see legend of Figure 5) were carried out. Sequences were aligned using Muscle and trimmed with Gblocks, and the resulting alignment was manually refined. An ML phylogeny was inferred using IQPNNI, and ML bootstrap support was determined based on 100 replicates.

## RESULTS AND DISCUSSION

### Genes in addition to the standard fungal set, *rnpB* and *rps3*

The mitochondrial genomes of the zygomycetes *R.oryzae* (previously listed incorrectly as *R.stolonifer*), *M.verticillata*



**Figure 1.** Genomic maps of the mtDNAs of *R.oryzae*, *M.verticillata* and *S.culisetae*. The inner circle gives a scale in kilo base pair. The outer circle indicates the location of genes, exons (black) and introns plus intronic ORFs (gray). Names of ORFs, *rps3* and *rnpB* are colored to distinguish them from standard fungal genes (black).

and *S.culisetae* were completely sequenced. Like most other fungal mtDNAs, they map as circular molecules (Figure 1), although they are most likely organized as linear multimeric concatamers *in vivo*, as in other fungi (22). mtDNAs of the three species carry the basic fungal set of genes (Table 1), and encode a full set of tRNAs [only *trnI*(cau) is absent in *R.oryzae*], the RNA component of mitochondrial RNase P (*rnpB*) and a ribosomal protein (*rps3*; lacking in *R.oryzae*). MtDNA-encoded *rps3* has previously been identified in several ascomycetes and in one chytridiomycete (*Allomyces macrogynus*; Blastocladales), but not in other chytridiomycete orders, including Monoblepharidales, Spizellomycetales and Chytridiales. It has been proposed that this gene has been lost independently three times from opisthokont mitochondrial genomes: in the chytridiomycete lineage, in the animal lineage (23) and now also in one of the three zygomycetes presented here. We assume that, in all cases, *rps3* has been transferred to the nuclear genome, like other ribosomal protein genes missing in fungal mitochondrial genomes (24).

The sizes of zygomycete mtDNAs are within a close range of 54–58 kb (Figure 1). Genes are encoded on both strands, but are not as tightly packed as in animals and some ascomycetes: only 40.6% in *R.oryzae*, 43.1% in *M.verticillata* and 35.3% in *S.culisetae* are coding. Nonetheless, the coding regions of *nad2/nad3* and *nad4L/nad5* of *R.oryzae*, respectively, overlap by 1 nt (i.e. the last nucleotide position of the UAA stop codon of the upstream gene is the first nucleotide of the AUG start codon of the downstream gene; Figure 1). No

**Table 1.** Overview of gene, ORF and intron content in zygomycete mtDNAs

Genes and introns	<i>R.oryzae</i>	<i>M.verticillata</i>	<i>S.culisetae</i>
<i>rns, rnl</i>	■	■	■
<i>atp6,8,9</i>	■	■	■
<i>cob, cox1,2,3</i>	■	■	■
<i>nad1-6,AL</i>	■	■	■
<i>trnA-W</i>	24 [ <i>trnI</i> (cau) missing]	26	27
<i>rnpB</i>	■	■	■
<i>rps3</i>	□	■	■
Group I introns (intronic ORFs)	9(5)	4(3)	14(13)
Intron locations (number)	<i>cox1</i> (3), <i>cox2</i> (1) <i>cox3</i> (1), <i>cob</i> (2) <i>nad3</i> (1), <i>atp9</i> (1)	<i>cox1</i> (3), <i>cox3</i> (1)	<i>rnl</i> (1), <i>cox1</i> (9) <i>cox2</i> (1), <i>cob</i> (3)
Other ORFs	4	7	3

conservation of mitochondrial gene order is observed between these species.

### First steps toward a derived genetic code in zygomycetes

Both *R.oryzae* and the fast-evolving *S.culisetae* have retained the standard translation code for protein coding genes, a trait inherited from the eubacterial ancestors of mitochondria. However, *M.verticillata* reassigns two UGA 'stop' codons as tryptophan, once each in *nad3* and *nad4*. UGA(Trp) codons are also present in the *S.culisetae* intronic ORF283 and

ORF248, both encoding group I introns homing endonucleases of the LAGLI-DADG type. UGA(Trp) at amino acid position 237 of ORF248 is part of a distinctive, highly conserved sequence motif of this class of endonuclease, strongly suggesting its translation as tryptophan. It is possible that the presence of this UGA(Trp) is a vestige of horizontal intron transfer from a fungus adapted to this translation code. In fact, according to our phylogenetic analyses of intron endonucleases (Figure 5A), *S.culisetae* ORF248 is closely related to ORF313 of *Podospora anserina*, which makes preferential use of UGA(Trp) in its genes and intronic ORFs. Like in the fission yeast *S.pombe* and the basidiomycete *Schizophyllum commune*, the mtDNAs of *S.culisetae* and *M.verticillata* do not encode *trnW*(uca), which would efficiently recognize both UGA and UGG tryptophan codons. We assume that in all these cases, UGA codons are (albeit inefficiently) decoded by *trnW*(cca) (25,26). However, it cannot be excluded that, alternatively, the C in the wobble position of the anticodon is either modified or partially edited to allow efficient recognition of UGA(Trp) codons.

The zygomycete mtDNAs described here encode complete sets of tRNAs sufficient to recognize all encountered codons (for codon usage, see Supplementary Table 1S). *R.oryzae* does not have *trnI*(cau); however, ATA(Ile) codons are absent in standard mitochondrial genes, although they occur in intronic ORFs. Incidentally, a strikingly similar scenario exists in *Schizosaccharomyces octosporus*. It has been suggested (27) that either (i) the tRNA required for translation of ATA(Ile) is imported from the cytoplasm to recognize these codons or (ii) the intronic ORFs are neither translated nor required for intron splicing. A further explanation is that these codon positions are recognized by other tRNAs at low efficiency, resulting in amino acid misincorporation, which might be permissible at poorly conserved amino acid positions of proteins.

Whatever the mechanism of codon recognition, we suggest that such unexpected codon usage in intronic ORFs reflects horizontal intron transfer from species that are adapted to the use of UGA(Trp) and/or ATA(Ile). This codon usage is common in fungi and several other eukaryotic lineages.

#### Eubacteria-like mtP-RNAs in zygomycete mitochondria

Mitochondrial *rnpB* genes (encoding the mitochondrial RNA subunit of RNase P, mtP-RNA) were identified by *in silico* analysis in all three zygomycetes using the previously described procedures (12), and their RNA secondary structures were modeled by phylogenetic comparative analysis (Figure 2) (see also <http://megasun.bch.umontreal.ca/People/lang/rnpB/>). The presence of this gene in all three zygomycetes is striking, because outside fungi, *rnpB* is only present in the green alga *Nephroselmis olivacea* (28) and in various jakobids (B. F. Lang and E. Seif, unpublished data), including *Reclinomonas americana* (29). Within fungi, it is only present in some ascomycetes, but absent in basidiomycetes and chytridiomycetes (12).

The inferred size of the *S.culisetae* mtP-RNA is 145 nt, close to the shortest known example (140 nt, in *Saccharomycopsis fibuligera*) (30). Most remarkably, the highly reduced mtP-RNA structures of *S.culisetae* and budding yeasts are almost identical (Figure 2), perfectly matching the minimum consensus secondary structure of fungal mtP-RNAs

(12). In contrast, *rnpB* from *M.verticillata* and *R.oryzae* are the largest genes of this class ever identified (980 and 830 bp, respectively), even larger than *rnpB* genes studied in bacteria (31). In addition, the zygomycete mitochondrial RNA secondary structures are the most bacteria-like among fungi.

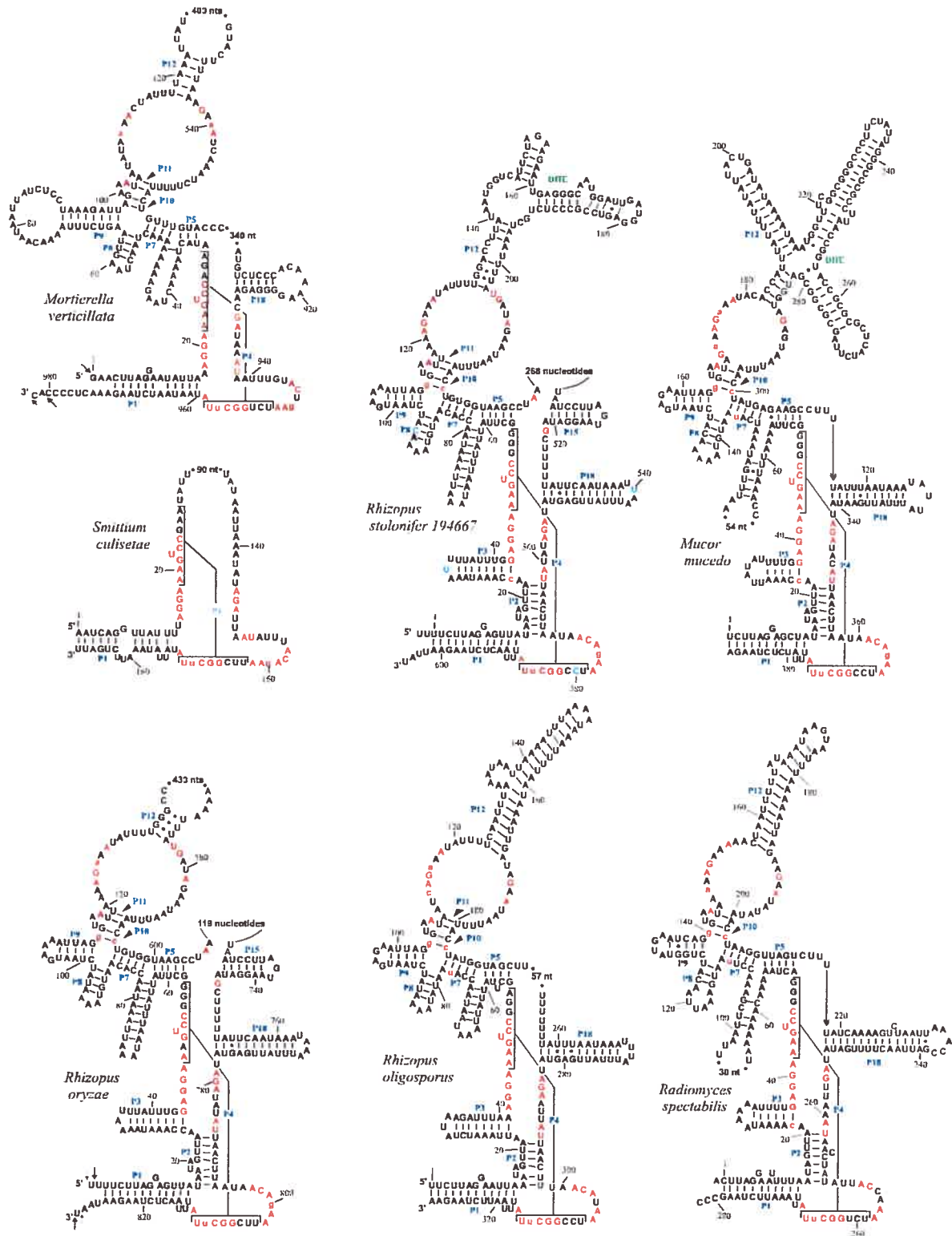
To verify the expression of the *R.oryzae* and *M.verticillata* genes, we determined their precise 5' and 3' ends by sequencing RT-PCR products of circularized mtP-RNAs (Figure 2). The 3' end of *M.verticillata* is 9 nt longer than anticipated, elongated by a cytidine-rich stretch of sequence. A similar extension is located at the 3' terminus of *Schizosaccharomyces octosporus* mtP-RNA (12) and downstream of protein coding genes in a variety of fungi (see below; Figure 3). The 5' end of *M.verticillata*, and both the 5' and 3' termini of *R.oryzae* mtP-RNAs, match the proposed secondary structure model and reveal little heterogeneity of mtP-RNA termini (Figure 2).

In evolutionary terms, zygomycete mtP-RNA structures cover an unprecedented wide range of intermediate stages in loss of RNA structural elements. The mtP-RNAs from *R.oryzae* and closely related species have the most bacteria-like secondary structures, containing almost all structural elements of the bacterial minimum secondary structure consensus (32). They are followed by *M.verticillata*, whose structure closely resembles the more derived mtP-RNA of the ascomycete *Taphrina deformans* (12). Finally, the tiny, yeast-like mtP-RNA molecule of *S.culisetae* has no P2 helix, which is otherwise almost omnipresent in mtP-RNAs. Note that this helix is also absent in *M.verticillata*, potentially indicating its loss in a common ancestor.

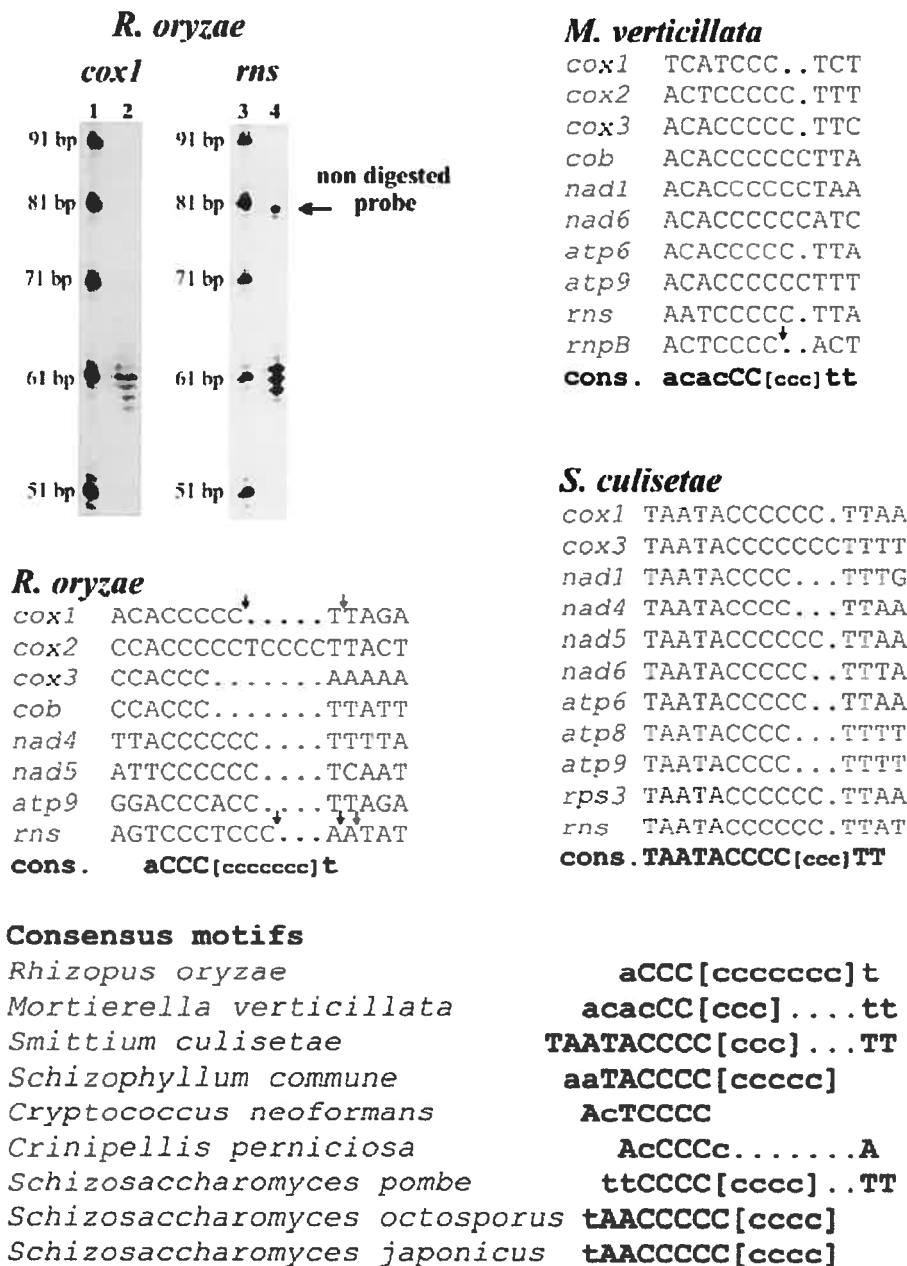
The most bacteria-like fungal mtP-RNA secondary structure is that of *R.oryzae*, only lacking P13, P14 and P19, which are otherwise only present in the protist mtP-RNAs of *N.olivacea* and *R.americana*. The large size of the *R.oryzae* and *M.verticillata* mtP-RNAs is due to insertions at the J5-15 and J5-18 junctions, respectively, and in the P12 helix. In order to determine whether these regions are conserved structural elements or more variable insertion elements or introns, we amplified the cDNA sequence of the *R.oryzae* mtP-RNA, and the genomic sequences from the closely related Mucorales *M.mucedo*, *R.spectabilis*, *R.oligosporus* and *R.stolonifer*. Figure 2 shows that the insertion sequences can be folded into double hairpin structures [DHEs (33)]. Because the cDNA sequence of the *R.oryzae* mtP-RNA is identical to the genomic sequence, these variable regions are not introns. Furthermore, the insertion points and sizes of these regions vary substantially, indicating that they have been acquired recently and independently. Their presence in mtP-RNAs pinpoints structural regions that are likely not critical for RNase P activity. An analogous situation has been described in some cyanobacteria, where P-RNAs contain short tandem repeats that increase the length of helix P12. Site-directed mutagenesis experiments have shown that this helix is not required for catalytic activity *in vitro*, implying that it is also unlikely to be crucial for the *in vivo* activity (34).

#### Conserved C-rich motifs in mRNAs and SSU-rRNA

Small C-rich clusters are present downstream of mitochondrial protein- and SSU-rRNA coding regions, in all three species (Figure 3). The consensus sequence of this motif varies only slightly among zygomycetes. It also exists in



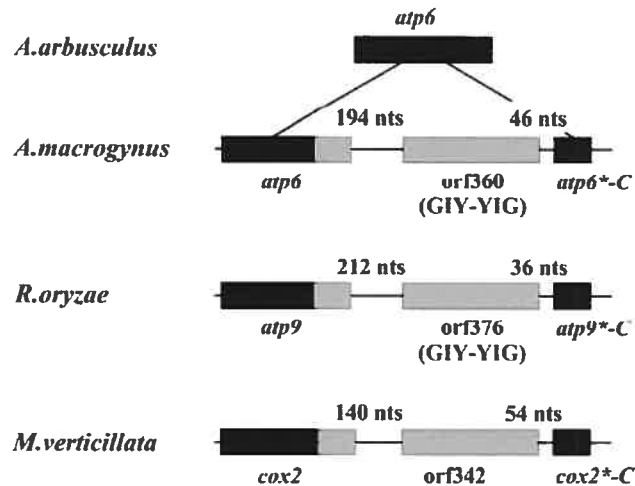
**Figure 2.** Secondary structure models for mtP-RNAs from *R.oryzae*, *R.stolonifer* 194667, *R.oligosporus*, *R.spectabilis*, *M.mucedo*, *S.culisetae* and *M.verticillata*. Positions in red are invariant in the minimum bacterial consensus (32); uppercase letters in the mtP-RNAs indicate 100%, lowercase 90%, conservation of the minimum bacterial consensus sequence. The arrows pinpoint experimentally determined termini; arrow length is proportional to the percentage of molecules ending at a defined position. Double hairpin elements are named in green. The few nucleotides colored blue in the *R.stolonifer* mtP-RNA model are different in its close relative *R.oryzae*.



**Figure 3.** 3' RNA processing motifs in zygomycetes, basidiomycetes and fission yeasts. The 3' termini of the *R. oryzae* mitochondrial SSU rRNA and of *cox1* mRNA were determined by nuclease S1 assays and run on a sequencing gel against a commercial 10 bp ladder (Invitrogen 10821-015) that was 3' labeled with ddATP (<sup>32</sup>P). For experimental details see Materials and Methods. The positions of 3' termini for both molecules are indicated in the derived consensus sequences by arrows. A small fraction of the undigested form of the SSU rRNA probe is apparent on the gel. In the lower part of the figure, additional, similar motifs in fission yeasts and basidiomycetes are presented. Uppercase letters indicate 100% conservation and lowercase letters correspond to at least 60% nucleotide conservation. Lowercase Cs between brackets indicate the C-clusters of variable length.

basidiomycetes and in fission yeasts, pointing to a shared function. Mapping of the 3' end of the *M. verticillata* mtP-RNA (which also terminates with this motif as discussed above; Figure 2), of *cox1* and the SSU-rRNA of *R. oryzae* shows that these C-rich motifs are the site of 3' RNA processing, and are retained in the mature RNA molecules. The presence of ragged 3' ends (Figure 3) indicates that these are

generated by an exonuclease trimming mechanism. Similar observations have been made in fission yeasts (12,27). This mechanism resembles that of nuclear and viral RNAs terminating in polyuridine motifs that serve as a binding site for the La protein implicated in RNA protection against exonucleases [reviewed in (35); homologs of La are known from a range of fungi and animals].



**Figure 4.** Schematic view of *atp6* regions of *A. macrogynus* and *A. arbusculus* (36), *atp9* of *R. oryzae* and *cox2* of *M. verticillata*. Coding sequences are enclosed in boxes and intergenic spacers are represented by a thick line. Black boxes indicate sequences present before the invasion by the corresponding ORF. Gray boxes represent ORFs and newly acquired sequences.

#### More instances of mobile endonuclease elements?

As reported earlier, the chytridiomycete fungus *A. macrogynus* has a novel mtDNA insertion element whose sequence is absent in a close relative, *Allomyces arbusculus*. This element consists of a duplicate C-terminus of a foreign *atp6* gene, plus an ORF encoding an endonuclease that is responsible for its mobility (36). The inserted *atp6* portion is fused in phase with the resident gene (Figure 4), reconstituting an obviously functional hybrid gene of standard length and amino acid conservation. In fact, it has been shown that homing endonucleases can be mobile even independent of introns and genes (37), and that they are capable of carrying genetic material from one site to another when they migrate.

Intriguingly, similar gene hybrids are present in the mtDNAs of *R. oryzae* (*atp9\*-C*) and *M. verticillata* (*cox2\*-C*), including ORF376 (*R. oryzae*) and ORF342 (*M. verticillata*) (Figure 3). ORF342 has no significant similarity to known endonucleases, but ORF376, like the mobile element endonuclease of *A. macrogynus* (36), encodes a protein related to homing endonucleases of the GIY-YIG type (38,39). The same structural organization is seen in the *atp6*-ORF360-*atp6\*-C* gene region in *A. macrogynus*, *atp9*-ORF376-*atp9\*-C* in *R. oryzae* and *cox2*-ORF342-*cox2\*-C* in *M. verticillata* (Figure 3), indicating the presence of similar mobile elements. However, contrary to the *A. macrogynus* case, we currently do not have biochemical evidence for the endonucleolytic activity. Both the *cox2\*-C* and *atp9\*-C* fragments of the two zygomycetes encode C-terminal ends that are 100% identical at the amino acid level, suggesting that the source of potential transfers are closely related zygomycete species.

#### Lateral transfer of a group I intron from zygomycete to angiosperm mitochondria

A significant portion of the genomes described here is occupied by introns (*R. oryzae* 15.8%; *M. verticillata* 8.6%; and

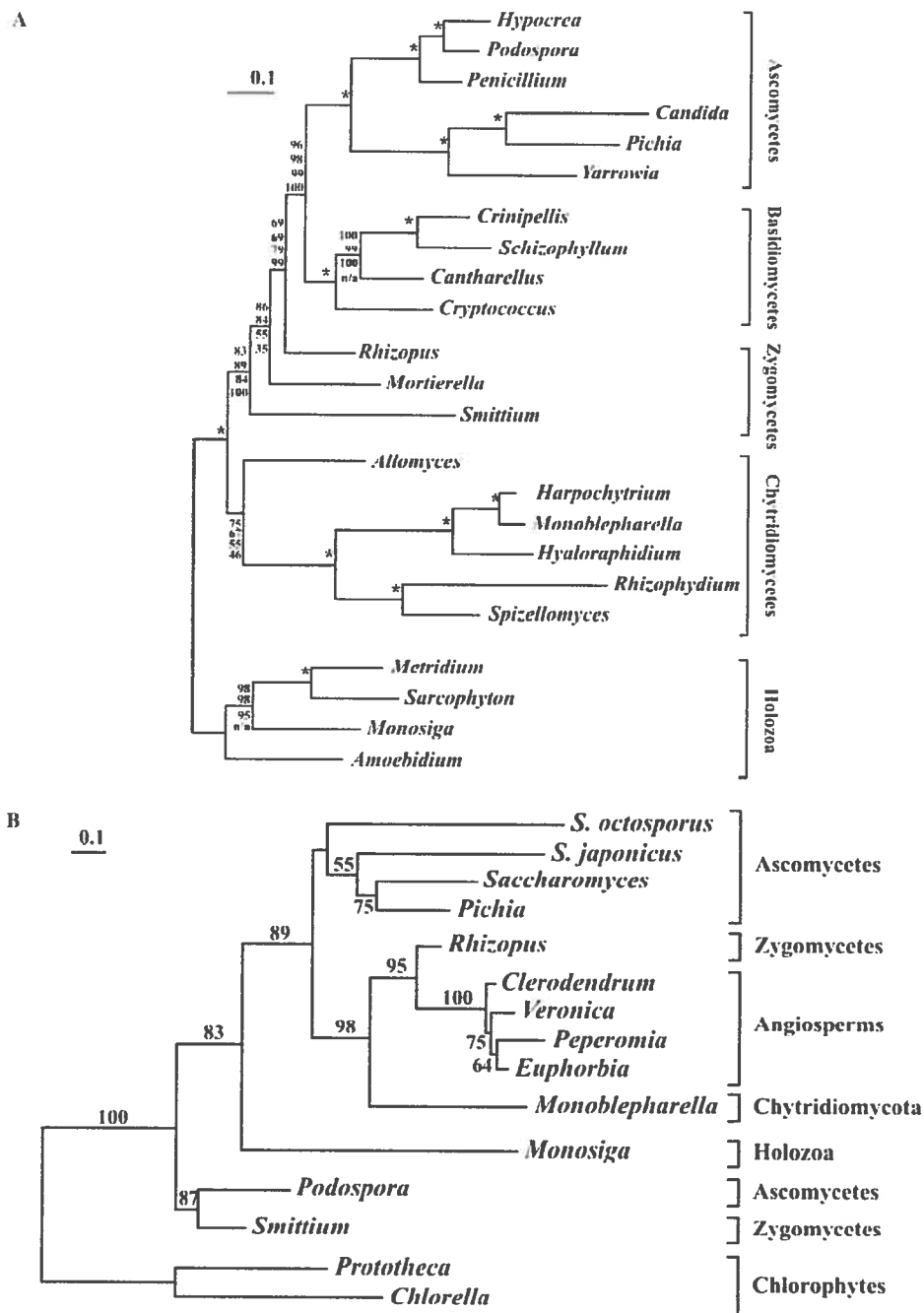
*S. culisetae* 27.4%). With 14 introns, the mtDNA of *S. culisetae* contains the largest number, 9 of which are located in the *cox1* gene (Table 1). Here, all the identified zygomycete introns are of group I and 22 contain intronic ORFs: sixteen of the LAGLI-DADG type and six of the GIY-YIG type. In *R. oryzae*, we identified one intron [*cox1*-i1(ORF305)], which is most similar to introns inserted at the same positions of angiosperm *cox1* genes (highest BLAST expect value of  $e^{-114}$  with *Philodendron oxycardium*, *Lamium* sp. and *Malpighia glabra*). Because this is the only group I intron in vascular plant mtDNAs, it has most likely been acquired by lateral transfer. The hypothesis of recent horizontal transfer of this intron from a fungal donor to flowering plants (40) is further based on the incongruence between the intron and organismal phylogenies, and its closer phylogenetic relationship to a fungal intron than to those of *Marchantia* and *Prototheca* (40–43). However, because of the high mobility of this intron and the possibility of multiple lateral transfers, the published phylogenetic inferences have to be interpreted with caution.

The presence of additional, highly similar introns in three fungi, *R. oryzae*, *S. culisetae* (this paper) and *Monoblepharella*15 (8), allows us to more rigorously test the hypothesis of Vaughn *et al.* by phylogenetic analysis (Figure 5A). Our analysis reveals that ORFs from *R. oryzae*, *Monoblepharella*15 and the three angiosperms group together with high support (98%). ORF305 from *R. oryzae* is the closest and most similar relative of the angiosperm ORFs, suggesting that the fungal donor of the group I intron and its resident ORF was a zygomycete. This scenario is biologically meaningful, because symbiotic mycorrhizal zygomycetes live in close association with most plants. However, note that ORF248 from *S. culisetae* branches with ORF313 from *P. anserina*. This observation, along with the presence of a UGA codon specifying tryptophan in ORF248 (see above), suggests a second case of lateral transfer, this time from a euascomycete close to *P. anserina*, into a zygomycete.

#### Phylogenetic analysis with standard mitochondrial proteins: are zygomycetes paraphyletic?

The availability of complete mtDNAs from three distant zygomycetes provides an opportunity for testing the monophyly of Zygomycota. Tree topologies were inferred from a 2890 position concatenated protein alignment using two standard ML-based methods, MrBayes and separate ML analysis. The results of these analyses are summarized in Figure 5B, which presents the tree produced by IQPNNI, along with bootstrap values from all methods. All methods produced similar results, except that separate analysis recovered monophyly of *S. culisetae* and *M. verticillata* (81% bootstrap support). The latter topology was also found in the 'credible set' in three independent runs of MrBayes (5.3, 44.7 and 99.0% posterior probability, respectively; MrBayes bootstrap support for this grouping was 45%). Although these data support both of these topologies, it is interesting that the monophyly of *S. culisetae* and *M. verticillata* is better supported under more sophisticated (yet statistically valid) models.

Although this phylogeny is generally robust, bootstrap support values indicate two major areas of uncertainty: the position of *A. macrogynus* and the relative branching positions of the zygomycetes. Indeed, also the approximately unbiased



**Figure 5.** (A) Fungal phylogeny based on multiple proteins. Mitochondrion-encoded protein sequences from *Harpochytrium* sp. 94, *Crinipellis perniciosa*, *Cryptococcus neoformans*, *Hypocrea jecorina*, *Amoebidium parasiticum*, *Spizellomyces punctatus*, *Yarrowia lipolytica*, *Monosiga brevicollis*, *R.oryzae*, *Rhizophydium* sp. JEL136, *Penicillium marneffei*, *Pichia canadensis*, *Cantharellus cibarius*, *Sarcophyton glaucum*, *S.culisetae*, *Monoblepharella15*, *Metridium senile*, *A.macrogynus*, *Hyaloraphidium curvatum*, *Candida albicans*, *P.anserina*, *S.comune* and *M.verticillata* were aligned, concatenated and trimmed. Phylogenies were inferred from the resulting 2890 character alignment using four different methods. Shown here is the ML tree inferred using IQPNNI, along with bootstrap support values from PHYML, IQPNNI, MrBayes and separate ML analysis, in order from top to bottom, based on 100 replicates. Nodes with 100% bootstrap support using all methods are indicated by an 'asterisk'. Clearly, both the position of *A.macrogynus* and the branching order of the zygomycetes remain unclear, although the topology is robust overall. (B) Phylogeny of intronic ORFs. Sequences of intronic ORFs inserted in *cox1* genes were obtained from the following species: ORF305 from *R.oryzae*, ORF248 from *S.culisetae*, ORF318 from *Monoblepharella15* (NP\_803527), ORF333 from *Schizosaccharomyces japonicus* (NP\_705621), ORF317 from *S.octosporus* (NP\_700369), ORF313 from *P.anserina* (NP\_074934), ORF319 from *Pichia canadensis* (NP\_038209), ia4 from *S.cerevisiae* (AAB21126), ORF251 *Chlorella vulgaris* (T07187), ORF234.2 *Prototheca wickerhamii* (NP\_042245), ORF280 from *Peperomia obtusifolia* (AAB86934), ORF279 from *Veronica catenata* (CAA11340) and ORF277 from *Marranta leuconeura* (CAA11350). Sequences were aligned as described in Materials and Methods, and a phylogeny was inferred by ML. Only bootstrap support values >50% are shown. This tree robustly supports the monophyly between sequences from *Monoblepharella15*, *R.oryzae*, and the angiosperms, strongly suggesting horizontal intron transfer between these two groups.



likelihood ratio test suggests these same problems. When site-wise likelihoods are calculated separately for each functional class, the confidence set contained a total of 17 tree topologies that failed to reject the data ( $P < 0.05$ ). Among these topologies, *A.macrogynus* branches immediately above, immediately below or monophyletic with the chytridiomycetes, and *R.oryzae* and *M.verticillata* are found to be either monophyletic or paraphyletic (with either species branching more deeply than the other). Although *S.culisetae* and *M.verticillata* are monophyletic in the best tree under this model of separately optimized functional classes, all topologies in which all three zygomycetes are monophyletic are rejected. It is worth noting that several obvious wrong positions for *S.culisetae* were also observed (e.g. as most ancestral among fungi, or among the chytridiomycetes), suggesting that the accelerated evolutionary rate in this species causes a long-branch attraction artefact (44). Similar results are obtained when sitewise likelihoods were calculated from the fully concatenated alignment. Clearly, these data are insufficient to resolve the phylogeny of the zygomycetes, most likely because these three species diverge deeply within fungi, and at relatively short distance from each other. In such a situation, two strategies can be used to resolve the dilemma, addition of more zygomycete and neighboring fungal lineages or addition of more sequence per species. As our protein dataset is already based on complete mtDNAs, this latter strategy implies resorting to expressed sequence tag and/or nuclear genome sequences, a currently ongoing project.

## SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Gertraud Burger (Université de Montréal) for helpful comments on the manuscript, Michael W. Gray and Murray Schnare for help in the analysis of the genes for small and large rRNAs, Kerry O'Donnell (National Center for Agricultural Utilization Research, Peoria, IL), R. W. Lichtwardt (Department of Botany, University of Kansas, Lawrence, KS) and Carolyn Babcock (Canadian Collection of Fungal Cultures, Ottawa) for supply of fungal strains. We further acknowledge the excellent technical assistance of Z. Wang. Salary and interaction support from the Canadian Institutes of Health Research (MOP 42475; B.F.L.), the Canadian Institute for Advanced Research (CIAR; B.F.L.), the Nova Scotia Health Research Foundation (J.L.), and the 'Bourses d'Excellence biT' (CIHR; Y.L.) and supply of laboratory equipment and informatics infrastructure by Genome Quebec/Canada is gratefully acknowledged. Funding to pay the Open Access publication charges for this article was provided by the CIHR.

## REFERENCES

- Voigt,K., Cigelnik,E. and O'Donnell,K. (1999) Phylogeny and PCR identification of clinically important Zygomycetes based on nuclear ribosomal-DNA sequence data. *J. Clin. Microbiol.*, **37**, 3957-3964.
- Tanabe,Y., O'Donnell,K., Saikawa,M. and Sugiyama,J. (2000) Molecular phylogeny of parasitic Zygomycota (Dimargaritales, zoopagales) based on nuclear small subunit ribosomal DNA sequences. *Mol. Phylogenet. Evol.*, **16**, 253-262.
- Voigt,K. and Wostemeyer,J. (2000) Reliable amplification of actin genes facilitates deep-level phylogeny. *Microbiol. Res.*, **155**, 179-195.
- Berbec,M.L., Carmean,D.A. and Winka,K. (2000) Ribosomal DNA and resolution of branching order among the ascomycota: how many nucleotides are enough? *Mol. Phylogenet. Evol.* **20**, 337-344.
- Kurtzman,C.P. (2003) Phylogenetic circumscription of Saccharomycetes, Kluyveromycetes and other members of the Saccharomycetaceae, and the proposal of the new genera Lachancea, Nakasomyces, Naumovia, Vanderwaltozyma and Zygotulasporea. *FEMS Yeast Res.*, **4**, 233-245.
- James,T.Y., Porter,D., Leander,C.A., Vilgalys,R. and Longcore,J.E. (2000) Molecular phylogenetics of the Chytridiomycota supports the utility of ultrastructural data in chytrid systematics. *Can. J. Bot.*, **78**, 226-350.
- Leigh,J., Seif,E., Rodriguez,N., Jacob,Y. and Lang,B.F. (2003) Fungal evolution meets fungal genomics. In Arora,D.K. (ed.), *Handbook of Fungal Biotechnology*. 2nd edn. Marcel Dekker Inc., NY, 145-161.
- Bullerwell,C.E., Forget,L. and Lang,B.F. (2003) Evolution of monolepharidalean fungi based on complete mitochondrial genome sequences. *Nucleic Acids Res.*, **31**, 1614-1623.
- Lang,B.F., O'Kelly,C., Nerad,T., Gray,M.W. and Burger,G. (2002) The closest unicellular relatives of animals. *Curr. Biol.*, **12**, 1773-1778.
- Alexopolous,C.J., Mims,C.W. and Blackwell,M. (1996) *Introductory Mycology*. John Wiley & Sons, NY.
- Ustinova,J., Krienitz,L. and Huss,V.A. (2000) *Hyaloraphidium curvatum* is not a green alga, but a lower fungus; *Amoebidium parasiticum* is not a fungus, but a member of the DRIPs. *Protist.*, **151**, 253-262.
- 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.
- Okpodu,C.M., Robertson,D., Boss,W.F., Togasaki,R.K. and Surzycki,S.J. (1994) Rapid isolation of nuclei from carrot suspension culture cells using a BioNebulizer. *BioTechniques*, **16**, 154-159.
- Edgar,R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.
- Castresana,J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.*, **17**, 540-552.
- Guindon,S. and Gascuel,O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.*, **52**, 696-704.
- Vinhle,S. and Von Haeseler,A. (2004) IQPNNI: moving fast through tree space and stopping in time. *Mol. Biol. Evol.*, **21**, 1565-1571.
- Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., Gordon,P., Duruffe,L., Gaasterland,T., Lopez,P., Muller,M. et al. (2002) The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. *Proc. Natl Acad. Sci. USA*, **99**, 1414-1419.
- Ronquist,F. and Huelsenbeck,J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572-1574.
- Schmidt,H.A., Strimmer,K., Vingron,M. and von Haeseler,A. (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, **18**, 502-504.
- Shimodaira,H. and Hasegawa,M. (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics*, **17**, 1246-1247.
- Bendich,A.J. (1996) Structural analysis of mitochondrial DNA molecules from fungi and plants using moving pictures and pulsed-field gel electrophoresis. *J. Mol. Biol.*, **255**, 564-588.
- Bullerwell,C.E., Burger,G. and Lang,B.F. (2000) A novel motif for identifying *rps3* homologs in fungal mitochondrial genomes. *Trends Biochem. Sci.*, **25**, 363-365.
- Lang,B.F., Gray,M.W. and Burger,G. (1999) Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.*, **33**, 351-397.
- Bullerwell,C.E., Leigh,J., Seif,E., Longcore,J.E. and Lang,B.F. (2003) Evolution of the fungi and their mitochondrial genomes. In Arora,D.K. and Khachatourians,G.G. (eds), *Applied Mycology and Biotechnology*. Elsevier Science, Amsterdam, Vol. 3, 133-159.

26. Paquin,B., Laforest,M.J., Forget,L., Roewer,I., Wang,Z., Longcore,J. and Lang,B.F. (1997) The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. *Curr. Genet.*, **31**, 380–395.
27. Bullerwell,C.E., Leigh,J., Forget,L. and Lang,B.F. (2003) A comparison of three fission yeast mitochondrial genomes. *Nucleic Acids Res.*, **31**, 759–768.
28. Turmel,M., Lemieux,C., Burger,G., Lang,B.F., Otis,C., Plante,I. and Gray,M.W. (1999) The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*. Two radically different evolutionary patterns within green algae. *Plant Cell*, **11**, 1717–1730.
29. 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.
30. Wise,C.A. and Martin,N.C. (1991) Dramatic size variation of yeast mitochondrial RNAs suggests that RNase P RNAs can be quite small. *J. Biol. Chem.*, **266**, 19154–19157.
31. Brown,J.W. (1999) The Ribonuclease P Database. *Nucleic Acids Res.*, **27**, 314.
32. Siegel,R.W., Banta,A.B., Haas,E.S., Brown,J.W. and Pace,N.R. (1996) *Mycoplasma fermentans* simplifies our view of the catalytic core of ribonuclease P RNA. *RNA*, **2**, 452–462.
33. Paquin,B., Laforest,M.J. and Lang,B.F. (2000) Double-hairpin elements in the mitochondrial DNA of *Allomyces*: evidence for mobility. *Mol. Biol. Evol.*, **17**, 1760–1768.
34. Vioque,A. (1997) The RNase P RNA from cyanobacteria: short tandemly repeated repetitive (STRR) sequences are present within the RNase P RNA gene in heterocyst-forming cyanobacteria. *Nucleic Acids Res.*, **25**, 3471–3477.
35. Wolin,S.L. and Cedervall,T. (2002) The La protein. *Annu. Rev. Biochem.*, **71**, 375–403.
36. Paquin,B., Laforest,M.J. and Lang,B.F. (1994) Interspecific transfer of mitochondrial genes in fungi and creation of a homologous hybrid gene. *Proc. Natl Acad. Sci. USA*, **91**, 11807–11810.
37. Eddy,S.R. and Gold,L. (1992) Artificial mobile DNA element constructed from the EcoRI endonuclease gene. *Proc. Natl Acad. Sci. USA*, **89**, 1544–1547.
38. Burger,G. and Werner,S. (1985) The mitochondrial *URF1* gene in *Neurospora crassa* has an intron that contains a novel type of URF. *J. Mol. Biol.*, **186**, 231–242.
39. Michel,F. and Cummings,D.J. (1985) Analysis of class I introns in a mitochondrial plasmid associated with senescence of *Podospora anserina* reveals extraordinary resemblance to the *Tetrahymena* ribosomal intron. *Curr. Genet.*, **10**, 69–79.
40. Vaughn,J.C., Mason,M.T., Sper-Whitis,G.L., Kuhlman,P. and Palmer,J.D. (1995) Fungal origin by horizontal transfer of a plant mitochondrial group I intron in the chimeric *Cox1* gene of *Peperomia*. *J. Mol. Evol.*, **41**, 563–572.
41. Adams,K.L., Clements,M.J. and Vaughn,J.C. (1998) The *Peperomia* mitochondrial *cox1* group I intron: timing of horizontal transfer and subsequent evolution of the intron. *J. Mol. Evol.*, **46**, 689–696.
42. 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. USA*, **95**, 14244–14249.
43. Palmer,J.D., Adams,K.L., Cho,Y., Parkinson,C.L., Qiu,Y.L. and Song,K. (2000) Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proc. Natl Acad. Sci. USA*, **97**, 6960–6966.
44. Felsenstein,J. (1978) Cases in which parsimony and compatibility methods will be positively misleading. *Syst. Zool.*, **27**, 401–410.

### **Article 3**

#### **Fusions of mitochondrial with *E. coli* ribonuclease P RNAs are catalytically active.**

Elias Seif, Alexandre Cadieux and B. Franz Lang

Manuscript in preparation.

Canadian Institute for Advanced Research, Centre Robert Cedergren, Département de Biochimie, Université de Montréal, 2900 Boulevard Edouard-Montpetit, Montréal, Québec, H3T 1J4, Canada.

#### Authors contribution :

Elias Seif : developed detailed experimental design and did most of the experiments, data analysis and preparation of manuscript and figures.

Alexandre Cadieux : did part of the experimental work.

Franz Lang : helped design the experiments, and prepare the manuscript and figures.

## Abstract

Mitochondrially encoded RNA subunits of mitochondrial RNase P (mtP-RNA) have been identified in jakobid flagellates such as *Reclinomonas americana*, the prasinophyte *Nephroselmis olivacea*, and in numerous ascomycete and zygomycete fungi. Whereas the secondary structures of ascomycete mtP-RNAs are highly reduced compared to their bacterial counterparts, the jakobid, prasinophyte and zygomycete mtP-RNAs retain most features of the minimum bacterial structure, and are potentially active *in vitro* in the absence of a protein subunit. Here we describe the precise mapping of *R. americana* and *N. olivacea* mtP-RNA ends, and the comparative structural analysis of six additional jakobid RNAs. We further investigate ribozyme activities of mtP-RNAs and find that even the most bacteria-like molecules of jakobids are inactive *in vitro*. However, when certain domains of jakobid and *N. olivacea* RNAs are exchanged with *E. coli* P-RNA sequences, they regain the capacity to specifically cleave tRNA precursors. *In vitro* mutagenesis of these hybrid mtP-RNAs provides evidence for the importance of various structural elements for the ribozyme activity, including a remnant P3 pairing in *S. ecuadoriensis*, and GNRA tetra-loops in L14 of *J. libera* and in L18 of both *Jakoba libera* and *R. americana* #50394. We will discuss why native mtP-RNAs fail to catalyze 5' tRNA processing *in vitro*, and discuss reasons for the evolutionary tendency to reduce if not eliminate the mitochondrial RNA subunit of RNase P.

## Introduction

Ribonuclease P (RNase P) is an endonuclease that removes 5' leader sequences from tRNA precursors in eubacteria, archaeobacteria, eukaryotes and organelles. In most studied cases, RNase P is a ribozyme composed of an RNA component (P-RNA), and one or several protein subunits (P-protein; (Altman et al., 1995; Frank & Pace, 1998; Hartmann & Hartmann, 2003)). Possible exceptions are in spinach chloroplasts (Thomas et al., 1995; Thomas et al., 2000b) and in human mitochondria (Rossmannith & Karwan, 1998; Puranam & Attardi, 2001; Rossmannith & Potuschak, 2001) where protein-only ribonucleases are suspected.

When present, the P-RNA subunit is indispensable for the endonucleolytic activity of RNase P, as shown by RNase treatment in a broad variety of cases including eubacteria, archaeobacteria, eukaryotes, plastids and mitochondria (e.g., (Stark et al., 1978; Gardiner & Pace, 1980; Kline et al., 1981; Hollingsworth & Martin, 1986; Jayanthi & Van Tuyle, 1992; Baum et al., 1996; Lee et al., 1996a; Han et al., 1998; Vainauskas et al., 1998; Stathopoulos et al., 2001)). Furthermore, all tested eubacterial and some archaeobacterial (Methanobacteria, Thermococci, and Halobacteria) P-RNAs, catalyse 5' tRNA processing *in vitro* in the absence of the P-protein (Guerrier-Takada et al., 1983; Gardiner et al., 1985; Pannucci et al., 1999; Wagner et al., 2001). Although eukaryotic P-RNAs lack such an *in-vitro* activity, the mechanism of cleavage of the RNase P holoenzyme is RNA-based, similar to that of eubacteria (Thomas et al., 2000a).

P-RNA from *E. coli* (M1 RNA) was the first multiple-turnover ribozyme to be identified (Guerrier-Takada et al., 1983), and is by far the best-studied P-RNA. Secondary and tertiary structure models have been proposed for this RNA, based on phylogenetic comparative analyses, enzymatic probing, and cross linking experiments (Brown et al., 1996; Westhof et al., 1996; Haas & Brown, 1998; Massire et al., 1998). All P-RNAs except those of some mitochondria (for details see below) contain two main folding domains (Loria & Pan, 1996), the catalytic domain (“C” domain, including helices P1-6 and P15-18), and the specificity domain (“S” domain; Fig. 1). The “C” domain extends from the 5’ and 3’ ends of the RNA and includes P1-6, 15-18 helices, loops (L), and single stranded junctions (J; Figure 1B). Two main functions were assigned to the C-domain: (i) binding of  $Mg^{2+}$  ions that are essential for catalysis (Christian et al., 2000; Kaye et al., 2002; Tanaka et al., 2004a), which is mainly accomplished by P4 and the single stranded region in the junction between P3 and P4 (J3/4); and (ii) binding of the 3’ terminal RCCA of tRNA precursors by the GGU motif that is present in the internal loop between P15 and P16 of the *E. coli* P-RNA (L15/16; (Kirsebom & Svard, 1994; Kufel & Kirsebom, 1998; Brannvall et al., 2004)). The S-domain, which includes P7-14 helices, loops and junctions, is mainly responsible for interactions with the T stem-loop region of pre-tRNA substrates. Recently, the structure of the S-domain of the *Thermus thermophilus* P-RNA (which is most similar to *E. coli* P-RNA, defined as type A) was resolved at 2.9 Å (Krasilnikov et al., 2004). Although the overall folding of the “S” domains of *T. thermophilus* and *Bacillus subtilis* (type B) differ, they share an opening structure

formed by P9, P10 and P11 and the non-helical module J11/12-J12/11 (Krasilnikov et al., 2003). It has been suggested that this common structure accommodates the pre-tRNA (specifically T-stem and T-loop) and orients it for interaction with the catalytic domain (LaGrandeur et al., 1994; Odell et al., 1998). It is known that tetraloops of the GNRA type confer thermodynamic stability of short RNA double helices (Heus & Pardi, 1991); further, that they stabilize tertiary structures of large RNAs by intramolecular interactions (Jaeger et al., 1994; Murphy & Cech, 1994; Cate et al., 1996). Phylogenetic comparative studies and cross-linking studies in *E. coli* suggest that C- and S- domains interact through tertiary interactions of GNRA tetraloops, between L18/P8, and L9/P1 (Brown et al., 1996; Massire et al., 1997). Similar interactions occur within the S-domain, between L14/P8 and L13/P12. GNRA motifs in *E. coli* P-RNA contribute to the catalytic activity as shown by *in-vitro* mutagenesis. A significant decrease of *E. coli* RNase P activity was noted when certain GNRA motifs were exchanged by UNGC tetraloop motifs (Pomeranz Krummel & Altman, 1999), which are otherwise common in RNA molecules (Proctor et al., 2002).

Genes encoding P-RNAs (*rnpB*) have been determined in hundreds of species, and secondary structure models were proposed for P-RNAs from the three major domains of life, including plastids (Brown, 1999) and mitochondria (Seif et al., 2003; Seif et al., 2005). Comparative studies suggest that all P-RNAs are homologs, sharing up to five sequence motifs (CRI-V; (Chen & Pace, 1997)), a cage-shaped pseudo-knot (P1-P4; (Forster & Altman, 1990)), a four-way junction (P7-10), and a large internal loop comprised between P11, P12, P13 and P14 (J11/12, J12/13, J14/11) (Chen &

Pace, 1997). The most notable departures from this established consensus structure exist in mitochondria of budding yeasts and the zygomycete *Smittium culisetae*, where P-RNA secondary structures are limited to P1 and P4 (Lee et al., 1996b; Seif et al., 2003; Seif et al., 2005; Talla et al., 2005). Mitochondrial P-RNAs (mtP-RNAs) from other fungal mitochondria are less derived, intermediate between those of yeast mitochondria and Eubacteria. As mitochondria originated from an  $\alpha$ -Proteobacterium, the reduced forms of fungal mtP-RNAs are best explained by loss of structural elements from an  $\alpha$ -Proteobacterium-like P-RNA (Seif et al., 2003). Indeed, mitochondria of the bacterivorous protist *Reclinomonas americana* (member of the jakobid flagellates) have the most bacteria-like mitochondrial genome based on various features (Lang et al., 1997; Jacob et al., 2004), including its mtP-RNA. It contains all structural elements present in the minimum eubacterial consensus (Lang et al., 1997), p13, P14, plus a P19 helix otherwise present in the mtP-RNA of the prasinophyte *Nephroselmis olivacea* (Turmel et al., 1999; Seif et al., 2005), and in  $\alpha$ -proteobacterial P-RNAs (Fig. 1A, B). Whereas mtP-RNAs have been well characterized in various fungi (Shu & Martin, 1991; Wise & Martin, 1991; Lee et al., 1996b; Seif et al., 2003; Seif et al., 2005), little is known about the more eubacteria-like mtP-RNAs from jakobids or from *N. olivacea*. In particular, it is currently unknown whether or not these mtP-RNAs have a catalytic activity *in vitro*, in the absence of a protein subunit.

In this study, we show that the mtP-RNAs from *R. americana* and *N. olivacea* are transcribed and processed according to predictions based on RNA secondary



structure models. To gain information on the evolution of the jakobid mtP-RNAs, we identified six *rnpB* genes in newly sequenced mitochondrial genomes. We further investigate the *in vitro* catalytic activity of several mtP-RNAs and show that the native RNA molecules are inactive, but that the activity can be partially restored in chimeras with *E. coli* P-RNA, and augmented through a variety of mutations. The reasons for the lack of activity of native mtP-RNAs will be discussed.

## **MATERIAL AND METHODS**

### **Cell cultures and nucleic acid extractions**

An axenic strain of *N. olivacea* (NIES-484; a prasinophyte green alga distributed by the National Institute for Environmental Forum, Japan) was kindly provided by Drs. M. Turmel and C. Lemieux. Cells from this species were grown as previously described (Turmel et al., 1999). The jakobids *Reclinomonas americana* (New Zealand strain, ATCC 50394), *Reclinomonas americana* (ATCC 50283), *Reclinomonas americana* (ATCC 50284), *Reclinomonas* sp. (ATCC 50633) and *Seculamonas ecuadoriensis* (ATCC 50688) were grown in WCL medium, whereas the two jakobids adapted to saline growth conditions, *Jakoba libera* (ATCC 50422) and *Jakoba bahamiensis* (ATCC 50695), were cultivated in F/2 medium. The recipes for these growth media are found at <http://megasun.bch.umontreal.ca/People/lang/FMGP/methods.html>. For large-scale cultures, jakobids were incubated in 2.5-liter culture flasks under gentle shaking, at 24°C for several days. They were daily fed with live *Enterobacter aerogenes* bacteria (ATCC 13048). When reaching the late logarithmic growth phase (after ~ 5 days), cells were harvested by centrifugation and the nucleic acids were extracted as described (Seif et al., 2003).

### **MtP-RNA 5' and 3' end mapping**

About 10 µg of mitochondrial RNA from were used to determine the precise 5' or 3' ends of mtP-RNAs, by RNA ligation, RT-PCR amplification and cDNA sequencing,

as previously described (Seif et al., 2003). Primer sequences for *R. americana* are 5' TAAATGAGTCCAGACTTTCCT 3' and 5' GAGCAAAACCATTA GGTGCT 3'; and 5' CTATTCATACTAGCTCT 3' and 5' GTGCAAGAAT TTAATTCGCT 3' for *N. olivacea*.

### **Construction of chimerical P-RNAs**

The *rnpB* genes of *E. coli*, *R. americana* #50394, *J. libera*, *S. ecuadoriensis*, *R. oryzae*, and *N. olivacea* were cloned in pFBS (a modified bluescript vector) and used as templates for the following PCR amplifications. Primer sequences for construction of chimerical P-RNAs and for other mutagenesis experiments are available as supplementary information. The cloning vector (pFBS) is amplified together with one of the two (mitochondrial or *E. coli*) domains, which permits direct cloning of chimerical *rnpB* genes, after DNA ligation.

For construction of type I chimeras (see Fig. 2A for the precise extensions of the mitochondrial and bacterial sequences), mitochondrial P-RNAs were amplified to include P1-4, P18, 19 plus J15/18. The respective *E. coli* P-RNA (M1 RNA) domain was amplified to include P5-P17. For type II chimera (see Figure 3A) the amplified mitochondrial domain contains P5, 7-15; and the *E. coli* domain of P1-4, J15/18 and P18. All amplification products were end-repaired (unmodified T7 DNA polymerase from New England Biolabs, #M02565S, plus the Klenow fragment from Roche, #1008404) at 14°C for 30 minutes. After denaturation of the polymerases at 65°C for 10 minutes, the DNA fragments were purified on low melting agarose gel (FMC

Bioproducts #50102). The PCR products including the vector sequence was dephosphorylated with alkaline phosphatase (Boehringer Mannheim #713023), followed by extraction with phenol/chloroform and ethanol precipitation, whereas the corresponding fragment containing the other half of the P-RNA sequence was phosphorylated with T4 polynucleotide kinase (USB #27073602). The two fragments were then ligated at 14°C overnight, transformed, and cloned with the expected sequence identified by DNA sequencing.

#### **In vitro synthesis of RNA and RNase P activity assays**

The internally radio-labelled precursor tRNA substrates, and the unlabelled P-RNAs were synthesized *in vitro* and purified as described previously (Jacob et al., 2004). DNA templates for *in vitro* synthesis of RNA were prepared by PCR amplification of respective P-RNA sequences, by inclusion of a T7 RNA polymerase promoter in one of the two primers used for amplification. Purified RNAs were dissolved in 1x PA buffer (50 mM Tris, pH 7.5, 100 mM NH<sub>4</sub>Cl, 10 mM MgCl<sub>2</sub>), heated for 5 min at 65°C and then allowed to slowly cool to 37 °C.

The catalytic activity of P-RNAs was tested using 1,000 cpm of substrate under two conditions: (i) as used for testing the activity of bacterial P-RNAs (Guerrier-Takada et al., 1986), including 1x PA buffer (50 mM Tris, pH 7.5, 100mM NH<sub>4</sub>Cl, 10 mM MgCl<sub>2</sub>), 100 mM MgCl<sub>2</sub> and 4% PEG (polyethylene glycol), or (ii) as used for testing the archaebacterial activity (50 mM Tris pH 8, 300 mM MgCl<sub>2</sub>, 2.5 M ammonium acetate, 0.1% SDS, and 0.05% Nonidet P-40; (Pannucci et al., 1999)).

For assays including the *E. coli* P-protein, 5nM of the protein was added to 0.5 nM of P-RNA in binding buffer (20mM Hepes-KOH, pH 8, 400mM NH<sub>4</sub>OAc, 10mM Mg(OAc)<sub>2</sub>, 5% glycerol), the mixture was heated for 5 minutes at 37°C, after which 1,000 cpm of tRNA precursor (substrate) was added and the mix incubated at 37°C overnight.

Reaction products were loaded on 7% denaturing polyacrylamide gels as described (Jacob et al., 2004), gels were exposed for 2 hours on a Biorad molecular imaging screen K (#1707841), and images were developed using the Biorad molecular imager FX and analyzed with the “Quantity One” software.

#### **Genetic complementation assay**

BL21 cells that carry a *rnpA* mutation in the *E. coli* chromosome (C5 R46H; (Guerrier-Takada et al., 1995)) and that contain  $\lambda$  prophages encoding the T7 RNA polymerase, were used for the genetic complementation assay (the permissive and non-permissive temperatures for this *E. coli* strain are 30°C and 43°C, respectively). The mutant cells were transformed with plasmid constructions encoding respective mtP-RNA constructions, and grown overnight at 30°C on LB media with 5  $\mu$ g/ml tetracycline (*rnpA* chromosomal selection) and 100  $\mu$ g/ml ampicillin (plasmid selection). Several independent transformants were then grown on two different (pre-warmed) agar plates at either 43°C or 30°C. The genetic complementation of the various P-RNAs was assessed by examining colony growth after overnight incubation at the non-permissive temperature (Baer et al., 1989). Cells transformed with vector

alone and a vector encoding M1 RNA served as the negative and positive controls, respectively.

## RESULTS

### ***R. americana* and *N. olivacea* mtP-RNAs of predicted size are expressed *in vivo*.**

To verify that the mtP-RNAs from *N. olivacea* and *R. americana* #50394 are transcribed, and processed *in vivo* at the sites as predicted (Lang et al., 1999; Turmel et al., 1999), we mapped their precise extremities by ligation of the 5' and 3' ends of mitochondrial RNAs, RT-PCR, and sequencing of the amplification products. In both cases, the 5' and 3' extremities matched the inferred secondary structure models, with little heterogeneity of the RNA extremities (Fig. 1A, B).

### **Bacteria-like mtP-RNAs in jakobids**

In order to obtain data for the comparative analysis of the bacteria-like mtP-RNAs in the jakobid *R. americana* #50394 (Lang et al., 1997), we sequenced mitochondrial DNAs (mtDNAs) of various other jakobids including *R. americana* #50283, *R. americana* #50284, *Reclinomonas* sp. #50633, *Seculamonas ecuadoriensis*, *Jakoba libera*, and *Jakoba bahamiensis* (Burger G and Lang BF; unpublished data). By searching universally conserved motifs of P-RNAs, and the universally conserved pseudo-knot (P4) (Forster & Altman, 1990; Chen & Pace, 1997; Seif et al., 2003), we identified *rnpB* genes in all jakobid mtDNAs. Their predicted mtP-RNA secondary structures are similar to those of *R. americana* #50394 (Lang et al., 1997), containing all structural elements of the minimum bacterial secondary structure (Brown et al., 1994). The only exception is in *S. ecuadoriensis*, which has a highly reduced P3 (Fig. 1B). Following a comparative analysis of all jakobid mtP-RNAs, we can update the

previously published secondary structure model for *R. americana* #50394 (Lang et al., 1997) by extending the P18 helix, and by assigning a GNRA tetra-loop to L18 that is conserved among jakobids except in *S. ecuadoriensis* (Fig. 1B). Also note that the P19 helix, which is a landmark feature of few eubacterial lineages including  $\alpha$ -Proteobacteria, is also present in jakobid (except *J. bahamiensis*), and in the prasinophyte mtP-RNAs.

### **Bacteria-like mtP-RNAs are catalytically inactive *in vitro***

To determine if any of the bacteria-like mtP-RNAs are catalytic *in vitro*, we have chosen seven candidates that contain all or most of the structural elements as defined in the minimum bacterial P-RNA structure (Fig. 1A), including the jakobids *R. americana* #50394, *S. ecuadoriensis*, *J. libera*, *J. bahamiensis*, the prasinophyte *N. olivacea*, and the zygomycetes *Rhizopus oryzae* and *Rhizopus oligosporus* (Seif et al., 2005). Respective mtP-RNAs were synthesized *in vitro*, and their catalytic activity tested *in vitro* using the RNA precursors of mitochondrial tRNA<sup>Pro</sup> from *R. americana* and tRNA<sup>Lys</sup> from *R. oryzae* as substrates. None of the mtP-RNAs exhibits a catalytic activity, under any testing conditions (ribozyme concentration, low and high salt, addition of PEG, etc). The activity of the *R. americana* #50394 mtP-RNA was further tested *in vitro*, in the presence of the *E. coli* P-protein (the protein subunit of *E. coli* RNase P; results not shown), and in an *in vivo* complementation assay in the thermo-sensitive *E. coli* A49 mutant, which has a defective P-protein (Baer et al., 1989). Also these additional assays turned out negative (data not shown).



### **MtP-RNA domains can be complementary to *E. coli* P-RNA sequences**

To understand why bacteria-like mtP-RNAs are not active *in vitro*, i.e., in order to locate specific structural elements that are lacking for activity, we constructed chimerical P-RNAs by combining half-molecules of *E. coli* with mitochondrial P-RNAs.

Type I chimeras contain the most conserved structural elements of mtP-RNAs including P1-4, 18, and 19; the remaining parts of the molecule are from *E. coli* (Fig. 2A). Type I chimeras have various degrees of *in-vitro* activities, highest in *J. libera* and followed in order by *R. americana* #50394, *S. ecuadoriensis*, and *N. olivacea* (Fig. 2B,C). The chimera formed with the *R. oryzae* mtP-RNA is inactive. In type II chimeras, the mitochondrial and *E. coli* portions of the P-RNA molecule are reversed with respect to type I chimeras (Fig. 3A). The *R. americana* #50394 P-RNA has the highest catalytic activity followed by *J. libera* and *S. ecuadoriensis*; *N. olivacea* is inactive (Fig 2B,C). Type II chimeras are inactive at lower salt concentrations (e.g., 100mM NH<sub>4</sub>Cl, 110 mM MgCl<sub>2</sub>, 50mM Tris, pH 7.5) and require ~10 times higher ribozyme concentrations to achieve levels of activity comparable to type I, at higher salt concentrations (Fig. 2B,C).

When incubated in the presence of the *E. coli* P-protein (C5), the catalytic activity of the type II chimera of *R. americana* #50394 is strongly increased, but only slightly in *J. libera* (Figure 3C). Other type II chimeras are unaffected by the addition of the protein subunit, as are all type I chimeras (data not shown).

### **Mutation of conserved GNRA tetra-loops in jakobid mtP-RNAs reduce catalytic activity**

Comparative sequence analysis suggests the conservation of GNRA tetra-loops in L18 of *Reclinomonas* and *Jakoba* species, and in L14 of *Jakoba* species (Fig. 1B), which is consistent with evidence from  $\alpha$ -Proteobacteria (Brown, 1999) and *E. coli* (Pomeranz Krummel & Altman, 1999). In order to test the impact of these motifs on the catalytic activity of mtP-RNAs, these tetra-loops were mutated in the corresponding chimerical P-RNAs. According to the results with *E. coli* P-RNA, this should lead to the loss of the catalytic activity at low (< 20 mM) MgCl<sub>2</sub> concentrations (Pomeranz Krummel & Altman, 1999). Indeed, mutations of the tetra-loops in L18 from GUAA to AUAA, in type I chimeras of *J. libera* and *R. americana* #50394, have no visible effect on catalytic activities at 100 mM MgCl<sub>2</sub> or higher (data not shown). However, both mutants fail to cut pre-tRNA<sup>Lys</sup> at 30mM MgCl<sub>2</sub> (Fig. 4A). Similarly, a mutation of the tetra-loop in L14 of *J. libera* from GAGA to GAGG (applied to the type II chimera) drastically decreases the catalytic activity even at 300mM MgCl<sub>2</sub> (Fig. 4B).

### **Functionally important P3 remnant in mtP-RNA of *S. ecuadoriensis***

P3 is conserved in all bacterial P-RNAs, varying in length between 11 and 65 nucleotides (Tanaka et al., 2004b), and is also present in jakobid mtP-RNAs except for *S. ecuadoriensis*, which has a potential remnant of this helix with a single A-U

pair (Fig. 1B). To verify its significance, the A-U pair was mutated to either A-A or C-G (Fig. 5B). At 300 mM MgCl<sub>2</sub>, no significant difference in catalytic activity is observed. However, at 100 mM the catalytic activity decreased more than twofold for the AxA mutant, but did not affect the mutation in which A-U was replaced with a C-G base pair (Fig. 5). These results support the contribution of P3, or at least a similar structure with a single Watson-Crick base pair, to the catalytic activity of mtP-RNA of *S. ecuadoriensis*.

#### **Other structural elements with impact on the catalytic activity of mtP-RNAs**

In *E. coli*, the presence of P6, a GGU motif in L15/16 (positions 292-294) have been shown to have an impact on the catalytic activity of the P-RNA, (Haas et al., 1991; Darr et al., 1992; Kirsebom & Svard, 1994). To evaluate the importance of these structural elements (lacking in all mtP-RNAs), the domains P5, P6, P15, P16 from *E. coli* were grafted into the *R. americana* #50394 wild type mtP-RNA. The addition of these elements was insufficient to produce a catalytically active RNA (data not shown).

In addition, we tested the impact of correcting a deviation in the consensus sequence of CRV of *R. americana* #50394 (mutation of G292 to A), and a deletion of P19, which is neither present in the *E. coli* P-RNA, nor in the mtP-RNA of *J. bahamiensis*. The deletion of P19 from the respective *R. americana* #50394 chimera P-RNA causes a minor (20%) decrease of catalytic activity (Table 1). In contrast, introduction of a GNRA motif in L9 of the *R. americana* #50394 type II chimera

increases the activity (50%), as does the replacement of A292 by a G in the universally conserved motif CRV in the type I chimera (370% at low  $\text{MgCl}_2$  concentrations; see Table 1).

Finally, as mitochondrial genes are A+T rich, we verified if an increase of the G+C content in helical regions of mtP-RNAs might stabilize the RNA secondary structure, thereby increasing the catalytic activity. In one experiment, two A-U pairs in P4 of a type I chimera were replaced with G-C pairs (Fig. 1B, Table 1), resulting in a 320 % increase of the catalytic activity at a lower salt concentration (100mM  $\text{MgCl}_2$ , 100 mM  $\text{NH}_4\text{Cl}$ ), but no significant change under higher salt conditions (300mM  $\text{MgCl}_2$ , 2.5M ammonium acetate).

## DISCUSSION

### Mitochondrial P-protein is essential for mtRNase P activity

Six new jakobid mitochondrial *rnpB* genes have been identified that encode mtP-RNAs with bacteria-like secondary structures similar to that of *R. americana* #50394. All tested jakobid, *N. olivacea* and fungal mtP-RNAs are catalytically inactive *in vitro*, a result that is not surprising because eukaryotic, chloroplast and most archaeobacterial P-RNAs require a protein subunit for activity, and because mtP-RNAs are more A+T rich than most bacterial P-RNAs and will less efficiently fold into a catalytically active structure. Therefore, we assume that jakobid, like all other mtP-RNAs (e.g., yeast or *A. nidulans*; (Lee et al., 1996a; Lee et al., 1996b; Groom et al., 1998)), require a more substantial participation of one or more P-proteins to produce an active RNase P.

Although the *E. coli* P-protein does not restore the activity of the native *R. americana* #50394 mtP-RNA, it significantly enhances the activity of a jakobid mitochondrial/*E. coli* P-RNA chimeras, suggesting little specificity of the protein subunit. This view is in line with the moderate conservation of eubacterial P-proteins (Gopalan et al., 1997), and the lack of conservation between the known yeast mitochondrial P-protein (Stribinskis et al., 2001) and its eubacterial counterparts. To address these questions in more detail, it is paramount to identify mitochondrial P-proteins in species other than yeast, and that have mtP-RNAs such as various fungi (Seif et al., 2003; Seif et al., 2005) or the jakobid flagellates.

### **Both mitochondrial and bacterial P-RNA sequences interact with the *E. coli* P-protein**

*R. americana* #50394 and *J. libera* type II P-RNA chimeras have low catalytic activities that significantly increase after the addition of the *E. coli* P-protein, whereas P-RNA chimeras containing *S. ecuadoriensis* and *N. olivacea* sequences remain inactive.

Investigations of bacterial P-RNAs indicate the existence of several P-Protein – P-RNA interactions. In *B. subtilis* and *E. coli*, both C- and S-domains contribute to protein binding according to RNase T1 protection or Dimethyl Sulfate (DMS) and Fe(II)-EDTA footprinting experiments (Vioque et al., 1988; Talbot & Altman, 1994; Loria et al., 1998; Biswas et al., 2000). These results are supported by nucleotide analogue interference mapping (NAIM) in *B. subtilis* (Rox et al., 2002), which predict the presence of three major domains of P-protein interaction, two of which located in the C domain. This might explain why in our experiments only type II chimeras are activated by the *E. coli* P-protein.

### **Secondary structure conservation strongly affects the activity of chimerical P-RNAs**

The bacteria-like mtP-RNAs of jakobids and *N. olivacea* have retained structural features that produce catalytically active *E. coli* - mtP-RNA chimeras of varying degrees. The efficiency of complementation appears to correlate with the overall degree of secondary structure conservation, in comparison to the *E. coli* structure, and

other structural features such as conserved tetra-loops. We suspect that the high activity of the type I chimera of *J. libera* is due to the relatively high G+C content of helical regions of its mtP-RNA, a view that is confirmed by the increase of the catalytic activity of the type I chimera of *R. americana* #50394 in which two A-U pairs are exchanged by G-C pairs (Table I).

The catalytic activity of type I chimeras is an order of magnitude higher than those of type II chimeras. Again, this can be explained by the different levels of secondary structure conservation of the two mtP-RNA half-molecules. In type I chimeras, the mitochondrial domain is best conserved, both among mtP-RNAs, and in comparison to eubacterial P-RNAs, whereas type II chimeras include the more variable mitochondrial S-domain. This region of mtP-RNAs is also more easily lost during the course of mitochondrial evolution (Seif et al., 2003).

#### **Possible tertiary interactions in mtP-RNAs**

The presence of GNRA-type motifs in L18 and L14 of mtP-RNAs are possibly involved in long range interactions with P8, like what has been proposed to exist in *E. coli* (Brown et al., 1996). Similar GNRA interactions predicted for group I and II self splicing introns (Michel & Westhof, 1990; Costa et al., 1997), ribosomal RNAs (Gutell, 1996) and bacterial RNase P RNAs (Brown et al., 1996) have allowed to establish the rules that govern GNRA loop interactions with RNA helices. The last adenine of a GNRA motif (A4) interacts invariably with a C-G base pair, and the base pair following this C-G interacts with the nucleotide at the third position of the tetra-

loop (R3), co-varying according to the nature of the purine. An A at the third position (A3) is predicted to interact preferably with a C-G, with some tolerance for other base pairs (Michel & Westhof, 1990), and a G at the third position (G3) is predicted to interact primarily with a U-A, and with minimal tolerance for A-U, and no tolerance for C-G or G-C. Following these rules, we expect L18 to perfectly interact with P8 in mtP-RNAs of *R. americana* #50394, *Reclinomonas* sp. #50633 and *R. americana* #50284. In these cases, the L18/P8 interactions are shown by broken lines in Fig. 1. In *R. americana* #50283 and *J. libera* these tertiary interactions are also possible, but not in *J. bahamiensis*. To test these and other possible tertiary interactions, *in vitro* assays of native jakobid mtP-RNAs have to be developed, which will require the identification and purification of the respective mitochondrial P-proteins that are expected to be more efficient in reconstitutions of the ribozyme activity *in vitro* than the *E. coli* P-protein.

### Conclusions

Bacteria-like mtP-RNAs are not catalytic *in vitro*, unlike their bacterial counterparts, but the conservation of some structural features that are required for activity has been demonstrated in chimeras with *E. coli* P-RNA. Our investigations indicate that the lack of activity of native jakobid mtP-RNAs is due to the cumulative effect of (i) their relatively high A+U content, disfavoured thermodynamic stability compared to G+C rich bacterial P-RNAs; (ii) the absence of P6 and of certain GNRA motifs affecting their ability to form long-range interactions; (iii) the absence of the GGU motif in



L15/L16, disfavouring substrate binding; and (iv) deviations from the eubacterial consensus sequences that have a significant impact on the catalytic activity of mtP-RNAs, as demonstrated for CRV of *R. americana*. Obviously, any of these factors alone does not explain the lack of activity of mtP-RNAs, since relatively A+U rich bacterial P-RNAs and a 'micro P-RNA' lacking some of the GNRA motifs and P6 (Siegel et al., 1996) are catalytically active, and because minor deviations from the consensus sequences are tolerated in bacterial P-RNAs (Rasmussen & Nolan, 2002). In addition, binding of the 3' terminal RCCA of tRNA by L15 is not essential for catalytic activity as shown in *E. coli* (Guerrier-Takada et al., 1984), and in Cyanobacteria (Pascual & Vioque, 1999).

It is obvious that imperfections of mtP-RNAs accumulate in the course of evolution, which must be compensated by substantial changes to the protein moiety of mitochondrial RNase P. The identification of mitochondrial P-proteins will be an important next step in understanding the function and evolution of mitochondrial RNase P.

**Acknowledgements:**

We wish to thank Dr. Sidney Altman (Yale University) and Dr. Cecilia Guerrier-Takada (Yale University) for technical advice and for providing *E. coli* P-protein and the A49 mutant strain. We further acknowledge the excellent technical assistance of Jean-Francois Bouffard and Lise Forget. Salary and interaction support from the Canadian Institutes of Health Research (MOP 42475; BFL), the Canadian Institute for Advanced Research (CIAR; BFL), and supply of laboratory equipment and informatics infrastructure by Genome Quebec/Canada is gratefully acknowledged.

**References**

- Altman S, Kirsebom LA, Talbot S. 1995. Recent studies of RNase P. In: Soll D, RajBhandry U, eds. *tRNA Structure, Biosynthesis, and Functions*. Washington DC: American Society for Microbiology Press. pp 67-78.
- Baer MF, Wesolowski D, Altman S. 1989. Characterization in vitro of the defect in a temperature-sensitive mutant of the protein subunit of RNase P from *Escherichia coli*. *J Bacteriol* 171:6862-6866.
- Baum M, Cordier A, Schon A. 1996. RNase P from a photosynthetic organelle contains an RNA homologous to the cyanobacterial counterpart. *J Mol Biol* 257:43-52.
- Biswas R, Ledman DW, Fox RO, Altman S, Gopalan V. 2000. Mapping RNA-protein interactions in ribonuclease P from *Escherichia coli* using disulfide-linked EDTA-Fe. *J Mol Biol* 296:19-31.
- Brannvall M, Kikovska E, Kirsebom LA. 2004. Cross talk between the +73/294 interaction and the cleavage site in RNase P RNA mediated cleavage. *Nucleic Acids Res* 32:5418-5429.
- Brown JW. 1999. The Ribonuclease P Database. *Nucleic Acids Res* 27:314.
- Brown JW, Haas ES, Gilbert DG, Pace NR. 1994. The Ribonuclease P database. *Nucleic Acids Res* 22:3660-3662.
- Brown JW, Nolan JM, Haas ES, Rubio MA, Major F, Pace NR. 1996. Comparative analysis of ribonuclease P RNA using gene sequences from natural microbial populations reveals tertiary structural elements. *Proc Natl Acad Sci U S A* 93:3001-3006.
- Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, Cech TR, Doudna JA. 1996. Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* 273:1678-1685.
- Chen JL, Pace NR. 1997. Identification of the universally conserved core of ribonuclease P RNA. *RNA* 3:557-560.

- Christian EL, Kaye NM, Harris ME. 2000. Helix P4 is a divalent metal ion binding site in the conserved core of the ribonuclease P ribozyme. *RNA* 6:511-519.
- Costa M, Deme E, Jacquier A, Michel F. 1997. Multiple tertiary interactions involving domain II of group II self-splicing introns. *J Mol Biol* 267:520-536.
- Darr SC, Zito K, Smith D, Pace NR. 1992. Contributions of phylogenetically variable structural elements to the function of the ribozyme ribonuclease P. *Biochemistry* 31:328-333.
- Forster AC, Altman S. 1990. Similar cage-shaped structures for the RNA components of all ribonuclease P and ribonuclease MRP enzymes. *Cell* 62:407-409.
- Frank DN, Pace NR. 1998. Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu Rev Biochem* 67:153-180.
- Gardiner K, Pace NR. 1980. RNase P of *Bacillus subtilis* has a RNA component. *J Biol Chem* 255:7507-7509.
- Gardiner KJ, Marsh TL, Pace NR. 1985. Ion dependence of the *Bacillus subtilis* RNase P reaction. *J Biol Chem* 260:5415-5419.
- Gopalan V, Baxevas AD, Landsman D, 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.
- Groom KR, Heyman HC, Steffen MC, Hawkins L, Martin NC. 1998. *Kluyveromyces lactis* SEF1 and its *Saccharomyces cerevisiae* homologue bypass the unknown essential function, but not the mitochondrial RNase P function, of the *S. cerevisiae* RPM2 gene. *Yeast* 14:77-87.
- Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35:849-857.
- Guerrier-Takada C, Haydock K, Allen L, Altman S. 1986. Metal ion requirements and other aspects of the reaction catalyzed by M1 RNA, the RNA subunit of ribonuclease P from *Escherichia coli*. *Biochemistry* 25:1509-1515.

- Guerrier-Takada C, Li Y, Altman S. 1995. Artificial regulation of gene expression in *Escherichia coli* by RNase P. *Proc Natl Acad Sci U S A* 92:11115-11119.
- Guerrier-Takada C, McClain WH, 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.
- Gutell RR. 1996. Comparative sequence analysis and the structure of 16S and 23S rRNA. In: Zimmermann MA, Dahlberg AE, eds. *Ribosomal RNA: Structure, Evolution, Processing, and Function in Protein Biosynthesis*. Boca Raton, FL: CRC press.
- Haas ES, Brown JW. 1998. Evolutionary variation in bacterial RNase P RNAs. *Nucleic Acids Res* 26:4093-4099.
- Haas ES, Morse DP, Brown JW, Schmidt FJ, Pace NR. 1991. Long-range structure in ribonuclease P RNA. *Science* 254:853-856.
- Han SJ, Lee BJ, Kang HS. 1998. Purification and characterization of the nuclear ribonuclease P of *Aspergillus nidulans*. *Eur J Biochem* 251:244-251.
- Hartmann E, Hartmann RK. 2003. The enigma of ribonuclease P evolution. *Trends Genet* 19:561-569.
- Heus HA, Pardi A. 1991. Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* 253:191-194.
- Hollingsworth MJ, Martin NC. 1986. RNase P activity in the mitochondria of *Saccharomyces cerevisiae* depends on both mitochondrion and nucleus-encoded components. *Mol Cell Biol* 6:1058-1064.
- Jacob Y, Seif E, Paquet PO, Lang BF. 2004. Loss of the mRNA-like region in mitochondrial tmRNAs of jakobids. *RNA* 10:605-614.
- Jaeger L, Michel F, Westhof E. 1994. Involvement of a GNRA tetraloop in long-range RNA tertiary interactions. *J Mol Biol* 236:1271-1276.
- Jayanthi GP, Van Tuyle GC. 1992. Characterization of ribonuclease P isolated from rat liver cytosol. *Arch Biochem Biophys* 296:264-270.

- Kaye NM, Zahler NH, Christian EL, Harris ME. 2002. Conservation of helical structure contributes to functional metal ion interactions in the catalytic domain of ribonuclease P RNA. *J Mol Biol* 324:429-442.
- Kirsebom LA, Svard SG. 1994. Base pairing between Escherichia coli RNase P RNA and its substrate. *EMBO J* 13:4870-4876.
- Kline L, Nishikawa S, Soll D. 1981. Partial purification of RNase P from Schizosaccharomyces pombe. *J Biol Chem* 256:5058-5063.
- Krasilnikov AS, Xiao Y, Pan T, Mondragon A. 2004. Basis for structural diversity in homologous RNAs. *Science* 306:104-107.
- Krasilnikov AS, Yang X, Pan T, Mondragon A. 2003. Crystal structure of the specificity domain of ribonuclease P. *Nature* 421:760-764.
- Kufel J, Kirsebom LA. 1998. The P15-loop of Escherichia coli RNase P RNA is an autonomous divalent metal ion binding domain. *RNA* 4:777-788.
- LaGrandeur TE, Huttenhofer A, Noller HF, Pace NR. 1994. Phylogenetic comparative chemical footprint analysis of the interaction between ribonuclease P RNA and tRNA. *EMBO J* 13:3945-3952.
- Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387:493-497.
- Lang BF, Gray MW, Burger G. 1999. Mitochondrial genome evolution and the origin of eukaryotes. *Annu Rev Genet* 33:351-397.
- Lee YC, Lee BJ, Hwang DS, Kang HS. 1996a. Purification and characterization of mitochondrial ribonuclease P from Aspergillus nidulans. *Eur J Biochem* 235:289-296.
- Lee YC, Lee BJ, Kang HS. 1996b. The RNA component of mitochondrial ribonuclease P from Aspergillus nidulans. *Eur J Biochem* 235:297-303.

- Loria A, Niranjanakumari S, Fierke CA, Pan T. 1998. Recognition of a pre-tRNA substrate by the *Bacillus subtilis* RNase P holoenzyme. *Biochemistry* 37:15466-15473.
- Loria A, Pan T. 1996. Domain structure of the ribozyme from eubacterial ribonuclease P. *RNA* 2:551-563.
- Massire C, Jaeger L, Westhof E. 1997. Phylogenetic evidence for a new tertiary interaction in bacterial RNase P RNAs. *RNA* 3:553-556.
- Massire C, Jaeger L, Westhof E. 1998. Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. *J Mol Biol* 279:773-793.
- Michel F, Westhof E. 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol* 216:585-610.
- Murphy FL, Cech TR. 1994. GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J Mol Biol* 236:49-63.
- Odell L, Huang V, Jakacka M, Pan T. 1998. Interaction of structural modules in substrate binding by the ribozyme from *Bacillus subtilis* RNase P. *Nucleic Acids Res* 26:3717-3723.
- Pannucci JA, Haas ES, Hall TA, Harris JK, Brown JW. 1999. RNase P RNAs from some Archaea are catalytically active. *Proc Natl Acad Sci U S A* 96:7803-7808.
- Pascual A, Vioque A. 1999. Substrate binding and catalysis by ribonuclease P from cyanobacteria and *Escherichia coli* are affected differently by the 3' terminal CCA in tRNA precursors. *Proc Natl Acad Sci U S A* 96:6672-6677.
- Pomeranz Krummel DA, Altman S. 1999. Verification of phylogenetic predictions in vivo and the importance of the tetraloop motif in a catalytic RNA. *Proc Natl Acad Sci U S A* 96:11200-11205.

- Proctor DJ, Schaak JE, Bevilacqua JM, Falzone CJ, Bevilacqua PC. 2002. Isolation and characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions. *Biochemistry* 41:12062-12075.
- Puranam RS, 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.
- Rasmussen TA, Nolan JM. 2002. G350 of Escherichia coli RNase P RNA contributes to Mg<sup>2+</sup> binding near the active site of the enzyme. *Gene* 294:177-185.
- Rossmannith W, Karwan RM. 1998. Characterization of human mitochondrial RNase P: novel aspects in tRNA processing. *Biochem Biophys Res Commun* 247:234-241.
- Rossmannith W, Potuschak T. 2001. Difference between mitochondrial RNase P and nuclear RNase P. *Mol Cell Biol* 21:8236-8237.
- Rox C, Feltens R, Pfeiffer T, Hartmann RK. 2002. Potential contact sites between the protein and RNA subunit in the Bacillus subtilis RNase P holoenzyme. *J Mol Biol* 315:551-560.
- Seif E, Leigh J, Liu Y, Roewer I, Forget L, Lang BF. 2005. Comparative mitochondrial genomics in zygomycetes: bacteria-like RNase P RNAs, mobile elements and a close source of the group I intron invasion in angiosperms. *Nucleic Acids Res* 33:734-744. Print 2005.
- Seif ER, Forget L, Martin NC, Lang BF. 2003. Mitochondrial RNase P RNAs in ascomycete fungi: lineage-specific variations in RNA secondary structure. *RNA* 9:1073-1083.
- Shu HH, Martin NC. 1991. RNase P RNA in Candida glabrata mitochondria is transcribed with substrate tRNAs. *Nucleic Acids Res* 19:6221-6226.
- Siegel RW, Banta AB, Haas ES, Brown JW, Pace NR. 1996. Mycoplasma fermentans simplifies our view of the catalytic core of ribonuclease P RNA. *RNA* 2:452-462.
- Stark BC, Kole R, Bowman EJ, Altman S. 1978. Ribonuclease P: an enzyme with an essential RNA component. *Proc Natl Acad Sci U S A* 75:3717-3721.



- Stathopoulos C, Tekos A, Zarkadis IK, Drainas D. 2001. Extensive deproteinization of *Dictyostelium discoideum* RNase P reveals a new catalytic activity. *Eur J Biochem* 268:2134-2140.
- Stribinskis V, Gao GJ, Sulo P, Ellis SR, Martin NC. 2001. Rpm2p: separate domains promote tRNA and Rpm1r maturation in *Saccharomyces cerevisiae* mitochondria. *Nucleic Acids Res* 29:3631-3637.
- Talbot SJ, Altman S. 1994. Gel retardation analysis of the interaction between C5 protein and M1 RNA in the formation of the ribonuclease P holoenzyme from *Escherichia coli*. *Biochemistry* 33:1399-1405.
- Talla E, Anthouard V, Bouchier C, Frangeul L, Dujon B. 2005. The complete mitochondrial genome of the yeast *Kluyveromyces thermotolerans*. *FEBS Lett* 579:30-40.
- Tanaka T, Ando T, Haga S, Kikuchi Y. 2004a. Examining the bases of the J3/4 domain of *Escherichia coli* ribonuclease P. *Biosci Biotechnol Biochem* 68:1388-1392.
- Tanaka T, Kanda N, Kikuchi Y. 2004b. The P3 domain of *E. coli* ribonuclease P RNA can be truncated and replaced. *FEBS Lett* 577:101-104.
- Thomas BC, Chamberlain J, Engelke DR, Gegenheimer P. 2000a. Evidence for an RNA-based catalytic mechanism in eukaryotic nuclear ribonuclease P. *RNA* 6:554-562.
- Thomas BC, Gao L, Stomp D, Li X, Gegenheimer PA. 1995. Spinach chloroplast RNase P: a putative protein enzyme. *Nucleic Acids Symp Ser*:95-98.
- Thomas BC, Li X, Gegenheimer P. 2000b. Chloroplast ribonuclease P does not utilize the ribozyme-type pre-tRNA cleavage mechanism. *RNA* 6:545-553.
- Turmel M, Lemieux C, Burger G, Lang BF, Otis C, Plante I, Gray MW. 1999. The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*. Two radically different evolutionary patterns within green algae. *Plant Cell* 11:1717-1730.
- Vainauskas S, Stribinskis V, Padegimas L, Juodka B. 1998. Partial characterization of the ribonuclease P from *Tetrahymena pyriformis*. *Biochimie* 80:595-604.

- Vioque A, Arnez J, Altman S. 1988. Protein-RNA interactions in the RNase P holoenzyme from *Escherichia coli*. *J Mol Biol* 202:835-848.
- Wagner M, Fingerhut C, Gross HJ, Schon A. 2001. The first phytoplasma RNase P RNA provides new insights into the sequence requirements of this ribozyme. *Nucleic Acids Res* 29:2661-2665.
- Westhof E, Wesolowski D, Altman S. 1996. Mapping in three dimensions of regions in a catalytic RNA protected from attack by an Fe(II)-EDTA reagent. *J Mol Biol* 258:600-613.
- Wise CA, Martin NC. 1991. Dramatic size variation of yeast mitochondrial RNAs suggests that RNase P RNAs can be quite small. *J Biol Chem* 266:19154-19157.

### Legends to figures

**Figure 1:** Secondary structure models of bacteria-like mtP-RNAs.

**A.** Minimum bacterial consensus (Brown et al., 1994) and mtP-RNA secondary structure models for *N. olivacea* (Turmel et al., 1999), and *Rhizopus oryzae* (Seif et al., 2005). The S- and C- domains are indicated in the minimum bacterial consensus, by a broken red line.

**B.** MtP-RNA secondary structure models for the jakobids *Reclinomonas*

*americana* #50294 (#50394), *R. americana* #50283, *R. americana* #50284, *S. ecuadoriensis*, *J. libera*, and *J. bahamiensis*. In the model of *R. americana* #50284, nucleotides in green denote the few sequence differences to *Reclinomonas* sp. #50633. Positions in red are invariant in the minimum bacterial consensus; uppercase letters indicate 100%, lowercase >80% conservation of the minimum bacterial consensus sequence. The arrows pinpoint experimentally determined RNA termini; the arrow length is proportional to the percentage of molecules ending at a given position. Nucleotides boxed in black correspond to the mutations that were introduced to increase the GC content of P4 (see also Table 1). A(292) that has been mutated to correct the deviation from CRV sequence (Table I), is circled in black.

**Figure 2:** Catalytic activity assays of type I chimeras.

**A.** Secondary structure model for type I chimeras. The *E. coli* domain is shown in red, and the mitochondrial domain in black.

**B.** Catalytic assays of type I chimeras with pre-tRNA<sup>Pro</sup>, under conditions optimal for archaeobacterial RNase P. 1,000 cpm of substrate was incubated at 37°C for 18 hours in the presence of 100 nM of ribozyme, in 300 mM MgCl<sub>2</sub>, 2.5 M ammonium acetate, 50 mM Tris pH 8, 0.1% SDS, and 0.05% Nonidet P-40. All tested RNAs show activity except the *Rhizopus oryzae* chimera. 10 nM of *E. coli* P-RNA was used as the positive control, and substrate only as negative control.

**C.** Catalytic assays of type I chimeras with pre-tRNA<sup>Pro</sup>, at high salt conditions. Assays were performed in 50 mM Tris, pH 7.5, 100mM NH<sub>4</sub>Cl, 110 mM MgCl<sub>2</sub>, with ribozyme concentrations as in Fig. 2B. Under these conditions, only the *J. libera* and *R. americana* #50394 chimeras are catalytically active.

**Figure 3:** Catalytic activity assays of type II chimeras.

**A.** Secondary structure model for type II chimeras. The *E. coli* domain is shown in red, and the mitochondrial domain in black.

**B.** Catalytic assays of type II chimeras with pre-tRNA<sup>Pro</sup> and pre-tRNA<sup>Lys</sup>, at salt conditions optimal for archaeobacterial RNase P. 1,000 cpm of substrate were incubated at 37°C for 18 hours in the presence of 1 μM of ribozyme, in 300 mM MgCl<sub>2</sub>, 2.5 M ammonium acetate, 50 mM Tris pH 8, 0.1% SDS, and 0.05% Nonidet P-40. *J. libera* and *R. americana* #50394 have the highest catalytic activity followed by *J. libera* and *S. ecuadoriensis*, where as *N. olivacea* is inactive.

**C.** Catalytic assays of type II chimeras in the presence of *E. coli* P-protein.

The assays were performed as described in material and methods except that the concentrations of P-RNAs and P-protein was five fold higher.

**Figure 4:** Catalytic assays of GNRA tetra-loop mutants in chimeras of *J. libera* and *R. americana* #50394.

**A.** Catalytic assays for mutants in L18 of *J. libera* and *R. americana* #50394 type I chimeras. Experiments with the respective wild type (GUAA) and mutant sequences are labelled in the figure. In the negative control, no ribozyme was included. The assays were performed with 100nM of the corresponding P-RNA, for 18 hours at 37° C, with 1,000 cpm of pre-tRNA<sup>Lys</sup> in a buffer containing 30 mM MgCl<sub>2</sub>, 100 mM NH<sub>4</sub>Cl, 50 mM Tris pH 7.5.

**B.** Catalytic assays of a mutant of L14 in the type II chimera of *J. libera*.

Experiments with the wild type (GAGA) and mutant sequence (GAGG) are labelled in the figure. Substrate, no ribozyme was included. The assays conditions were" 1,000 nM of ribozyme for 18 hours at 37° C, with the conditions optimal for archaeobacterial P-RNAs (see material and methods).

**Figure 5:** P3 mutagenesis in *S. ecuadoriensis*.

**A.** Catalytic assays for mutants at the base of P3.

Position 32 is marked by an asterisk (\*) in both A and B. 100 nM of the corresponding RNAs were incubated with 1,000 cpm of pre-tRNA<sup>Pro</sup> in a buffer containing 100 mM MgCl<sub>2</sub>, 2.5 M ammonium acetate, 50 mM Tris pH 8, 0.1% SDS, and 0.05% Nonidet P-

40. The \*AxA mutant has a twofold decrease in catalytic activity compared to the \*G-C mutant and the \*U-A wild type sequence. "Sub." denotes substrate only (negative control).

**B.** Secondary structure model of P3 and surrounding nucleotides.

**Table I:** Effect of specific mutations in *R. americana* type I and type II chimeras.

Mutations	Relative catalytic activity		
	300 mM MgCl <sub>2</sub> 2.5 M CH <sub>3</sub> COONH <sub>4</sub> (50 mM Tris pH 8, 0.1% SDS, and 0.05% NP-40)	100 mM MgCl <sub>2</sub> 100 mM NH <sub>4</sub> Cl (50 mM Tris pH 7.5)	20 mM MgCl <sub>2</sub> 100 mM NH <sub>4</sub> Cl (50 mM Tris pH 7.5)
ΔP6	30%	-	-
L15/16-CCA	50%	-	-
ΔP19	80%	-	-
P4 G+C rich	100%	320%	-
A-G 292	110%	120%	370%
L9 GAGA	150%	-	-

Key to mutant code:

ΔP6; mispairs in P6. L15/16-CCA: mutation of GGU motif to CCA. ΔP19: deletion of P19. P4 G+C rich: exchange of A-U base pairs by G-C pairs in P4 (black boxed in Fig. 1). A-G 292: mutation of A292 to G in CRV (black encircled in Fig. 1). L9 GAGA: GAGA introduced into L9.

Catalytic efficiency was defined as the percentage of substrate cleaved by the mutant, compared to the original chimera. For all tests, 100 nM of type I chimera mutant P-RNA (or 1,000 nM of type II chimera mutant P-RNA) were incubated for 18 hours at 37° C, in presence of 1,000 cpm pre-tRNA<sup>Pro</sup>. The reaction buffers are described in the Table.





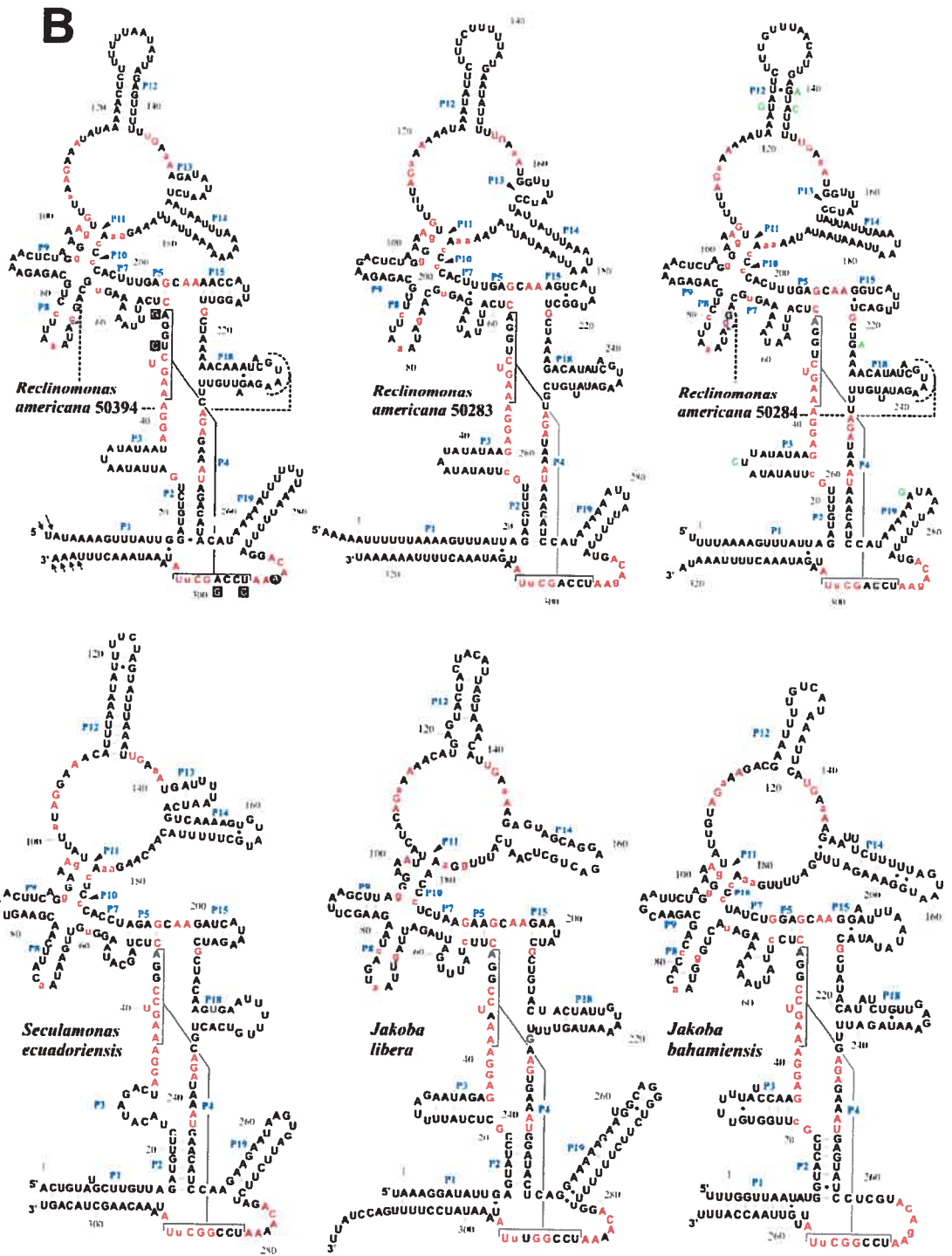
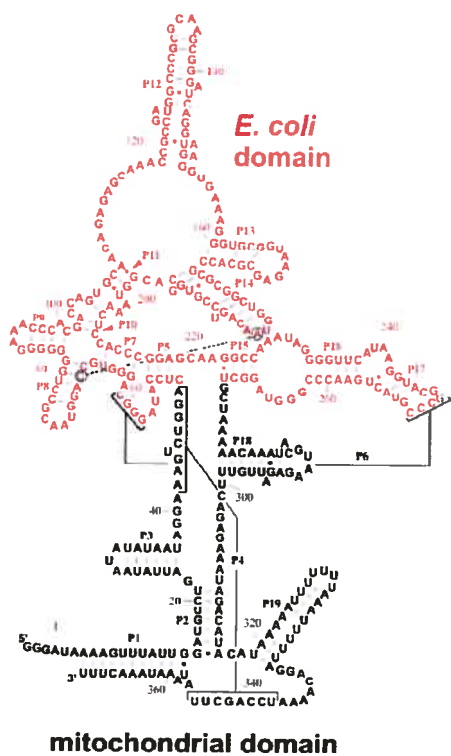
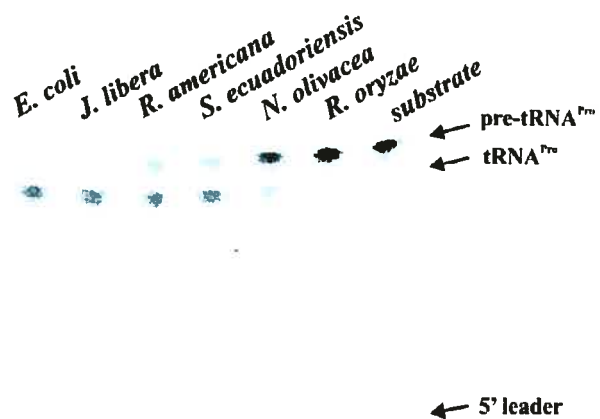


Figure 1B

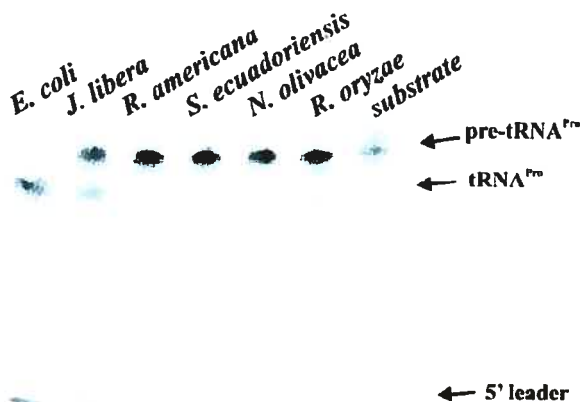
**A type I chimera**



**B**

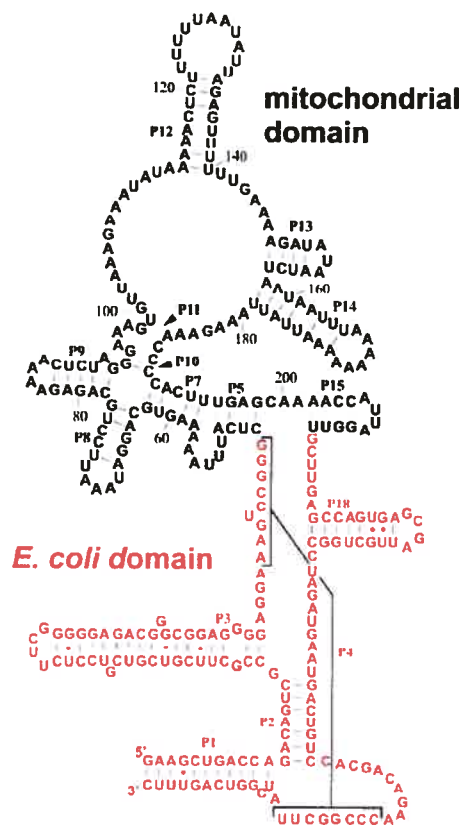


**C**

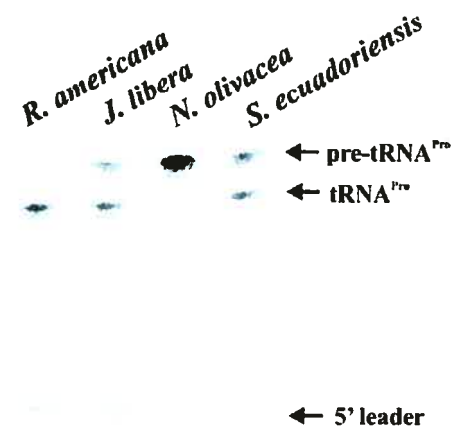


**Figure 2**

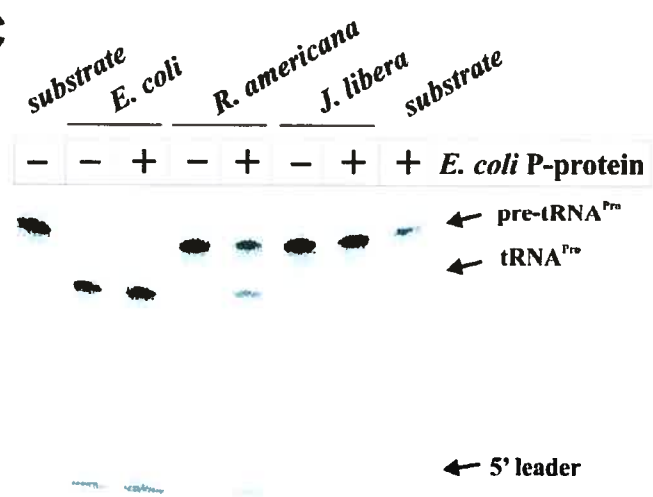
**A type II chimera**



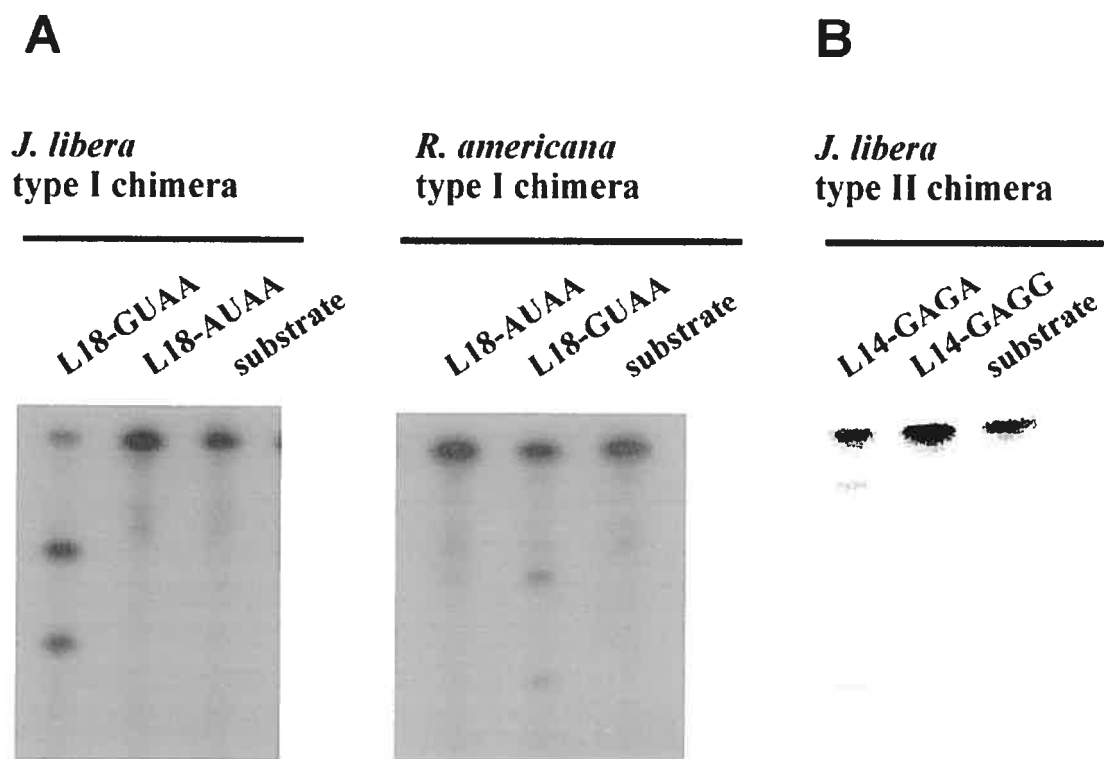
**B**



**C**



**Figure 3**

**Figure 4**

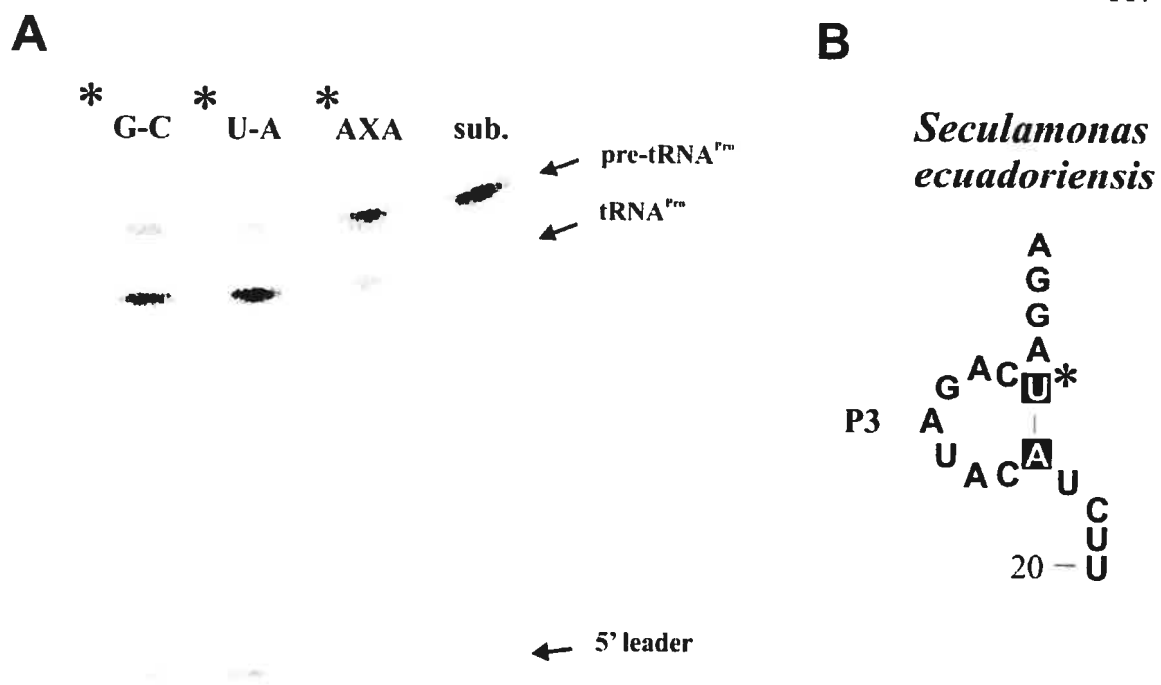


Figure 5

**Supplementary information Table S1:** Oligonucleotides used for *in vitro* mutagenesis.

Species	Type	5' end	3' end
<i>Reclinomonas americana</i>	mtP-RNA	5'GAAATTAATACGACTCACTATAGGGA TAAAAGTTTATTGGATGTCTGA3'	5'TAAAGTTTATTTATAAGCTGGATT3'
	Type I	5'TCCAGACTTTCCTATTATATA3'	5'GCTAAAAACAAATACGTAAGAGA3'
	Type II	5'CTCATTTAAAAAGTGCAGGAT3'	5'ACCTAATGGTTTTGCTCAAAG3'
	L9GNA	5'AGTTTCCTCTGCAGGAATTTATCC3'	5'CTAGGGAAAGTGTAAAGAA3'
	L18ANAA	5'GCTAAAAACAAATACATAAGAG 3'	5'AGCCTACCCGGGTTTCAGTACGG3'
	P19deletion	5'ATACAAAATCCAGCTTATAAATAAAC3'	5'TATATGTCTATTTCTCTGAAACAA3'
	P6 deletion	5'CTCCATAACACAGGGTGCCAGGTAAC3'	5'AGCCTACCCGGGTTTCAGTACGG3'
	L16 CCA	5'GCTAAAAACAAATACGTAAGAGA3'	5'AGCCTTGGCGGGTTTCAGTACGGGC3'
	P4 G+C rich	5'AAACCCGGCTTATAAATAAACTTT3'	5'ATATAATAGGAAAGTCCGGGCTC3'
		5'ATTATAATCAGACATCCAATAAAC3'	5'TGTCCTAAAAATTTAAAAAAATTT3'
A to G292	5'GAATCCAGCTTATAAATAAACTTT3'	5'TGTCCTAAAAATTTAAAAAAATTT3'	
<i>Jakoba libera</i>	mtP-RNA	5'GAAATTAATACGACTCACTATAGGGT AAAGGATATTGAGTATCCGCTC3'	5'TAATAGGTCAAAGGATATTTATA A3'
	Type I	5'TCCGGATTTTCCTCTCTATTC3'	5'GCTGTACTACTATTGTA AAAA3'
	Type II	5'CTTCTATTTGATTAGATTAGTTAATG3'	5'GATATTCTTGCTTCTTAGAGGG3'
	L18ANAA	5'GCTGTACTACTATTATAAAA3'	5'GATATTCTTGCTTCTTAGAGGG3'
	L14GNRG	5'GGCTGCTCAATCATTGGAAA3'	5'TCCTGCTACTCTTTTCAATGTT3'
<i>Jakoba bahamiensis</i>	mtP-RNA	5'GAAATTAATACGACTCACTATAGGGT TGGTTAATATGGTACTCGCT3'	5'ATGGTTAACAATAAGCCGGATTC3'
	Type II	5'CTCCTTATTA AAAAAGATAC3'	5'GTATATATTA AATTCCTTGCTCC3'
<i>Seculamonas ecuadoriensis</i>	mtP-RNA	5'GAAATTAATACGACTCACTATAGGGT AACTGTATGCTTGTTAGATGTTTC3'	5'ACTGTAGCTTGTTTATAAGCCGG 3'
	Type I	5'TCCGGACTTTCCTAGTCTATG3'	5'GCTACAAGTGAATTTTGTCA3'
	Type II	5'CTCTAGCATAAGGTGTTGAATAAC3'	5'GATCTTATGATCTTGCTCTAGGTG3'
	P3 A-A	5'ACATAGACAAGGAAAGTCCGGA 3'	5'AGAACATCTAACAAGCATAACAG3'
	P3 G-C	5'CCATAGACGAGGAAAGTCCGGA 3'	5'AGAACATCTAACAAGCATAACAG3'
<i>Nephroselmis olivacea</i>	mtP-RNA	5'GAAATTAATACGACTCACTATAGGGA AAAAGTTTTCAGAGAATTTGATGC3'	5'ATTCAGAGAATTTCTAAGCTGGG3'
	Type I	5'CCCAGACTTTCCTTTTACATTATTC3'	5'GCTAAAAATCTGAAGCTCCTTGAT3'
	Type II	5'CTCCAATAGAATAAAAATATAC3'	5'GAATTA AATCTTGACCAAGT3'
<i>Rhizopus oryzae</i>	mtP-RNA	5'GAAATTAATACGACTCACTATAGGGT CTTAGGAGTTATAATAGTTA3'	5'TTCTTAGAGTTAAATAAGCCGA3'
	Type I	5'CCCGGACTTTCCTCAA3'	5'GCTTTTTATTCAATAAATAAATT3'
<i>Rhizopus oligosporus</i>	mtP-RNA	5'GAAATTAATACGACTCACTATAGGGT TCTTAGGAATTAATAGTTATTA3'	5'TTCTTAGAATTA AATAAGCCGGT 3'
<i>E. coli</i>	mtP-RNA	5'GAAATTAATACGACTCACTATAGGGA AGCTGACCAGACAGTCGC3'	5'AGGTGAAACTGACCGATAAGCC3'
	Type I	5'CTCCATAGGGCAGGGTGCCAGGTA3'	5'AGCCTACCCGGGTTTCAGTACGG3'

### **Chapitre III : Discussion**

### D.1 Une nouvelle approche pour identifier les gènes *rnpB*

Les outils informatiques utilisés pour l'identification des gènes protéiques, comme « Blast » et « Fasta », se basent essentiellement sur la séquence primaire de ces gènes comme critère de recherche, et ils ne sont pas efficaces pour l'identification des gènes *rnpB* pour deux raisons. Premièrement, la conservation chez les gènes codant des ARN structuraux est essentiellement au niveau de la structure secondaire. Deuxièmement, les gènes codant la composante ARN de la RNase P mitochondriale sont riches en A et T, particulièrement chez les ascomycètes, ce qui cause des artefacts majeurs avec les approches informatiques traditionnelles.

Pour remédier à ces problèmes, la première étape du projet était d'établir une nouvelle approche pour identifier les gènes *rnpB* dans les génomes mitochondriaux complètement séquencés d'une façon systématique, afin d'avoir le plus grand échantillonnage taxonomique possible. Nous avons adopté une approche qui est basée sur l'identification de motifs de séquences universellement conservés et sur l'identification des hélices P1 et P4 simultanément. Cette méthode a été décrite dans l'article 1. Elle nous a permis d'identifier de nouveaux gènes *rnpB* chez les ascomycètes, les zygomycètes, et les jakobides (voir les articles 1, 2, 3). Une trouvaille inattendue était celle du gène *rnpB* dans le génome mitochondrial de *S. pombe*, dont la séquence était disponible depuis le début des années 1980, mais personne n'avait réussi à identifier le gène *rnpB*. Les candidats identifiés par cette méthode chez les champignons, *Schizosaccharomyces pombe*, *Schizosaccharomyces octosporus*, *Mortierella verticillata*, et *Rhizopus oryzae* ont été cartographiés, et nous avons montré que ces candidats sont transcrits et maturés *in vivo* comme prédit par les modèles de structures secondaires que nous avons établis (voir articles 1 et 2). Notre approche a fait ses preuves une fois de plus, lorsqu'elle a été utilisée par le groupe de chercheurs qui a séquencé et analysé le génome mitochondrial de l'ascomycète *Kluyveromyces thermotolerans* (Talla et al., 2005).



## **D.2 Evolution de la structure secondaire des ARN-Pmt**

### **D.2.1 Modélisation par comparaison phylogénétique**

Suite à l'identification de nouveaux gènes *rnpB*, nous avons établi des modèles de structures secondaires pour les ARN-Pmt correspondants, par la méthode de comparaison phylogénétique. Cette approche a déjà fait ses preuves dans le passé (Pace et al., 1989), et reste la méthode la plus efficace pour établir des modèles de base. Par la suite, ces modèles peuvent être validés et raffinés par les méthodes en laboratoire, comme la cristallisation suivie par une étude aux rayons X (Ferre-D'Amare et al., 1998), la résonance magnétique nucléaire (Shen et al., 1995), le pontage (crosslinking)(Chen et al., 1998), la mutagenèse, ainsi que les méthodes de protection enzymatique et chimique (Brenowitz et al., 2002; Brunel et Romby, 2000; Ehresmann et al., 1987; Knapp, 1989). Pour établir nos modèles, nous nous sommes basés sur les structures secondaires déjà publiées chez les eubactéries et chez la mitochondrie de *Reclinomonas americana* NZ (Brown, 1999; Lang et al., 1997). Dans certains cas, nous avons utilisé le logiciel «Mfold» pour générer les hélices possibles dans des petites régions qui varient entre 20 et 50 nucléotides (Zuker, 1989). L'utilité de ce logiciel était limitée car ce dernier ne peut détecter les pseudo-nœuds, et donc ne pouvait recréer la région entourant le P4, la partie la plus conservée chez les ARN-Pmt.

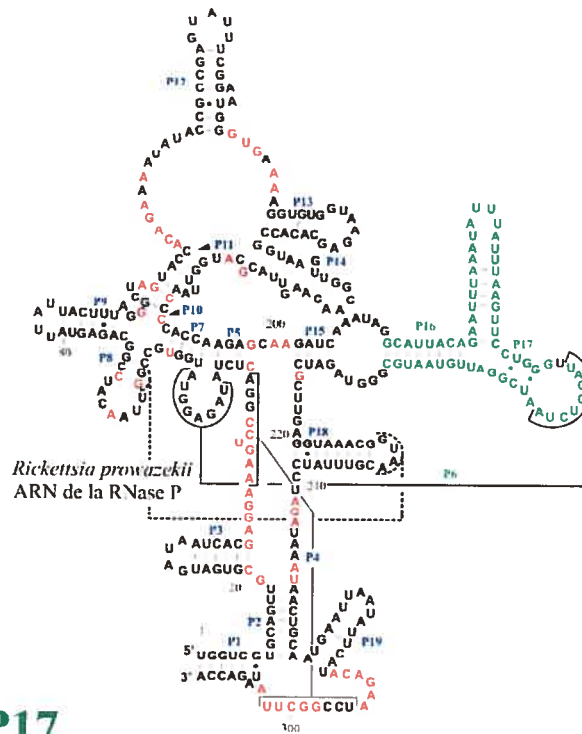
### **D.2.2 Les structures secondaires des ARN-Pmt des isolats des jakobides sont les mieux conservées**

Grâce aux nouveaux modèles établis chez les jakobides et suite à l'identification de cinq nouveaux gènes *rnpB* chez *Seculamonas ecuadoriensis*, *Jakoba libera*, *Jakoba bahamiensis*, *Reclinomonas americana* 50283 et *R. americana* 50284, nous avons mis à

jour le modèle de la structure secondaire de l'ARN-Pmt de *R. americana* NZ en allongeant le P18 (voir article 3). La structure secondaire des isolats de *R. americana* est la plus similaire à celle des eubactéries parmi les ARN-Pmt connus. Si on compare cette structure à celles des ARN-P des  $\alpha$ -protéobactéries, et plus précisément à celle de *Rickettsia prowazekii*, nous remarquons l'absence des hélices P6, P16, et P17 (Figure D.1). Parmi les jakobides, nous notons aussi l'absence du P3 chez *S. ecuadoriensis*, du P13 chez *J. libera* et *J. bahamiensis*, et du P19 chez *J. bahamiensis* seulement. Le P14 est absent chez le prasinophyte (algue verte) *Nephroselmis olivacea* (Turmel et al., 1999) (voir Figure 1 de l'article 3).

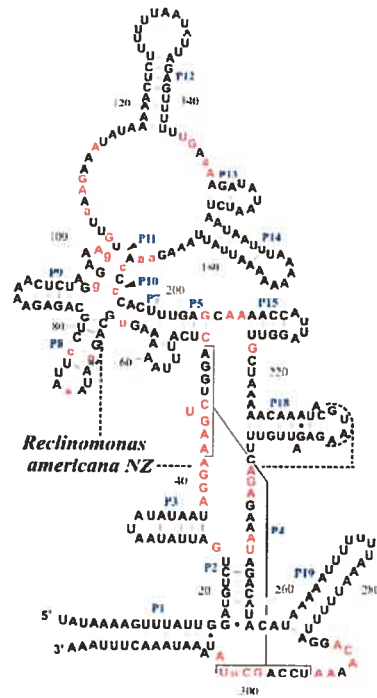
Figure D.1 : Modèle de structure secondaire de l'ARN-P de *Rickettsia prowazekii* et de l'ARN-Pmt de *R. americana* NZ.

Les nucléotides en rouge sont universellement conservés (Chen et Pace, 1997), alors que les nucléotides en vert dans le modèle de *R. prowazekii* montrent les éléments structuraux absents dans le modèle de l'ARN-Pmt de *R. americana* NZ.



**Perte des  
hélices**

**P6, P16, P17.**



### D.2.3 Evolution convergente vers une structure minimale chez les champignons

Chez les champignons, les ARN-Pmt du genre *Rhizopus* partagent le plus grand nombre d'hélices avec le modèle de *R. americana* NZ (voir article 2). L'ARN-Pmt chez *Rhizopus oryzae* n'a perdu que trois hélices, les P13, P14, et P19. Les données présentement disponibles montrent que ces hélices sont absentes chez tous les champignons et laissent croire qu'elles ont été perdues chez leur ancêtre commun. Nous retrouvons une grande variation dans les modèles de structure secondaire de l'ARN-Pmt chez les champignons, mais nous avons noté que deux hélices se trouvent chez tous les ARN-Pmt, le P1, le P4 et la région qui les entourent et qui inclut les régions conservées CRI et CRV. Nous avons même retrouvé des ARN-Pmt qui se limitent à ces deux hélices chez les levures, et chez le zygomyète *Smittium culisetae*. Est-ce un hasard que dans deux lignées évolutives différentes, celles des zygomyètes et des ascomycètes, nous retrouvons deux ARN-Pmt qui n'ont retenu que ces deux hélices? Cette observation laisse croire que le P1, le P4 et les régions en simple brin avoisinantes, qui forment le site catalytique chez les ARN-P eubactériens, ont un rôle important et incontournable dans la fonction des ARN-Pmt.

L'étude structurale et comparative que nous avons menée nous apprend que les ARN-Pmt ont perdu plusieurs des éléments structuraux présents chez les ARN-P des  $\alpha$ -protéobactéries (Figure D.1), et que cette perte est plus accentuée chez les champignons que chez les protistes (articles 1, 2 et 3). Est-ce que cette perte est nécessairement accompagnée par une réduction de la taille de ces ARN? Et est-ce qu'elle suit un ordre défini?

#### D.2.4 Les ARN-Pmt cibles d'invasion des éléments à double hélices «DHE »

Pour répondre à la première question, nous avons comparé la taille des ARN-Pmt, ou du gène *rnpB*, d'onze espèces, et le degré de réduction structurale de ces ARN (Table D.1). Nous avons défini la réduction structurale chez un ARN-Pmt, comme étant l'absence d'hélices qui ont déjà existé dans un modèle ancestral. Pour des fins pratiques, nous avons considéré l'ARN-Pmt de *R. americana* NZ comme étant le modèle ancestral, puisque ce dernier contient le plus grand nombre d'hélices et donc ressemble le plus à l'ancêtre commun des ARN-Pmt. Nous avons remarqué que les ARN-Pmt de *Mortierella verticillata*, *Rhizopus oryzae*, *Taphrina deformans*, *S. cerevisiae*, et *N. olivacea* ont une taille plus grande que celle de l'ARN-Pmt de *R. americana* NZ (voir Table D.1 et les articles 2 et 3), alors que leur structure secondaire est moins conservée que chez ce dernier. Néanmoins, les ARN-Pmt qui montrent les modèles structuraux les plus réduits, ceux de *S. fibuligera* et *S. culisetae*, ont les plus petites tailles, 140 et 188 nucléotides respectivement. Ces résultats montrent que la taille d'un ARN-Pmt ne reflète pas nécessairement le degré de conservation de sa structure secondaire, même si dans certains cas nous trouvons une corrélation. En fait, la taille des ARN-Pmt ne diminue pas nécessairement suite à la perte des éléments structuraux présents dans les formes plus ancestrales, pour la simple raison que d'autres éléments structuraux apparaissent chez ces ARN.

Certains éléments structuraux sont apparus suite à l'invasion du gène *rnpB* par des éléments mobiles comme c'est le cas chez les ARN-Pmt de plusieurs zygomycètes (article 2). Dans ce cas, nous avons observé l'apparition de grands inserts, dont les tailles varient entre 57 et 430 nucléotides, dans le P12 et dans les jonctions entre le P5 et le P18 ou P15, selon l'espèce. La séquence de ces régions peut être repliée en éléments doubles hélices (Double Helical Elements)(Paquin et al., 2000). Ces éléments sont des éléments

mobiles riches en GC, formés de deux hélices séparées souvent par un nucléotide. Ils envahissent les régions intergéniques des génomes mitochondriaux ainsi que les ARN ribosomiques. Notre étude a montré que ces éléments ne sont pas éliminés dans l'ARN-Pmt mature de *R. oryzae*, ce qui suggère qu'ils sont insérés dans des régions non-critique pour la fonction des ARN-Pmt (article 2).

Nom de l'espèce	Groupe phylogénétique	hélice(s) absente(s)	Taille de l'ARN-Pmt
<i>M. verticillata</i>	Zygomycète	P2, 3, 13-15, 19	982
<i>R. oryzae</i>	Zygomycète	P13, 14, 19	829
<i>S. cerevisiae</i>	Ascomycète	P2, 3, 5, 7-15, 19	423
<i>N. olivacea</i>	Prasinophyte	P13	359
<i>T. deformans</i>	Ascomycète	P3, 11, 13-15, 19	<i>319</i>
<i>J. libera</i>	Jakobide	P13	<i>307</i>
<i>S. ecuadoriensis</i>	Jakobide	P3	<i>306</i>
<i>S. pombe</i>	Ascomycète	P5, 7-14, 19	273
<i>A. nidulans</i>	Ascomycète	P3, 11, 13-15, 18, 19	232
<i>S. culisetae</i>	Zygomycète	P2, 3, 5, 7-15, 18, 19	<i>188</i>
<i>S. fibuligera</i>	Ascomycète	P2, 3, 5, 7-15, 18, 19	140

**Table D.1:** Comparaison entre les tailles des ARN-Pmt et la réduction de leur structure secondaire.

L'échantillon des espèces choisies représente la diversité des ARN-Pmt, et contient les espèces qui ont la plus grande variation par rapport à la taille de leur ARN-Pmt. Les hélices énumérées sont celles présentes chez *R. americana* NZ, mais absentes chez les espèces correspondantes. La taille de l'ARN-Pmt est celle déterminée expérimentalement lorsque cette information est disponible. Les tailles indiquées en italique ont été calculées à partir de la taille du gène *rnpB*.



Un autre élément apparu chez les ARN-Pmt des zygomycètes est l'hélice située entre le P5 et le P7. Elle varie de taille entre 20 et 100 nucléotides. Une telle hélice existe dans les modèles de structure secondaire des ARN-P eubactériens du type B similaire à *Bacillus subtilis* (Haas et Brown, 1998), où elle est identifiée comme P5.1. Comme pour les ARN-Pmt, les ARN-P eubactériens de type B sont caractérisés par leur séquence riche en A et U. Est-ce que l'apparition de cette hélice à la même position que chez les eubactéries implique qu'elle est importante pour la stabilisation de la structure des ARN-Pmt riches en A et U? Où est-ce simplement une coïncidence? Vu que la fonction exacte de cette hélice est inconnue pour l'instant, nous ne favorisons aucune hypothèse.

#### **D.2.5 La perte des éléments structuraux ne suit pas un ordre défini**

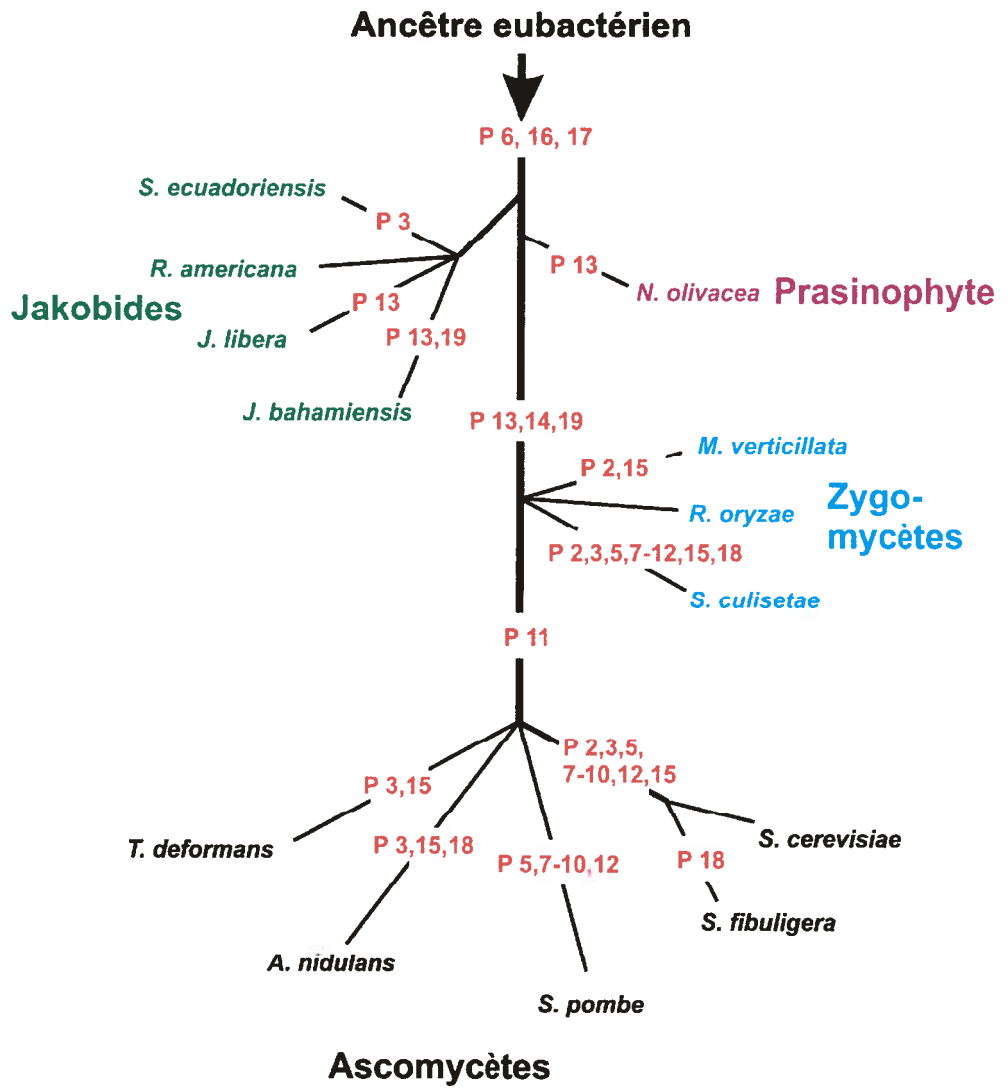
Pour déterminer si la perte des éléments structuraux chez les ARN-Pmt suit un ordre défini, nous avons établi un scénario de la disparition de ces éléments, en se basant sur les connaissances disponibles concernant la phylogénie des organismes où ces ARN ont été identifiés, et le modèle de structure secondaire de l'ARN-Pmt chez *R. americana* NZ (Figure D.2).

Chez les jakobides, les premières hélices perdues sont soit le P3 ou le P13. Chez *N. olivacea*, la seule hélice perdue est le P13. Quant aux champignons, nous remarquons qu'il y a une multitude de pertes, qui ne semblent pas suivre un ordre défini. Par contre, nous avons noté que les P13, 14 et 19 sont absents dans tous les modèles (Figure D.2), et que les hélices P5,7-10,12, qui forment en grande partie le domaine « S » chez les eubactéries sont perdues en un seul bloc. Nous avons remarqué que la présence de certaines hélices est plus fréquente que d'autres; alors nous avons calculé la fréquence de présence de chacune des hélices. Pour le faire, nous avons choisi un échantillon de modèle de structure secondaire d'ARN-Pmt d'une manière à représenter toutes les

variations observées, mais en éliminant les modèles provenant d'espèces proches, afin de ne pas surreprésenter les groupes phylogénétiques où les séquences sont plus disponibles, comme les levures (Table D.2). Mis à part le P1 et le P4 présents chez tous les espèces, cette étude montre que le P18 est le plus fréquent à 75%, suivi du P2 et du domaine P5, 7-10,12 à 67%. D'autre part, le P13 est le moins fréquent à 17% suivi par le P14 et le P19 à 33%. Ces données suggèrent que chez les ARN-Pmt, le domaine « C » (P1, 4 100%; P18 75%; P2 67%; P15 50%; P3 42%; P19 33%) est plus conservé que le domaine « S » (P7-10 67%; P11 50%; 14 33%; P13 17%).

Figure D.2 : Scénario hypothétique décrivant la perte des éléments structuraux chez les ARN-Pmt.

L'emplacement des espèces sur cette figure est basé sur les phylogénies décrites dans l'article 2 et (Gray et al., 1998). Les hélices perdues sont indiquées en rouge.



Identification de l'hélice	Nombre de présence d'une hélice / nombre d'espèces	Fréquence en pourcentage
P1,4	12/12	100%
P18	9/12	75%
P2	8/12	67%
P5,7,8,9,10,12	8/12	67%
P15	6/12	50%
P11	6/12	50%
P3	5/12	42%
P19	4/12	33%
P14	4/12	33%
P13	2/12	17%

**Table D.2** : Fréquences de présence des hélices d'origine eubactérienne chez les ARN-Pmt.

L'échantillon des espèces qui ont été considérées pour cette étude inclut *R. americana* NZ, *J. libera*, *S. ecuadoriensis*, *N. olivacea*, *R. oryzae*, *M. verticillata*, *S. culisetae*, *T. deformans*, *A. nidulans*, *S. pombe*, *S. fibuligera*, et *S. cerevisiae*.

### D.2.6 Des motifs GNRA et YNMG dans les tétra-boucles du P3, P9, P12 et P18

L'étude structurale que nous avons menée jusqu'à maintenant s'est limitée au niveau de la présence ou l'absence des hélices héritées de l'ancêtre eubactérien, mais d'autres informations peuvent en être retirées. Par exemple, nous avons remarqué la présence des motifs GNRA (N=G, A, T, C; R=G, A) dans (i) le P9 de *R. oryzae*, *Mucor mucedo*, *Rhizopus stolonifer* 194667, et *Rhizopus oligosporus*, (ii) le P18 de tous les jakobides à part *S. ecuadoriensis*, et (iii) le P12 de *J. libera* et *J. bahamiensis*. Comme mentionné dans l'article 3, les tétra-boucles de la famille GNRA peuvent stabiliser les hélices dans lesquelles elles sont présentes, interagir avec des régions éloignées de l'ARN ou avec une protéine (Belanger et al., 2004; Correll et al., 1998; Legault et al., 1998). Chez *E. coli*, le motif GNRA du L18 interagit avec le P8, celui du L9 avec le P1, et celui du L13 avec le P8 (Brown et al., 1996; Harris et Pace, 1995a; Massire et al., 1998). La mutation de ces motifs diminue l'activité catalytique de la RNase P (Pomeranz Krummel et Altman, 1999). Nous avons remarqué aussi la présence du motif UUAG dans la tétra-boucle du P3 chez *J. libera* (article 3). Ce motif appartient à la famille des tétra-boucle YNMG (Y=U, C; M= A, C)(Proctor et al., 2002). Cette famille de tétra-boucles est particulièrement stable, et peut servir de site d'attache pour les protéines, ou interagir former des interactions tertiaires à l'intérieur de l'ARN (Ghisolfi-Nieto et al., 1996; Proctor et al., 2002). Même s'il est présent chez l'ARN-P d'*E. coli*, ce motif n'a jamais fait l'objet d'une étude par mutagenèse afin d'évaluer son rôle ou son importance pour l'activité catalytique. Il est connu que le P3 contribue à l'attache de l'ARN-P d'*E. coli* à la protéine-P (Biswas et al., 2000), ce qui soulève la possibilité que ce motif y soit impliqué aussi.

Tous les motifs décrits plus haut sont présents chez l'ARN-P d'*E. coli*. Chez les  $\alpha$ -protéobactéries le motif GNRA est présent dans le L9 des ARN-P de *Caulobacter crescentus*, *Agrobacterium tumefaciens*, *Rhodobacter capsulatus*, *Rhodospirillum rubrum* et chez onze des quinze isolats de *Wolbachia* (Andersson et al., 1998; Brown et

al., 1991; Brown et al., 1996; Harris et al., 2001) (<http://www.mbio.ncsu.edu/RNaseP/home.html>). Le motif GNRA est omniprésent dans le L14 et L18, de toutes les  $\alpha$ -protéobactéries étudiées; alors que le motif YNMG du P3 y est absent. La présence des motifs GNRA chez les ARN-Pmt des jakobides et chez les  $\alpha$ -protéobactéries suggère que ces motifs étaient présents chez l'ancêtre commun de ces deux groupes. En ce qui concerne le motif de la famille YNMG du P3 chez *J. libera*, nos données ne nous permettent pas de supporter l'hypothèse que ce motif provient de l'ancêtre eubactérien, même si cette possibilité demeure une hypothèse pertinente.

En résumé, notre étude structurale comparative nous mène à conclure que les ARN-Pmt ont perdu plusieurs des éléments structuraux présents chez leur ancêtre eubactérien, et que cette réduction structurale a créé des modèles de structures secondaires très hétérogènes, spécialement chez les mitochondries des champignons. Dans ce cas, nous observons une corrélation entre la perte accélérée d'éléments structuraux et la longueur des branches des arbres phylogénétiques basées sur les séquences mitochondriales. En fait, les ARN-Pmt des levures et du zygomycète *S. culisetae* sont les plus réduits (articles 1 et 2); et en parallèle, leur phylogénie mitochondriale montre de longues branches ce qui d'ailleurs crée des artefacts « d'attraction des longues branches » qui rendent la reconstruction phylogénétique difficile (article 2)(Bullerwell et al., 2003; Leigh et al., 2003).

Cette étude nous emmène à poser des questions quant à l'effet de cette réduction structurale sur l'activité catalytique des ARN-Pmt. Est-ce qu'ils sont catalytiques comme leur ancêtre eubactérien en absence de la composante protéique? Si oui, comment cette réduction structurale affecte l'activité catalytique? Sinon, quels éléments sont cruciaux pour maintenir une activité catalytique chez les ARN-P? Une autre observation intrigante est la disparition des hélices P5, P7-10, et P12 en un seul bloc. Pourquoi est-ce que toutes ces hélices disparaissent d'une façon simultanée? Et comment la RNase P compense-t-

elle cette perte? Et finalement, pourquoi le P1 et le P4 sont-ils omniprésents chez toutes les espèces étudiées?



### D.3 Effets de la réduction structurale sur l'activité catalytique des ARN-Pmt

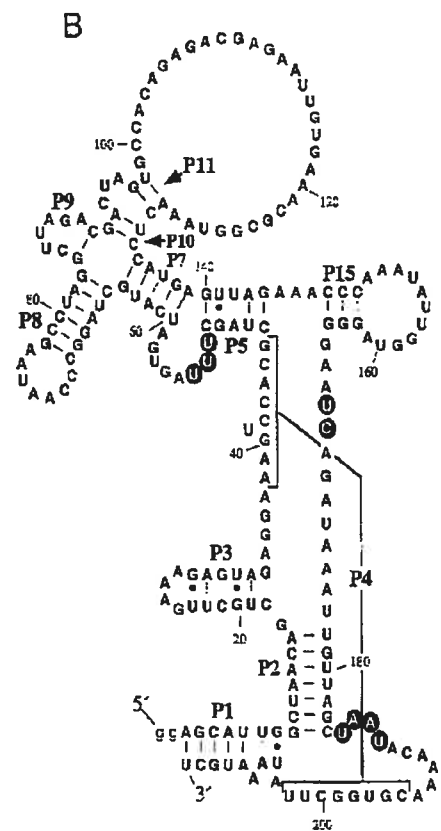
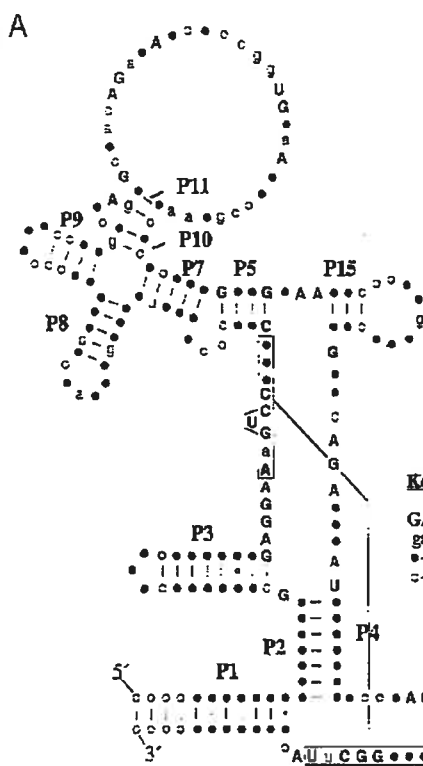
#### D.3.1 Les ARN-Pmt ne sont pas catalytiques *in vitro*

Les modèles de structure secondaire des ARN-Pmt des jakobides, de *N. olivacea*, et du genre *Rhizopus* chez les zygomycètes ont retenu les mêmes hélices qu'on retrouve dans le modèle du consensus minimal de la structure de l'ARN-P eubactérien (Figure D.3A), en plus du P12 et P18, ainsi que les P13, 14 et 19 dans certains cas (voir articles 2 et 3). Un ARN-P avec une structure similaire à ce modèle théorique a été synthétisé *in vitro* à partir de la séquence d'ARN-P de *Mycoplasma fermentas* (Micro ARN-P). Cette molécule a montré de l'activité catalytique 600 fois plus faible que celle de l'ARN natif (Siegel et al., 1996)(Figure D.3B). En se basant sur cette observation, nous avons produit et testé l'activité catalytique des ARN-Pmt de *R. americana* NZ, *J. libera*, *S. ecuadoriensis*, *N. olivacea*, *R. oryzae* et *R. oligosporus in vitro* (voir article 3). Nos tests ont été négatifs pour toutes ces molécules. Chez *S. cerevisiae*, un groupe de chercheurs a déjà purifié la RNase P mitochondriale, et il a rapporté que l'activité catalytique dépend des deux composantes, l'ARN-Pmt et la protéine-P (Hollingsworth et Martin, 1986). Donc, nous estimons que ce soit le cas aussi pour les ARN-Pmt que nous avons étudiés. Présentement, les composantes protéiques de la RNase P mitochondriale chez ces espèces sont inconnues; alors nous avons complété l'ARN-Pmt de *R. americana* NZ avec la protéine-P d'*E. coli*, la protéine C5. Encore une fois, notre résultat était négatif (article 3). Ce résultat n'est pas surprenant vu la distance phylogénétique et évolutionnaire qui sépare *E. coli* et les mitochondries, et, sachant que la protéine-P d'*E. coli* ne peut compléter l'ARN-P du plastide du glaucocystophyte (algue) *Cyanophora paradoxa*, alors que celle de la cyanobactérie *Synechocystis* sp. PCC 6803 en est capable (Pascual et Vioque, 1999). Dans l'avenir, il faudra essayer de compléter les ARN-Pmt avec une protéine-P provenant d'une  $\alpha$ -protéobactérie, le groupe eubactérien le plus proche phylogénétiquement des mitochondries (Gray et al., 1999).

Figure D.3: Modèle du consensus minimal eubactérien (A), et du « Micro ARN-P » (B).

Les nucléotides entourés de noir sont des nucléotides qui remplacent des hélices présentes dans la structure originale de l'ARN-P de *M. fermentas*.

Cette figure a été tirée de (Siegel et al., 1996).



### **D.3.2 Le domaine P6, 16-17 ne suffit pas pour rendre l'ARN-Pmt de *R. americana* NZ actif**

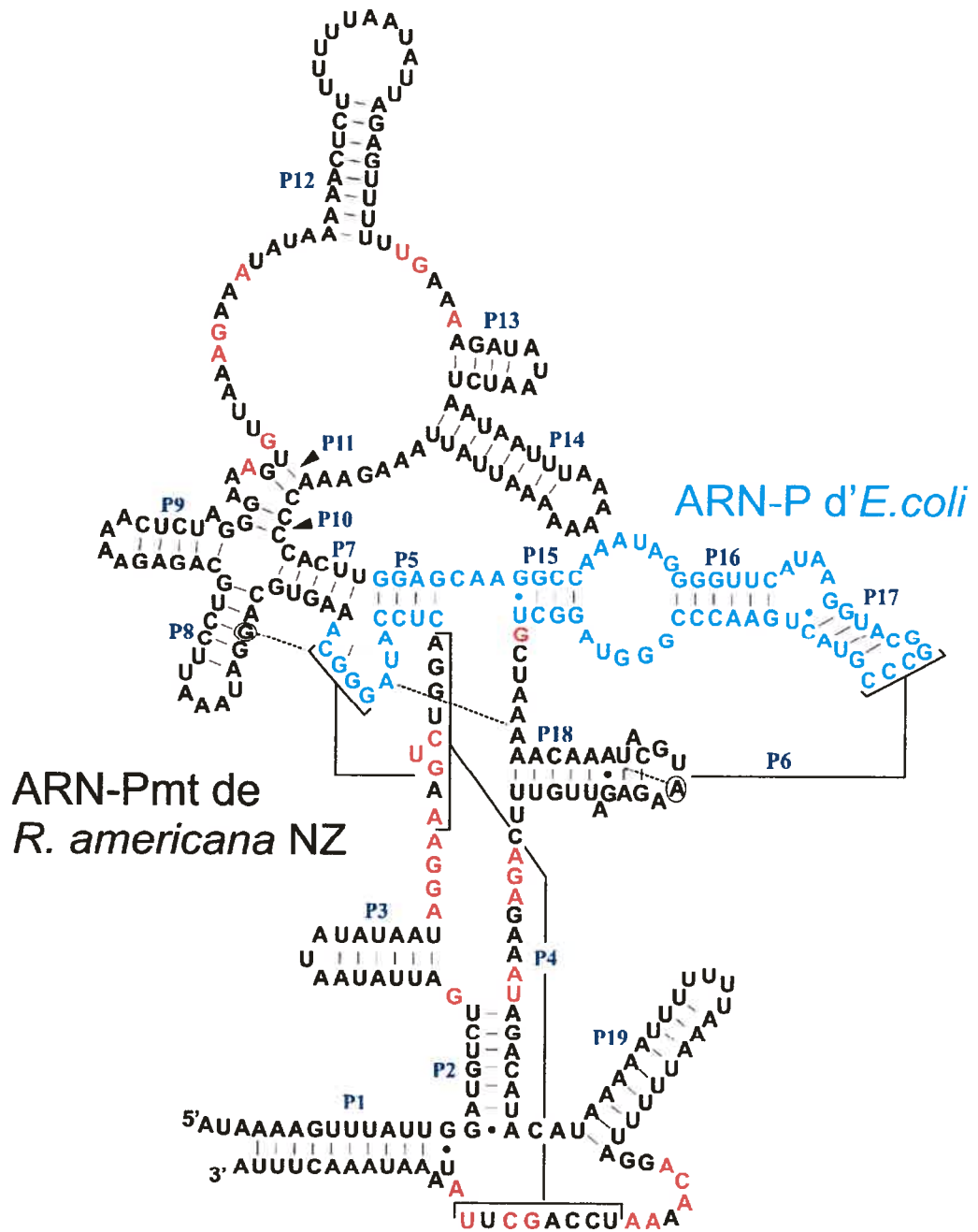
A première vue, le domaine P6, P16-17 qui est présent chez *R. prowazekii* et *E. coli* mais absent chez les ARN-Pmt, pourrait être la raison pour laquelle ces derniers sont non-catalytiques *in vitro* (Figure D.1). En conséquence, nous avons initié une étude de mutagenèse en ajoutant ce domaine provenant de l'ARN-P d'*E. coli* chez l'ARN-Pmt de *R. americana* NZ (Figure D.4). L'ARN-P chimérique que nous avons créé contient deux éléments connus pour avoir un impact majeur sur l'activité catalytique de l'ARN-P (i) l'hélice P6 (Darr et al., 1992; Haas et al., 1991), (ii) et le L15 incluant le motif GGU qui s'attache au pré-ARNt en liant le RCCA en 3' de ce dernier (voir la section I.3.3). Malgré ces ajouts, nous n'avons détecté aucune activité catalytique pour cet ARN-P chimérique (article 3). Ce résultat nous a emmené à conclure que la perte de ce domaine, seule, ne justifie pas l'absence de l'activité catalytique chez les ARN-Pmt. Notre conclusion est renforcée par deux autres observations. Premièrement, le domaine en question existe chez l'ARN-P du plastide de *C. paradoxa* (Shevelev et al., 1995), et ce dernier n'est pas actif *in vitro* (Baum et al., 1996). Deuxièmement, chez le micro ARN-P, ces domaines sont manquants et ce dernier est actif (Siegel et al., 1996)(Figure D.3B).

### **D.3.3 La fusion de l'ARN-P d'*E. coli* avec les ARN-Pmt produit des ARN catalytiques**

Afin de vérifier si certaines hélices ou régions en simple brin manquent des éléments cruciaux pour l'activité catalytique, nous avons construit deux ARN chimériques en fusionnant les séquences de l'ARN-P d'*E. coli* avec celles provenant de l'ARN-Pmt de *R. americana* NZ (article 3). Dans la chimère de type I, la région P1-4, P18, P19, qui selon nos études comparatives est la mieux conservée à travers tous les ARN-Pmt (voir D.2.5), a été fusionnée au domaine complémentaire d'*E. coli* (article 3).

Figure D.4: Chimère d'ARN-P de *R. americana* NZ après l'ajout des P5,6,15,16,17.

La séquence provenant d' *E. coli* est en bleu, et le reste de la séquence provient de l'ARN-Pmt de *R. americana* NZ. Les nucléotides en rouge sont les positions conservées dans le consensus minimal eubactérien (voir la figure D.3).



La chimère de type II a été construite à l'inverse de la chimère de Type I; la région P1-4, P18 provient d'*E. coli* alors que la partie complémentaire provient de *R. americana* NZ. Nous avons trouvé que les deux chimères sont actives (article 3), ce qui suggèrent que chez l'ARN-Pmt de *R. americana* NZ, l'absence d'activité catalytique ne peut être expliquée par le manque d'un seul élément critique.

#### **D.3.4 Raison multifactorielle pour l'absence d'activité catalytique chez les ARN-Pmt**

Les études de mutagenèse que nous avons effectuées sur les chimères de *R. americana* NZ ont montré que l'augmentation du contenu en GC du P4, la correction de l'identité du nucléotide A292 dans le motif universellement conservé « CRV », et l'ajout de motifs GNRA dans la tétra-boucle L9 augmentent l'activité catalytique des chimères (article 3). Ces résultats ainsi que ceux de la section précédente nous emmènent à suggérer que l'absence de l'activité catalytique chez les ARN-Pmt est causée par les facteurs suivants : (i) la séquence riche en A et U des ARN-Pmt, (ii) l'absence des motifs GNRA et du P6, (iii) la déviation même mineur dans les séquences des motifs universellement conservés, (iv) et finalement l'absence du motif GGU dans le L15. Les deux premiers éléments contribuent à la stabilisation de la conformation active de la ribozyme en stabilisant les hélices, et les interactions tertiaires. La déviation des séquences peut avoir un rôle dans la liaison des ions  $Mg^{2+}$  dans le cas du CRV, et l'absence du motif GGU dans le L15 défavorise la liaison au substrat (voir article 3).

Nous proposons que l'effet additif de tous ces éléments rend les ARN-Pmt non catalytiques, puisque chacune des raisons énumérées, à elle seule, ne justifie pas qu'un ARN-P avec une structure aussi conservée que celle de l'ARN-Pmt de *R. americana* NZ soit inactif. En fait, ces éléments ont déjà été rapportés manquants dans des ARN-P catalytiques, (i) les séquences des ARN-P eubactériens de type B sont riches en A et U, (ii) 92 % des modèles eubactériens ont la déviation de G à A dans le CRV à la position correspondante au G350 chez *E. coli* (Rasmussen et Nolan, 2002), (iii) le P6 et les tétra-boucles du motif GNRA qu'on retrouve chez *E. coli* sont absentes chez le micro ARN-P (Siegel et al., 1996), (iv) et finalement, la mutation du motif GGU du L15 n'est pas suffisante pour démolir l'activité catalytique de notre chimère (article 3).

### **D.3.5 Relation entre la conservation de la structure secondaire et l'activité catalytique des chimères**

De la même façon que pour l'ARN-Pmt de *R. americana* NZ, nous avons construit des chimères de type I et II pour les jakobides *J. libera*, *S. ecuadoriensis*, et le prasinophyte *N. olivacea*. Nous avons aussi construit une chimère de type I seulement pour le zygomycète *Rhizopus oryzae* (article 3). Nous avons noté que l'activité catalytique des chimères de type I est plus grande que celle des ARN chimériques de type II. Il a fallu augmenter la concentration d'ARN des chimères de type II d'un facteur dix, pour produire une activité catalytique équivalente à celle des chimères de type I. Cette observation peut être expliquée par le fait que le domaine « C » des ARN-Pmt est mieux conservé que le domaine « S » (voir la section D.2.5), et ceci se traduit par la capacité du domaine « C » de mieux compléter les séquences d'*E. coli*.



Nous avons remarqué aussi que l'activité catalytique des chimères varie selon le degré de conservation des ARN-Pmt originaux (article 3). Chez les chimères de type I, les ARN-Pmt des jakobides forment les chimères les plus actives catalytiquement, suivies par *N. olivacea*. Le zygomycète *R. oryzae* n'a pas montré d'activité (article 3). Pour les chimères de type II, celles provenant des jakobides sont actives, alors que celle de *N. olivacea* ne l'est pas. A l'intérieur des jakobides, les différences d'activité catalytiques sont plus petites, vu que la structure secondaire est très similaire. Les chimères construites pour *S. ecuadoriensis* sont les moins actives. Pour le type I, *J. libera* est plus active que *R. americana* NZ et vice-versa pour le type II. L'observation que la chimère de type I de *J. libera* montre une plus grande activité catalytique que la chimère de *R. americana* NZ, même si les deux ARN-Pmt originaux ont le même nombre d'hélices dans ce domaine, pourrait être expliquée par deux hypothèses. Premièrement, la séquence mitochondriale du domaine « C » de *J. libera* est plus riche en GC que celle de *R. americana* NZ, ce qui pourrait être un avantage si, comme nous avons proposé plus haut (section D.4), le contenu en A et U des ARN-Pmt a un impact négatif sur leur activité catalytique. Deuxièmement, la tétra-boucle du P3 chez *J. libera* appartient à la famille des motifs YNMG comme décrit plus haut. Pour vérifier l'importance de cette boucle, nous avons détruit le motif UUCG en échangeant le G terminal par un A. Suite aux tests catalytiques *in vitro*, nous n'avons noté aucun impact significatif sur l'activité catalytique, et par conséquent nous avons écarté la deuxième hypothèse (résultats non montrés).

En résumé, notre étude montre que les ARN-Pmt ont subi une réduction graduelle de leur structure secondaire, et qu'ils ont perdu des éléments clés pour l'activité catalytique au niveau de leur séquence primaire : les motifs GNRA, la déviation des motifs universellement conservés. Ces déviations de la forme ancestrale font que les ARN-Pmt sont non catalytiques et que leur capacité de compléter les séquences d'*E. coli* est affectée par cette perte. Mais comment cette perte est-elle compensée *in vivo*?

#### D.4 La composante protéique compense-t-elle la réduction structurale?

Chez *S. cerevisiae*, l'ARN-Pmt dépend de la composante protéique pour son activité catalytique (Hollingsworth et Martin, 1986). Nous proposons que ce soit aussi le cas pour les ARN-Pmt étudiés, et que cette protéine compense pour la perte structurale des ARN-Pmt. Pour l'instant, seule la séquence de la protéine-P mitochondriale (protéine-Pmt) de *S. cerevisiae* est connue (Morales et al., 1992). Cette protéine a une taille de 119 kDa, 10 fois plus grande que la protéine-P eubactérienne qui a une taille de 14 kDa, et ne montre pas de similarité avec les séquences eubactériennes. Il a été montré que la protéine-P de *S. cerevisiae* (RPM2) fait partie de la RNase P mitochondriale par des expériences génétique et par immunoprécipitation (Morales et al., 1992). Les études de délétion du gène *rpm2* chez des levures haploïdes ont montré que ce gène est impliqué dans une deuxième fonction essentielle pour la survie de la cellule, indépendante de la RNase P (Kassenbrock et al., 1995). Il a été proposé que la partie C-terminale de la protéine est responsable de la fonction reliée à la composante protéique de la RNase Pmt, alors que la partie N-terminale est impliquée dans la fonction essentielle pour la survie de la levure en présence de sources de carbone fermentable (glucose, éthanol), donc indépendante de la fonction de la RNase P mitochondriale (Stribinskis et al., 2001a; Stribinskis et al., 2001b). La taille relativement grande de la protéine-Pmt peut être expliquée par le fait qu'elle accomplit deux fonctions, et peut-être que cette augmentation de taille compense pour la perte structurale chez la composante ARN.

Pour comprendre comment la protéine-P compense pour la réduction structurale de l'ARN-Pmt, il faudra faire une étude comparative des séquences protéiques provenant de plusieurs groupes phylogénétiques pour élucider l'évolution de cette composante, comme nous avons fait pour l'ARN-Pmt. Ce genre d'étude pourra dévoiler si on retrouve encore des composantes protéiques qui ressemblent à la protéine-P eubactérienne, et on

pourrait faire une corrélation entre la disparition de certains domaines chez l'ARN-Pmt et l'apparition de domaines compensatoires chez la protéine-P. Pour l'instant une telle étude n'est pas possible à cause du manque des séquences, mais grâce aux projets de séquençages génomiques et des banques d'ADNc en cours, une telle étude sera possible dans un avenir rapproché.

## D.5 Scénarios possibles pour expliquer la distribution discontinue du gène *rnpB*

Des gènes *rnpB* ont été identifiés chez tous les jakobides connus, et chez certains zygomycètes et ascomycètes. Mais où se cache ce gène chez la grande majorité des organismes chez lesquels un tel gène n'a pas été trouvé? Est-ce qu'il existe ? Plusieurs scénarios sont possibles pour expliquer cette observation, et dans ce qui suit nous allons discuter ces scénarios, et évaluer la pertinence de chacun en se basant sur les données disponibles.

### D.5.1 Les gènes *rnpB* sont dérivés et non-reconnaissables avec notre approche

Chez *S. pombe*, le gène *rnpB* n'a été identifié qu'après presque deux décennies de la disponibilité de la séquence du génome mitochondrial, à cause de la réduction au niveau de la séquence primaire de ce gène. Alors, il est possible que la séquence de certains gènes soit dérivée, et qu'elle ne soit plus reconnaissable. Comme décrit dans l'article 1, cette hypothèse, bien que possible, n'est surtout pas applicable chez certains chytridiomycètes, où la séquence des génomes mitochondriaux est complètement annotée, et où il ne reste plus assez de régions non codantes pour inclure un gène *rnpB*, à moins que ce gène enchevauche un autre. Deuxièmement, malgré le fait que les gènes *rnpB* chez les champignons subissent une réduction structurale accélérée, nous avons remarqué que ce n'est pas le cas chez les protistes qui évoluent plus lentement. Alors dans les meilleures des chances, ce scénario pourrait être applicable chez les organismes dont le génome mitochondrial évolue vite (en conséquence forment de longues branches dans une phylogénie mitochondriale) mais n'explique pas la répartition globale du gène *rnpB*.

### **D.5.2 Le gène *rnpB* codé par le génome nucléaire**

Si ce gène n'est pas codé par le génome mitochondrial, il est possible qu'il le soit par le génome nucléaire, comme pour la composante protéique de la RNase P mitochondriale de *S. cerevisiae*. En fait, les composantes protéiques et ARN de la RNase MRP (Mitochondrial RNA Processing) sont codées par le noyau et importées par la mitochondrie chez les animaux (Davis et al., 1995; Li et al., 1994). Alors est-ce que c'est le cas pour la RNase P mitochondriale? Même si pour l'instant aucun gène qui code un ARN-Pmt n'a été retrouvé dans les génomes nucléaires complètement séquencés, cette hypothèse est probable et pourrait être vérifiée dans un avenir rapproché vu la multitude de projet de séquençage en cours.

### **D.5.3 Perte de la composante ARN**

Si nous avons de la difficulté à identifier le gène *rnpB*, une raison possible pourrait être l'absence de ce dernier. Car il est possible que la RNase P mitochondriale soit formée d'une composante protéique seulement; ou encore que la RNase P nucléaire soit importée dans la mitochondrie.

#### **D.5.3.1 La composante ARN est absente chez la RNase P du chloroplaste des épinards**

Chez les épinards, l'activité catalytique de la RNase P dans le chloroplaste (RNase P-cp) a été purifiée, et il a été montré que cette activité est résistante aux traitements aux nucléases (Thomas et al., 1995). En plus, les caractéristiques physiques de l'extrait suggèrent l'absence d'acides nucléiques. Finalement, les études sur le mécanisme chimique de coupure des pré-ARNt montrent qu'il s'agit d'un mécanisme

différent de celui adopté par la RNase P eubactérienne et eucaryote (Thomas et al., 2000). Toutes ces observations laissent croire que la RNase P-cp, dans ce cas, ne contient pas de composante ARN.

Ceci suggère qu'il est fort possible que la même situation se soit produite dans la lignée mitochondriale, et que l'ARN-Pmt a été perdu au profit de la protéine-P. Cette hypothèse a même été proposée pour la RNase P mitochondriale chez les humains.

#### **D.5.3.2 La RNase P mitochondriale chez les humains : protéique ou nucléaire?**

Chez les humains, deux hypothèses ont été avancées pour décrire la RNase P mitochondriale. La première, avancée en 1998 par le groupe des docteurs Rossmannith et Karwan, stipule que cette activité est formée d'une protéine seulement (Rossmannith et Karwan, 1998). En 2001, le groupe des docteurs Puranam et Attardi proposa que la RNase P mitochondriale contient une sous-unité ARN identique à celle de la RNase P nucléaire (Puranam et Attardi, 2001).

##### **D.5.3.2.1 RNase P mitochondriale humaine protéique**

L'équipe Rossmannith-Karwan a purifié l'activité catalytique de la RNase Pmt, en utilisant la capacité de cette dernière de cliver le pré-ARNt<sup>tyr</sup> mitochondrial, alors que la RNase P nucléaire (RNase Pnu) ne reconnaît pas ce substrat (Rossmannith et al., 1995). Ils ont rapporté que la RNase Pmt montre les propriétés physiques d'une enzyme protéique en ce qui concerne la vitesse de sédimentation sur un gradient de glycérol, et la densité sur un gradient de Cs<sub>2</sub>SO<sub>4</sub> (Rossmannith et Karwan, 1998). Selon les auteurs, le coefficient de sédimentation de la RNase Pmt est 9S alors que celui de la RNase Pnu est 16S. Ils ont noté que l'activité catalytique mitochondriale résiste aux traitements à la nucléase de *Micrococcus* et à la ribonucléase A. Finalement, ils ont essayé en vain de

détecter des ARN potentiels dans l'extrait par marquage radioactif des extrémités 5' et 3', une technique régulièrement utilisée pour ce genre d'expérience. Selon les auteurs, seul un ARN d'une taille de 25 nucléotides ou moins aurait pu passer inaperçu durant leurs analyses, mais vu que la taille du plus petit ARN-P connu est autour de 140 nucléotides (Wise et Martin, 1991), celui de la levure *Saccharomyces fibuligera*, ils ont exclu la présence d'un ARN-P potentiel. Toutes ces observations les ont emmené à proposer que la mtRNase P est une enzyme protéique; mais ils suggèrent d'interpréter ces résultats avec prudence, car la résistance à la nucléase de *micrococcus* a déjà été rapportées chez la RNase P de l'archéobactérie *Sulfolobus acidocaldarius*, et les nuRNase P de *Xenopus laevis* et de *S. cerevisiae*. Cependant, dans ces cas, la sous-unité ARN a été identifiée par marquage radioactif, ce qui n'est pas le cas pour la RNase Pmt chez les humains (voir les références dans (Rossmanith et Karwan, 1998)).

#### **D.5.3.2.2 L'ARN-Pmt humaine est identique à l'ARN-P de la RNase P nucléaire**

L'équipe du docteur Attardi a identifié en 1985 une activité RNase Pmt humaine en purifiant les mitochondries, et en les traitant avec de la nucléase de *micrococcus* pour éliminer l'activité RNase Pnu contaminante. Cette étude a montré que l'activité catalytique de la RNase Pmt est sensible à la pronase (protéinase) et la nucléase de *micrococcus* (Doersen et al., 1985). Cette étude a été répétée en 2001 en traitant la fraction mitochondriale avec de la digitonine afin de solubiliser la membrane externe et les contaminants nucléaires qu'elle pourrait emmener (Puranam et Attardi, 2001). Les résultats de cette étude confirment ce qui a été rapporté en 1985, et les auteurs ajoutent que la densité de la RNase Pmt est similaire à celle de la RNase Pnu, et que la composante ARN des deux est identique. Dans la publication de leurs résultats, les auteurs proposent que le groupe du docteur Karwan a effectué ses études avec des cellules congelées ce qui cause une dégradation des ARN; et ils ajoutent qu'ils n'ont pas

trouvé l'activité catalytique rapportée dans l'étude de ce groupe qui montre un coefficient de sédimentation différent de la RNase Pnu.

En résumé, il n'existe pas de consensus pour décrire la RNase Pmt chez les humains, chacune des équipes a défendu ses résultats dans une lettre publiée en 2001 (Rossmanith et Potuschak, 2001). Les résultats discordants me semblent être dus au fait que chacune des équipes a caractérisé une activité catalytique différente, si on se fie aux coefficients de sédimentation et à la résistance aux nucléases. Est-ce qu'il existe deux activités RNase P dans la mitochondrie des humains? Si c'est le cas ce serait une première. Est-ce que la rigueur des expérimentations est la cause de ces observations? Pour clarifier la situation, la détermination du gène codant la « RNase Pmt protéique » identifiée par l'équipe du docteur Karwan serait cruciale, et elle incarnerait la première séquence d'une RNase P protéique.



## Conclusions

En conclusion, le projet que nous avons amorcé a permis d'identifier quatre nouveaux gènes *rnpB* chez les ascomycètes, sept chez les zygomycètes, et cinq chez les jakobides. Nous avons établis de nouveaux modèles de structure secondaire pour tous les ARN-P connus par comparaison phylogénétique. La cartographie des extrémités 5' et 3' de six ARN-Pmt a montré que ces ARN sont transcrits et clivés à des positions concordantes avec nos modèles. L'étude structurale comparative nous emmène à conclure que ces modèles sont réduits comparés aux modèles eubactériens, et que cette réduction, perte d'hélices, s'accélère chez les champignons et spécialement chez les ascomycètes et le zygomycète *S. culisetae*. Cette perte ne suit pas un ordre défini et la taille de ces ARN ne reflète pas nécessairement leur état de conservation. Les P1 et P4 sont universellement présents, et le domaine « C » (catalytique) est mieux conservé que le domaine « S » (spécificité). Malgré la grande similarité entre les ARN-P eubactériens et les ARN-Pmt, ces derniers ne sont pas catalytiques. Nous proposons que plusieurs facteurs en sont responsables. La séquence riche en A et U, l'absence des motifs GNRA dans le L9 et L12, l'absence du P6, et du motif GGU dans le L15, et finalement la déviation des séquences consensus dans les régions universellement conservées. Cependant, certains domaines des ARN-Pmt des jakobides et de *N. olivacea* deviennent catalytiques, lorsque fusionnés à des domaines complémentaires provenant d'*E. coli*. Il existe une corrélation entre le degré de conservation de la structure secondaire et l'activité catalytique des chimères. Le domaine « C », chimère type I, complémente ~10 fois mieux les séquences d'*E. coli* que le reste de la molécule, chimère type II. Et les chimères produites avec les séquences d'ARN-Pmt originaux mieux conservés produisent des chimères avec une plus grande activité catalytique.

Nous proposons qu'*in vivo*, la composante protéique compense pour la réduction structurale chez les ARN-Pmt, et nous espérons qu'une étude comparative au niveau des

protéines P révélera la corrélation entre l'évolution des deux composantes. Pour l'instant, les données disponibles ne permettent pas une telle étude, mais grâce aux projets génomiques en cours, une telle étude sera possible dans un avenir rapproché.

Nous n'avons pas réussi à identifier un gène *rnpB* dans les génomes mitochondriaux de la majorité des groupes taxonomiques. Nous proposons plusieurs scénarios pour expliquer cette observation, (i) ces ARN sont trop réduits pour être détectés, (ii) ils sont codés par le génome nucléaire, (iii) la composante ARN n'est plus requise, ou (iv) la RNase P nucléaire effectue la maturation des ARNt mitochondriaux.

## Bibliographie

- Alifano, P., Fani, R., Lio, P., Lazcano, A., Bazzicalupo, M., Carlomagno, M.S., Bruni, C.B., Rivellini, F., Piscitelli, C., Arraiano, C.M., Blasi, V. and Nappo, A.G. (1996) Histidine biosynthetic pathway and genes: structure, regulation, and evolution. *Microbiol Rev*, **60**, 44-69.
- Alifano, P., Piscitelli, C., Blasi, V., Rivellini, F., Nappo, A.G., Bruni, C.B. and Carlomagno, M.S. (1992) Processing of a polycistronic mRNA requires a 5' cis element and active translation. *Mol Microbiol*, **6**, 787-798.
- 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.
- Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H. and Kurland, C.G. (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature*, **396**, 133-140.
- Baer, M.F., Wesolowski, D. and Altman, S. (1989) Characterization in vitro of the defect in a temperature-sensitive mutant of the protein subunit of RNase P from *Escherichia coli*. *J Bacteriol*, **171**, 6862-6866.
- Barrick, J.E., Corbino, K.A., Winkler, W.C., Nahvi, A., Mandal, M., Collins, J., Lee, M., Roth, A., Sudarsan, N., Jona, I., Wickiser, J.K. and Breaker, R.R. (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc Natl Acad Sci USA*, **101**, 6421-6426.
- Bartenschlager, R. and Lohmann, V. (2000) Replication of hepatitis C virus. *J Gen Virol*, **81**, 1631-1648.
- Baum, M., Cordier, A. and Schon, A. (1996) RNase P from a photosynthetic organelle contains an RNA homologous to the cyanobacterial counterpart. *J Mol Biol*, **257**, 43-52.
- Belanger, F., Gagnon, M.G., Steinberg, S.V., Cunningham, P.R. and Brakier-Gingras, L. (2004) Study of the functional interaction of the 900 Tetraloop of 16S ribosomal RNA with helix 24 within the bacterial ribosome. *J Mol Biol*, **338**, 683-693.
- Biswas, R., Ledman, D.W., Fox, R.O., Altman, S. and Gopalan, V. (2000) Mapping RNA-protein interactions in ribonuclease P from *Escherichia coli* using disulfide-linked EDTA-Fe. *J Mol Biol*, **296**, 19-31.
- Bothwell, A.L., Garber, R.L. and Altman, S. (1976) Nucleotide sequence and in vitro processing of a precursor molecule to *Escherichia coli* 4.5 S RNA. *J Biol Chem*, **251**, 7709-7716.
- Bouvet, P. and Belasco, J.G. (1992) Control of RNase E-mediated RNA degradation by 5'-terminal base pairing in *E. coli*. *Nature*, **360**, 488-491.
- Brenowitz, M., Chance, M.R., Dhavan, G., Takamoto, K., Brunel, C. and Romby, P. (2002) Probing the structural dynamics of nucleic acids by quantitative time-resolved and equilibrium hydroxyl radical "footprinting". *Curr Opin Struct Biol*, **12**, 648-653.
- Briani, F., Deho, G., Forti, F. and Ghisotti, D. (2001) The plasmid status of satellite bacteriophage P4. *Plasmid*, **45**, 1-17.

- Briani, F., Zangrossi, S., Ghisotti, D. and Deho, G. (1996) A Rho-dependent transcription termination site regulated by bacteriophage P4 RNA immunity factor. *Virology*, **223**, 57-67.
- Brown, J.W. (1999) The Ribonuclease P Database. *Nucleic Acids Res*, **27**, 314.
- 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.
- Brown, J.W., Nolan, J.M., Haas, E.S., Rubio, M.A., Major, F. and Pace, N.R. (1996) Comparative analysis of ribonuclease P RNA using gene sequences from natural microbial populations reveals tertiary structural elements. *Proc Natl Acad Sci U S A*, **93**, 3001-3006.
- Brown, S. (1991) Genes for 7S RNAs can replace the gene for 4.5S RNA in growth of *Escherichia coli*. *J Bacteriol*, **173**, 1835-1837.
- Brunel, C. and Romby, P. (2000) Probing RNA structure and RNA-ligand complexes with chemical probes. *Methods Enzymol*, **318**, 3-21.
- Bullerwell, C.E., Leigh, J., Forget, L. and Lang, B.F. (2003) A comparison of three fission yeast mitochondrial genomes. *Nucleic Acids Res*, **31**, 759-768.
- Cassano, A.G., Anderson, V.E. and Harris, M.E. (2004) Analysis of solvent nucleophile isotope effects: evidence for concerted mechanisms and nucleophilic activation by metal coordination in nonenzymatic and ribozyme-catalyzed phosphodiester hydrolysis. *Biochemistry*, **43**, 10547-10559.
- 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.
- 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.
- Chen, J.L., Nolan, J.M., Harris, M.E. and Pace, N.R. (1998) Comparative photocross-linking analysis of the tertiary structures of *Escherichia coli* and *Bacillus subtilis* RNase P RNAs. *EMBO J*, **17**, 1515-1525.
- Chen, J.L. and Pace, N.R. (1997) Identification of the universally conserved core of ribonuclease P RNA. *RNA*, **3**, 557-560.
- Chen, Y., Li, X. and Gegenheimer, P. (1997) Ribonuclease P catalysis requires Mg<sup>2+</sup> coordinated to the pro-RP oxygen of the scissile bond. *Biochemistry*, **36**, 2425-2438.
- Cherayil, B., Krupp, G., Schuchert, P., Char, S. and Soll, D. (1987) The RNA components of *Schizosaccharomyces pombe* RNase P are essential for cell viability. *Gene*, **60**, 157-161.
- Christian, E.L. and Harris, M.E. (1999) The track of the pre-tRNA 5' leader in the ribonuclease P ribozyme-substrate complex. *Biochemistry*, **38**, 12629-12638.
- Christian, E.L., Kaye, N.M. and Harris, M.E. (2000) Helix P4 is a divalent metal ion binding site in the conserved core of the ribonuclease P ribozyme. *RNA*, **6**, 511-519.
- Christian, E.L., Kaye, N.M. and Harris, M.E. (2002a) Evidence for a polynuclear metal ion binding site in the catalytic domain of ribonuclease P RNA. *EMBO J*, **21**, 2253-2262.

- Christian, E.L., McPheeters, D.S. and Harris, M.E. (1998) Identification of individual nucleotides in the bacterial ribonuclease P ribozyme adjacent to the pre-tRNA cleavage site by short-range photo-cross-linking. *Biochemistry*, **37**, 17618-17628.
- Christian, E.L., Zahler, N.H., Kaye, N.M. and Harris, M.E. (2002b) Analysis of substrate recognition by the ribonucleoprotein endonuclease RNase P. *Methods*, **28**, 307-322.
- Coburn, G.A. and Mackie, G.A. (1999) Degradation of mRNA in *Escherichia coli*: an old problem with some new twists. *Prog Nucleic Acid Res Mol Biol*, **62**, 55-108.
- Correll, C.C., Munishkin, A., Chan, Y.L., Ren, Z., Wool, I.G., Steitz, T.A. (1998) Crystal structure of the ribosomal RNA domain essential for binding elongation factors. *Proc Natl Acad Sci U S A*, **95**, 13436-13441.
- 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<sup>Asp</sup>. *Biochemistry*, Vol. 37, pp. 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.
- Darr, S.C., Zito, K., Smith, D. and Pace, N.R. (1992) Contributions of phylogenetically variable structural elements to the function of the ribozyme ribonuclease P. *Biochemistry*, **31**, 328-333.
- Davis, A.F., Jeong-Yu, S. and Clayton, D.A. (1995) Distribution of RNase MRP RNA during *Xenopus laevis* oogenesis. *Mol Reprod Dev*, **42**, 359-368.
- de la Cruz, J. and Vioque, A. (2003) A structural and functional study of plastid RNAs homologous to catalytic bacterial RNase P RNA. *Gene*, **321**, 47-56.
- 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.
- Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.P. and Ehresmann, B. (1987) Probing the structure of RNAs in solution. *Nucleic Acids Res*, **15**, 9109-9128.
- Feldstein, P.A., Buzayan, J.M. and Bruening, G. (1989) Two sequences participating in the autolytic processing of satellite tobacco ringspot virus complementary RNA. *Gene*, **82**, 53-61.
- Feng, D.F., Cho, G. and Doolittle, R.F. (1997) Determining divergence times with a protein clock: update and reevaluation. *Proc Natl Acad Sci U S A*, **94**, 13028-13033.
- Ferre-D'Amare, A.R., Zhou, K. and Doudna, J.A. (1998) Crystal structure of a hepatitis delta virus ribozyme. *Nature*, **395**, 567-574.
- Forster, A.C. and Altman, S. (1990a) External guide sequences for an RNA enzyme. *Science*, **249**, 783-786.
- Forster, A.C. and Altman, S. (1990b) Similar cage-shaped structures for the RNA components of all ribonuclease P and ribonuclease MRP enzymes. *Cell*, **62**, 407-409.
- Forster, A.C. and Symons, R.H. (1987) Self-cleavage of virusoid RNA is performed by the proposed 55-nucleotide active site. *Cell*, **50**, 9-16.

- Forti, F., Dragoni, I., Briani, F., Deho, G. and Ghisotti, D. (2002) Characterization of the small antisense CI RNA that regulates bacteriophage P4 immunity. *J Mol Biol*, **315**, 541-549.
- Forti, F., Sabbattini, P., Sironi, G., Zangrossi, S., Deho, G. and Ghisotti, D. (1995) Immunity determinant of phage-plasmid P4 is a short processed RNA. *J Mol Biol*, **249**, 869-878.
- Frank, D.N., Adamidi, C., Ehringer, M.A., Pitulle, C. and Pace, N.R. (2000) Phylogenetic-comparative analysis of the eukaryal ribonuclease P RNA. *RNA*, **6**, 1895-1904.
- Frank, D.N. and Pace, N.R. (1997) In vitro selection for altered divalent metal specificity in the RNase P RNA. *Proc Natl Acad Sci USA*, **94**, 14355-14360.
- Franklin, S.E., Zwick, M.G. and Johnson, J.D. (1995) Characterization and partial purification of two pre-tRNA 5'-processing activities from *Daucus carota* (carrot) suspension cells. *Plant J*, **7**, 553-563.
- Gardiner, K.J., Marsh, T.L. and Pace, N.R. (1985) Ion dependence of the *Bacillus subtilis* RNase P reaction. *J Biol Chem*, **260**, 5415-5419.
- Ghisolfi-Nieto, L., Joseph, G., Puvion-Dutilleul, F., Amalric, F. and Bouvet, P. (1996) Nucleolin is a sequence-specific RNA-binding protein: characterization of targets on pre-ribosomal RNA. *J Mol Biol*, **260**, 34-53.
- Gray, M. and Spencer, D. (1996) Organellar evolution. In Roberts, D., Sharp, P., Alderson, G. and Collins, M. (eds.), *Evolution of Microbial Life*. Cambridge University Press, pp. 109-126.
- Gray, M.W. (1989) Origin and evolution of mitochondrial DNA. *Annu Rev Cell Biol*, **5**, 25-50.
- Gray, M.W., Burger, G. and Lang, B.F. (1999) Mitochondrial evolution. *Science*, **283**, 1476-1481.
- Gray, M.W., Lang, B.F., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., Brossard, N., Delage, E., Littlejohn, T.G., Plante, I., Rioux, P., Saint-Louis, D., Zhu, Y. and Burger, G. (1998) Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Res*, **26**, 865-878.
- Gueneau de Novoa, P. and Williams, K.P. (2004) The tmRNA website: reductive evolution of tmRNA in plastids and other endosymbionts. *Nucleic Acids Res*, **32**, D104-108.
- Guerrier-Takada, C. and Altman, S. (1984) Catalytic activity of an RNA molecule prepared by transcription *in vitro*. *Science*, **223**, 285-286.
- Guerrier-Takada, C., Eder, P.S., Gopalan, V. and Altman, S. (2002) Purification and characterization of Rpp25, an RNA-binding protein subunit of human ribonuclease P. *RNA*, **8**, 290-295.
- 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.
- Guerrier-Takada, C., Haydock, K., Allen, L. and Altman, S. (1986) Metal ion requirements and other aspects of the reaction catalyzed by M1 RNA, the RNA subunit of ribonuclease P from *Escherichia coli*. *Biochemistry*, **25**, 1509-1515.
- Guerrier-Takada, C., van Belkum, A., Pleij, C.W. and Altman, S. (1988) Novel reactions of RNAase P with a tRNA-like structure in turnip yellow mosaic virus RNA. *Cell*, **53**, 267-272.

- Haas, E.S., Armbruster, D.W., Vucson, B.M., Daniels, C.J. and Brown, J.W. (1996) Comparative analysis of ribonuclease P RNA structure in Archaea. *Nucleic Acids Res*, **24**, 1252-1259.
- Haas, E.S. and Brown, J.W. (1998) Evolutionary variation in bacterial RNase P RNAs. *Nucleic Acids Res*, **26**, 4093-4099.
- Haas, E.S., Morse, D.P., Brown, J.W., Schmidt, F.J. and Pace, N.R. (1991) Long-range structure in ribonuclease P RNA. *Science*, **254**, 853-856.
- Hall, T.A. and Brown, J.W. (2001) The ribonuclease P family. *Methods Enzymol*, **341**, 56-77.
- 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.
- Hampel, A. and Tritz, R. (1989) RNA catalytic properties of the minimum (-)sTRSV sequence. *Biochemistry*, **28**, 4929-4933.
- Hansen, A., Pfeiffer, T., Zuleeg, T., Limmer, S., Ciesiolka, J., Feltens, R. and Hartmann, R.K. (2001) Exploring the minimal substrate requirements for trans-cleavage by RNase P holoenzymes from *Escherichia coli* and *Bacillus subtilis*. *Mol Microbiol*, **41**, 131-143.
- 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 Christian, E.L. (2003) Recent insights into the structure and function of the ribonucleoprotein enzyme ribonuclease P. *Curr Opin Struct Biol*, **13**, 325-333.
- Harris, M.E. and Pace, N.R. (1995a) Analysis of the tertiary structure of bacterial RNase P RNA. *Mol Biol Rep*, **22**, 115-123.
- Harris, M.E. and Pace, N.R. (1995b) Identification of phosphates involved in catalysis by the ribozyme RNase P RNA. *RNA*, **1**, 210-218.
- Haseloff, J. and Gerlach, W.L. (1989) Sequences required for self-catalysed cleavage of the satellite RNA of tobacco ringspot virus. *Gene*, **82**, 43-52.
- 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.
- Jackson, R.J. and Wickens, M. (1997) Translational controls impinging on the 5'-untranslated region and initiation factor proteins. *Curr Opin Genet Dev*, **7**, 233-241.
- Jacob, Y., Seif, E., Paquet, P.O. and Lang, B.F. (2004) Loss of the mRNA-like region in mitochondrial tmRNAs of jakobids. *RNA*, **10**, 605-614.
- Kassenbrock, C.K., Gao, G.J., Groom, K.R., Sulo, P., Douglas, M.G. and Martin, N.C. (1995) RPM2, independently of its mitochondrial RNase P function, suppresses an ISP42 mutant defective in mitochondrial import and is essential for normal growth. *Mol Cell Biol*, **15**, 4763-4770.
- Kaye, N.M., Christian, E.L. and Harris, M.E. (2002a) NAIM and site-specific functional group modification analysis of RNase P RNA: magnesium dependent structure

- within the conserved P1-P4 multihelix junction contributes to catalysis. *Biochemistry*, **41**, 4533-4545.
- Kaye, N.M., Zahler, N.H., Christian, E.L. and Harris, M.E. (2002b) Conservation of helical structure contributes to functional metal ion interactions in the catalytic domain of ribonuclease P RNA. *J Mol Biol*, **324**, 429-442.
- Kazantsev, A.V. and Pace, N.R. (1998) Identification by modification-interference of purine N-7 and ribose 2'-OH groups critical for catalysis by bacterial ribonuclease P. *RNA*, **4**, 937-947.
- Kikuchi, Y. (1995) RNase P as hyperprocessing enzyme: a model for formation of a biologically functional tRNA fragment. *Mol Biol Rep*, **22**, 171-175.
- Kikuchi, Y., Ando, Y. and Shiba, T. (1986) Unusual priming mechanism of RNA-directed DNA synthesis in copia retrovirus-like particles of *Drosophila*. *Nature*, **323**, 824-826.
- Kikuchi, Y., Sasaki, N. and Ando-Yamagami, Y. (1990) Cleavage of tRNA within the mature tRNA sequence by the catalytic RNA of RNase P: implication for the formation of the primer tRNA fragment for reverse transcription in copia retrovirus-like particles. *Proc Natl Acad Sci U S A*, **87**, 8105-8109.
- King, T.C., Sirdeskmukh, R. and Schlessinger, D. (1986) Nucleolytic processing of ribonucleic acid transcripts in procaryotes. *Microbiol Rev*, **50**, 428-451.
- Kirsebom, L.A. and Svard, S.G. (1994) Base pairing between *Escherichia coli* RNase P RNA and its substrate. *EMBO J*, **13**, 4870-4876.
- 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.
- Knap, A.K., Wesolowski, D. and Altman, S. (1990) Protection from chemical modification of nucleotides in complexes of M1 RNA, the catalytic subunit of RNase P from *E coli*, and tRNA precursors. *Biochimie*, **72**, 779-790.
- Knapp, G. (1989) Enzymatic approaches to probing of RNA secondary and tertiary structure. *Methods Enzymol*, **180**, 192-212.
- 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.
- 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.
- Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell*, **31**, 147-157.
- Kufel, J. and Kirsebom, L.A. (1996) Different cleavage sites are aligned differently in the active site of M1 RNA, the catalytic subunit of *Escherichia coli* RNase P. *Proc Natl Acad Sci U S A*, **93**, 6085-6090.
- Kurz, J.C., Niranjanakumari, S. and Fierke, C.A. (1998) Protein component of *Bacillus subtilis* RNase P specifically enhances the affinity for precursor-tRNA<sup>Asp</sup>. *Biochemistry*, **37**, 2393-2400.
- LaGrandeur, T.E., Huttenhofer, A., Noller, H.F. and Pace, N.R. (1994) Phylogenetic comparative chemical footprint analysis of the interaction between ribonuclease P RNA and tRNA. *EMBO J*, **13**, 3945-3952.



- 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.
- Lang, B.F., Gray, M.W. and Burger, G. (1999) Mitochondrial genome evolution and the origin of eukaryotes. *Annu Rev Genet*, **33**, 351-397.
- Lee, B., Matera, A.G., Ward, D.C. and Craft, J. (1996a) Association of RNase mitochondrial RNA processing enzyme with ribonuclease P in higher ordered structures in the nucleolus: a possible coordinate role in ribosome biogenesis. *Proc Natl Acad Sci U S A*, **93**, 11471-11476.
- 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, Y.C., Lee, B.J., Hwang, D.S. and Kang, H.S. (1996b) Purification and characterization of mitochondrial ribonuclease P from *Aspergillus nidulans*. *Eur J Biochem*, **235**, 289-296.
- Lee, Y.C., Lee, B.J. and Kang, H.S. (1996c) The RNA component of mitochondrial ribonuclease P from *Aspergillus nidulans*. *Eur J Biochem*, **235**, 297-303.
- Legault, P., Li, J., Mogridge, J., Kay, L.E. and Greenblatt, J. (1998) NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell*, **93**, 289-299.
- Leigh, J., Seif, E., Rodriguez, N., Jacob, Y. and Lang, B.F. (2003) Fungal Evolution meets Fungal Genomics. In Arora, D. (ed.), *Handbook of Fungal Biotechnology*. Marcel Dekker Inc., New York, pp. 145-161.
- Li, K., Smagula, C.S., Parsons, W.J., Richardson, J.A., Gonzalez, M., Hagler, H.K. and Williams, R.S. (1994) Subcellular partitioning of MRP RNA assessed by ultrastructural and biochemical analysis. *J Cell Biol*, **124**, 871-882.
- 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. Epub 12003 Oct 13229.
- Li, Y. and Altman, S. (2004) Polarity effects in the lactose operon of *Escherichia coli*. *J Mol Biol*, **339**, 31-39.
- Liu, F. and Altman, S. (1994) Differential evolution of substrates for an RNA enzyme in the presence and absence of its protein cofactor. *Cell*, **77**, 1093-1100.
- Liu, F. and Altman, S. (1996) Requirements for cleavage by a modified RNase P of a small model substrate. *Nucleic Acids Res*, **24**, 2690-2696.
- Loria, A. and Pan, T. (1997) Recognition of the T stem-loop of a pre-tRNA substrate by the ribozyme from *Bacillus subtilis* ribonuclease P. *Biochemistry*, **36**, 6317-6325.
- Loria, A. and Pan, T. (2001) Modular construction for function of a ribonucleoprotein enzyme: the catalytic domain of *Bacillus subtilis* RNase P complexed with *B. subtilis* RNase P protein. *Nucleic Acids Res*, **29**, 1892-1897.
- Lyons, A.J. and Robertson, H.D. (2003) Detection of tRNA-like structure through RNase P cleavage of viral internal ribosome entry site RNAs near the AUG start triplet. *J Biol Chem*, **278**, 26844-26850.
- Mak, J. and Kleiman, L. (1997) Primer tRNAs for reverse transcription. *J Virol*, **71**, 8087-8095.

- Mandal, M., Boese, B., Barrick, J.E., Winkler, W.C. and Breaker, R.R. (2003) Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell*, **113**, 577-586.
- Margulis, L. (1970) Origin of Eukaryotic Cells. In. Yale Univ. Press., New Haven, CT, p. 349.
- Martin, R.P., Sibler, A.P., Bordonne, R. and Dirheimer, G. (1983) [Yeast mitochondrial transfer RNA. Structure, coding properties and genome organization]. *Mol Biol (Mosk)*, **17**, 1126-1146.
- Martin, W. and Müller, M. (1998) The hydrogen hypothesis for the first eukaryote. *Nature*, **392**, 37-41.
- Massire, C., Jaeger, L. and Westhof, E. (1998) Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. *J Mol Biol*, **279**, 773-793.
- Matsuda, D. and Dreher, T.W. (2004) The tRNA-like structure of Turnip yellow mosaic virus RNA is a 3'-translational enhancer. *Virology*, **321**, 36-46.
- 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.
- Moreira, D. and Lopez-Garcia, P. (1998) Symbiosis between methanogenic archaea and delta-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. *J Mol Evol*, **47**, 517-530.
- Morrissey, J.P. and Tollervey, D. (1995) Birth of the snoRNPs: the evolution of RNase MRP and the eukaryotic pre-rRNA-processing system. *Trends Biochem Sci*, **20**, 78-82.
- Nadal, A., Martell, M., Lytle, J.R., Lyons, A.J., Robertson, H.D., Cabot, B., Esteban, J.I., Esteban, R., Guardia, J. and Gomez, J. (2002) Specific cleavage of hepatitis C virus RNA genome by human RNase P. *J Biol Chem*, **277**, 30606-30613.
- Nagai, K., Oubridge, C., Kuglstatter, A., Menichelli, E., Isel, C. and Jovine, L. (2003) Structure, function and evolution of the signal recognition particle. *EMBO J*, **22**, 3479-3485.
- Nahvi, A., Barrick, J.E. and Breaker, R.R. (2004) Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res*, **32**, 143-150. Print 2004.
- Nakamura, K., Miyamoto, H., Suzuma, S., Sakamoto, T., Kawai, G. and Yamane, K. (2001) - Minimal functional structure of Escherichia coli 4.5 S RNA required for. *J Biol Chem*, Vol. 276, pp. 22844-22849.
- Niranjanakumari, S., Stams, T., Cray, 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.
- Nolan, J.M., Burke, D.H. and Pace, N.R. (1993) Circularly permuted tRNAs as specific photoaffinity probes of ribonuclease P RNA structure. *Science*, **261**, 762-765.
- Oh, B.K., Frank, D.N. and Pace, N.R. (1998) Participation of the 3'-CCA of tRNA in the binding of catalytic Mg<sup>2+</sup> ions by ribonuclease P. *Biochemistry*, **37**, 7277-7283.
- Pace, N.R., Smith, D.K., Olsen, G.J. and James, B.D. (1989) Phylogenetic comparative analysis and the secondary structure of ribonuclease P RNA--a review. *Gene*, **82**, 65-75.

- Pan, T., Loria, A. and Zhong, K. (1995) Probing of tertiary interactions in RNA: 2'-hydroxyl-base contacts between the RNase P RNA and pre-tRNA. *Proc Natl Acad Sci U S A*, **92**, 12510-12514.
- 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.
- Paquin, B., Laforest, M.J. and Lang, B.F. (2000) Double-hairpin elements in the mitochondrial DNA of allomyces: evidence for mobility. *Mol Biol Evol*, **17**, 1760-1768.
- Pascual, A. and Vioque, A. (1999) Functional reconstitution of RNase P activity from a plastid RNA subunit and a cyanobacterial protein subunit. *FEBS Lett*, **442**, 7-10.
- Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A. and Cheng, H.L. (1986) A self-splicing RNA excises an intron lariat. *Cell*, **44**, 213-223.
- Pleij, C.W., Abrahams, J.P., van Belkum, A., Riefveld, K. and Bosch, L. (1987) The spatial folding of the 3' non-coding region of aminoacylatable plant viral RNAs. In Brinton, M.A. and Rueckert, R. (eds.), *Positive Strand Viruses*. Alan R Liss, New York, pp. 299-316.
- Pomeranz Krummel, D.A. and Altman, S. (1999) Verification of phylogenetic predictions in vivo and the importance of the tetraloop motif in a catalytic RNA. *Proc Natl Acad Sci U S A*, **96**, 11200-11205.
- Poritz, M.A., Strub, K. and Walter, P. (1988) Human SRP RNA and E. coli 4.5S RNA contain a highly homologous structural domain. *Cell*, **55**, 4-6.
- Proctor, D.J., Schaak, J.E., Bevilacqua, J.M., Falzone, C.J. and Bevilacqua, P.C. (2002) Isolation and characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions. *Biochemistry*, **41**, 12062-12075.
- 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.
- Rasmussen, T.A. and Nolan, J.M. (2002) G350 of *Escherichia coli* RNase P RNA contributes to Mg<sup>2+</sup> binding near the active site of the enzyme. *Gene*, **294**, 177-185.
- Ribes, V., Romisch, K., Giner, A., Dobberstein, B. and Tollervey, D. (1990) *E. coli* 4.5S RNA is part of a ribonucleoprotein particle that has properties related to signal recognition particle. *Cell*, **63**, 591-600.
- Rossmannith, W. and Karwan, R.M. (1998) Characterization of human mitochondrial RNase P: novel aspects in tRNA processing. *Biochem Biophys Res Commun*, **247**, 234-241.
- Rossmannith, W. and Potuschak, T. (2001) Difference between mitochondrial RNase P and nuclear RNase P. *Mol Cell Biol*, **21**, 8236-8237.
- Rossmannith, W., Tullo, A., Potuschak, T., Karwan, R. and Sbisà, E. (1995) Human mitochondrial tRNA processing. *J Biol Chem*, **270**, 12885-12891.
- Saenger. (1984) *Principles of nucleic acid structure*.
- Sagar, M.B., Lucast, L. and Doudna, J.A. (2004) Conserved but nonessential interaction of SRP RNA with translation factor EF-G. *RNA*, **10**, 772-778.

- Saville, B.J., Collins, R.A., Wu, H.N., Lin, Y.J., Lin, F.P., Makino, S., Chang, M.F. and Lai, M.M. (1990) A site-specific self-cleavage reaction performed by a novel RNA in *Neurospora* mitochondria. *Cell*, **61**, 685-696.
- Schlegl, J., Hardt, W.D., Erdmann, V.A. and Hartmann, R.K. (1994) Contribution of structural elements to *Thermus thermophilus* ribonuclease P RNA function. *EMBO J*, **13**, 4863-4869.
- Schmitz, M. and Tinoco, I., Jr. (2000) Solution structure and metal-ion binding of the P4 element from bacterial RNase P RNA. *RNA*, **6**, 1212-1225.
- Shen, L.X., Cai, Z. and Tinoco, I., Jr. (1995) RNA structure at high resolution. *Faseb J*, **9**, 1023-1033.
- 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.
- Siegel, R.W., Banta, A.B., Haas, E.S., Brown, J.W. and Pace, N.R. (1996) *Mycoplasma fermentans* simplifies our view of the catalytic core of ribonuclease P RNA. *RNA*, **2**, 452-462.
- 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.
- Smith, D. and Pace, N.R. (1993) Multiple magnesium ions in the ribonuclease P reaction mechanism. *Biochemistry*, **32**, 5273-5281.
- Stark, B.C., Kole, R., Bowman, E.J. and Altman, S. (1978) Ribonuclease P: an enzyme with an essential RNA component. *Proc Natl Acad Sci US A*, **75**, 3717-3721.
- Stribinskis, V., Gao, G.J., Ellis, S.R. and Martin, N.C. (2001a) Rpm2, the protein subunit of mitochondrial RNase P in *Saccharomyces cerevisiae*, also has a role in the translation of mitochondrially encoded subunits of cytochrome c oxidase. *Genetics*, **158**, 573-585.
- 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. (2001b) Rpm2p: separate domains promote tRNA and Rpm1r maturation in *Saccharomyces cerevisiae* mitochondria. *Nucleic Acids Res*, **29**, 3631-3637.
- Talla, E., Anthouard, V., Bouchier, C., Frangeul, L. and Dujon, B. (2005) The complete mitochondrial genome of the yeast *Kluyveromyces thermotolerans*. *FEBS Lett*, **579**, 30-40.
- Thomas, B.C., Gao, L., Stomp, D., Li, X. and Gegenheimer, P.A. (1995) Spinach chloroplast RNase P: a putative protein enzyme. *Nucleic Acids Symp Ser*, 95-98.
- 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.
- Tranguch, A.J. and Engelke, D.R. (1993) Comparative structural analysis of nuclear RNase P RNAs from yeast. *J Biol Chem*, **268**, 14045-14055.
- Turmel, M., Lemieux, C., Burger, G., Lang, B.F., Otis, C., Plante, I. and Gray, M.W. (1999) The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*. Two radically different evolutionary patterns within green algae. *Plant Cell*, **11**, 1717-1730.

- Tyagi, J.S. and Kinger, A.K. (1992) Identification of the 10Sa RNA structural gene of *Mycobacterium tuberculosis*. *Nucleic Acids Res*, **20**, 138.
- Underbrink-Lyon, K., Miller, D.L., Ross, N.A., Fukuhara, H. and Martin, N.C. (1983) Characterization of a yeast mitochondrial locus necessary for tRNA biosynthesis. Deletion mapping and restriction mapping studies. *Mol Gen Genet*, **191**, 512-518.
- Ushida, C., Himeno, H., Watanabe, T. and Muto, A. (1994) tRNA-like structures in 10Sa RNAs of *Mycoplasma capricolum* and *Bacillus subtilis*. *Nucleic Acids Res*, **22**, 3392-3396.
- Vioque, A., Arnez, J. and Altman, S. (1988) Protein-RNA interactions in the RNase P holoenzyme from *Escherichia coli*. *J Mol Biol*, **202**, 835-848.
- Wagner, M., Fingerhut, C., Gross, H.J. and Schon, A. (2001) The first phytoplasma RNase P RNA provides new insights into the sequence requirements of this ribozyme. *Nucleic Acids Res*, **29**, 2661-2665.
- Waugh, D.S. and Pace, N.R. (1990) Complementation of an RNase P RNA (*rnpB*) gene deletion in *Escherichia coli* by homologous genes from distantly related eubacteria. *J Bacteriol*, **172**, 6316-6322.
- Weiner, A.M. and Maizels, N. (1987) tRNA-like structures tag the 3' ends of genomic RNA molecules for replication: implications for the origin of protein synthesis. *Proc Natl Acad Sci U S A*, **84**, 7383-7387.
- Welting, T.J., van Venrooij, W.J. and Pruijn, G.J. (2004) Mutual interactions between subunits of the human RNase MRP ribonucleoprotein complex. *Nucleic Acids Res*, **32**, 2138-2146. Print 2004.
- Wientges, J., Putz, J., Giege, R., Florentz, C. and Schwienhorst, A. (2000) Selection of viral RNA-derived tRNA-like structures with improved valylation activities. *Biochemistry*, **39**, 6207-6218.
- Williams, K.P. and Bartel, D.P. (1996) Phylogenetic analysis of tmRNA secondary structure. *RNA*, **2**, 1306-1310.
- 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.
- Wise, C.A. and Martin, N.C. (1991) Dramatic size variation of yeast mitochondrial RNAs suggests that RNase P RNAs can be quite small. *J Biol Chem*, **266**, 19154-19157.
- Wu, H.N., Lin, Y.J., Lin, F.P., Makino, S., Chang, M.F. and Lai, M.M. (1989) Human hepatitis delta virus RNA subfragments contain an autocleavage activity. *Proc Natl Acad Sci U S A*, **86**, 1831-1835.
- 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.
- Zuker, M. (1989) On finding all suboptimal foldings of an RNA molecule. *Science*, **244**, 48-52.