HAL
open science

Caractérisation biochimique et structurale des RNases P et MRP chez la levure Saccharomyces cerevisiae

Claire Blachier Batisse

To cite this version:

Claire Blachier Batisse. Caractérisation biochimique et structurale des RNases P et MRP chez la levure Saccharomyces cerevisiae. Sciences agricoles. Université de Strasbourg, 2013. Français. NNT : 2013STRAJ001 . tel-00804257

HAL Id: tel-00804257

https://theses.hal.science/tel-00804257v1

Submitted on 25 Mar 2013

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

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

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE Sciences de la Vie et de la Santé

Service VAE

THÈSE présentée par :

Claire Batisse

soutenue le : **23 Janvier 2013**

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

Discipline/ Spécialité : Sciences de la Vie et de la Santé

spécialité Aspects moléculaires et cellulaires de la biologie

Caractérisation biochimique et structurale des RNases P et MRP chez la levure *Saccharomyces cerevisiae*

Président du jury :

Monsieur Pr. Hugues Dreyssé	Université de Strasbourg

Rapporteurs :

Monsieur le Dr. Jean Paul Renaud	Entreprise Novalyx
Monsieur le Dr. Bertrand Séraphin	Université de Strasbourg
Monsieur le Pr. Bruno Charpentier	Université de Lorraine
Madame le Dr. Béatrice Golinelli-Pimpaneau	Collège de France

Superviseur scientifique :

Monsieur le Dr. Philippe Giegé	Université de Strasbourg

Superviseur VAE :

Madame Danielle Haug	Université de Strasbourg

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude à l'ensemble des membres de mon jury pour avoir accepter d'évaluer ma demande de validation des acquis de l'expérience.

Je voudrais également exprimer ma gratitude à Mme Danielle Haug qui m'a accompagnée tout au long de cette démarche en me guidant avec ses critiques constructives et ses encouragements.

Je remercie plus particulièrement le Dr. Philippe Giégé pour m'avoir accompagnée scientifiquement dans cette aventure par ses conseils judicieux et son enthousiasme scientifique.

Je remercie le FONGECIF Alsace qui m'a soutenue en finançant le coût de l'accompagnement universitaire de cette démarche.

Je remercie le Pr. Calender pour m'avoir accueillie au sein de son laboratoire où j'ai découvert la recherche en biologie humaine. Je remercie le Dr. Stéphane Pinson pour m'avoir fait partager sa passion pour la NF1, pour m'avoir laissé une grande liberté d'entreprise et avoir toujours été prêt à tester de nouvelles approches. Je remercie tous les membres du laboratoire pour leur gentillesse, leur disponibilité et pour les moments que nous avons partagés autour d'une tablette de chocolat.

Je remercie le Pr. Ed Hurt pour m'avoir accueillie dans son équipe en Allemagne quand mon aptitude pour les langues étrangères était encore à l'état de balbutiement, pour m'avoir permis d'entrer dans le monde de la recherche fondamentale et m'avoir donné les bases pour aller vers d'autres projets. Je remercie l'ensemble des membres passés et présents de l'équipe Hurt et plus particulièrement Andréa Schliwa et Michal Skruzny pour leur accueil chaleureux et leur souplesse d'esprit qui m'ont permis de prendre un bon départ en Allemagne et de communiquer, même par geste quand les mots manquaient.

Je tiens à exprimer ma profonde reconnaissance au Dr. Bettina Böttcher pour m'avoir donné la possibilité de travailler dans un environnement scientifique de qualité en m'accueillant au sein de son équipe à l'EMBL, pour m'avoir associée à de nombreux projets, pour m'avoir toujours encouragée dans mes initiatives et m'avoir guidée durant ces années

allemandes où j'ai pu ainsi progresser dans mon compréhension de la recherche fondamentale. Je remercie les Drs. Katharina Hipp et Kyriaki Galani pour leur contribution au projet RNases P/MRP, pour leur bonne humeur et leur enthousiasme qui m'ont soutenue face au challenge présenté par cette étude. Enfin, je remercie les autres membres de l'équipe Böttcher ainsi que tous nos collaborateurs pour leur patience, leur ouverture d'esprit et leur disponibilité.

Je remercie le Dr. Marcus Heisler pour m'avoir accueillie dans son équipe et m'avoir permis de découvrir l'univers de la biologie végétale, pour m'avoir fait confiance pour installer son laboratoire et pour sa patience face à toutes mes questions sur la biologie végétale et la biologie du développement. Merci à Carolyn et Monica pour nos discussions autour de la vie en Allemagne et l'arrivée d'un bébé.

Je remercie le Dr. Patrick Schultz pour m'avoir accueillie dans son équipe qui me permet de profiter d'un environnement de haute qualité scientifique. Je tiens à exprimer ma gratitude au Dr. Valérie Lamour pour m'avoir accueillie dans son groupe à mon retour d'Allemagne, pour m'avoir soutenue dans cette démarche de VAE et pour sa franchise et son enthousiasme en toutes circonstances. Je remercie Julie Papillon et Claire Bedez pour leur écoute, leur patience, leur compréhension et particulièrement pour le dynamisme dont elles m'entourent.

Je remercie tous les membres du département de biologie structurale intégrative et particulièrement ceux du premier étage sud de l'IGBMC pour les gâteaux du Jeudi, pour leur patience face à mes questions de vocabulaire et de formulations et pour toutes nos discussions si riche d'échanges sur la science et sur la vie alsacienne.

Enfin, merci à ma famille qui a compris ma démarche, m'a soutenue et m'a laissé du temps pour la finaliser. Un grand merci à Julien pour son soutien de chaque instant, ses relectures pertinentes et sa patience durant ces années de paillasse et ce temps de rédaction. Enfin, merci à ma fille Laure pour avoir enrichi ma vie.

SOMMAIRE

Remerciements i

SOMMAIRE iii

INTRODUCTION 1

1. Les RNases P bactériennes 5

1.1. Structure des RNases P bactériennes 5
1.1.1. Caractérisation de la composante ARN 5
1.1.1.1. Les différents types de composantes ARN des RNases P bactériennes 5
1.1.1.2. L'organisation structurale de la composante ARN des RNases P bactériennes 7
1.1.1.3. Caractérisation fonctionnelle de la composante ARN des RNases P bactériennes 9
1.1.2. Caractérisation de la composante protéique 10
1.1.2.1. Les différentes fonctions de la composante protéique des RNases P bactériennes 10
1.1.2.2. L'organisation structurale de la composante protéique 11
1.1.2.3. Interaction entre la composante protéique et la composante ARN 11
1.2. Substrats de la RNase P bactérienne 12
1.2.1. Reconnaissance du substrat 12
1.2.2. Mécanisme de coupure du substrat 15

2. Les RNases P archées 17

2.1. Caractérisation de la composante ARN 17
2.2. Caractérisation de la composante protéique archée 19
2.2.1. Caractérisation fonctionnelle des sous-unités de la composante protéique archée 19
2.2.2. Caractérisation structurale de la composante protéique archée 21

3. Les RNases P eucaryotes 25

3.1. Les RNases P ribonucléoprotéiques 25
3.1.1. Les RNases P ribonucléoprotéiques nucléaires 25
3.1.1.1. Structure des RNases P nucléaires 25
3.1.1.1.1. Composition des RNases P nucléaires 25
3.1.1.1.2. Caractérisation de la composante ARN des RNases P nucléaires 29
3.1.1.1.2.1. Organisation structurale de la composante ARN 29
3.1.1.1.2.2. Caractérisation fonctionnelle de la composante ARN 34
3.1.1.1.3. Caractérisation structurale des RNases P nucléaires 34
3.1.1.1.3.1. L'assemblage in vivo de la RNase P nucléaire 35
3.1.1.1.3.2. Les interactions ARN-protéine au sein de la RNase P nucléaire 36
3.1.1.1.3.3. Les interactions protéine-protéine au sein de la RNase P nucléaire 37
3.1.1.1.4. Etude des substrats des RNases P nucléaires 39
3.1.1.1.4.1. Particularités dans la reconnaissance du substrat 39
3.1.1.1.4.2. Particularités dans la coupure du substrat 42
3.1.2. La RNase P RNP mitochondriale de *S. cerevisiae* 43
3.2. Les RNases P protéiques 43
3.2.1. La RNase P mitochondriale multiprotéique de *H. sapiens* 44

3.2.2. Les RNases P chez les végétaux 44
3.2.3. Les RNases P de Trypanosoma brucei 45
3.3. Les RNases MRP 46
3.3.1. Structure des RNases MRP 46
3.3.1.1. Composition des RNases MRP 46
3.3.1.2. Caractérisation de la composante ARN des RNases MRP 48
3.3.1.2.1. Structure de la composante ARN 48
3.3.1.2.2. Mutation de la composante ARN associées à des maladies humaines 49
3.3.1.3. Caractérisation structurale des RNases MRP 50
3.3.2. Etude des substrats des RNases MRP 51
3.3.2.1. Les différents substrats connus des RNases MRP 51
3.3.2.2. Particularité dans la reconnaissance du substrat 53

4. Description du sujet de recherche 55

MATERIELS ET METHODES 57

1. Description des souches de levure *S. saccharomyces* 59

1.1. Préparation des cassettes à intégrer 59
1.2. Transformation par la méthode LiAc/SS DNA/PEG 60
1.3. Test de l'intégration génomique d'une étiquette à la localisation souhaitée 61
1.4. Test de l'expression des protéines recombinantes 61

2. Purification des complexes RNases P et MRP 62

2.1. Purification des complexes natifs 62
2.2. Purification de complexes marqués avec un anticorps 62

3. Méthodes de caractérisations biochimiques 63

3.1. Western-blot 63
3.2. Spectrométrie de masse 63
3.3. RT-PCR quantitative 63
3.4. Synthèse *in vitro* des pré-ARNt 64
3.5. Test de l'activité endonucléolytique *in vitro* de la RNase P purifiée 65
3.6. Test de la capacité de liaison des pré-ARNt à la RNase P purifiée 65

4. Microscopie électronique à transmission 66

4.1. Préparation des échantillons 66
4.2. Traitement des images des complexes RNase P et MRP 66
4.3. Traitement des images des complexes RNases P marqués avec un anticorps 67
4.4. Positionnement des modèles atomiques dans la carte de microscopie électronique 67

RESULTATS **69**

1. Expression et caractérisation biochimique des complexes RNase P et MRP **71**

1.1. Marquage et Expression des protéines d'intérêt 71
1.1.1. Principe du marquage de protéine de levure par recombinaison homologue 71
1.1.2. Résultats 71
1.1.2.1. Test de l'intégration de l'étiquette 73
1.1.2.2. Test d'expression 73
1.2. Purification séquentielle des complexes RNase P et MRP 75
1.2.1. Principe de la purification 75
1.2.2. Vérification de l'efficacité de la purification 75
1.2.3. Détermination de la composition des complexes purifiés 77
1.2.4. Quantification des noyaux ARN RPR1 et NME1 purifiés 77
1.2.5. Contrôle de la qualité des échantillons purifiés par microscopie électronique 79
1.2.5.1. La méthode de coloration négative en sandwich 79
1.2.5.2. Résultats 80
1.3. Test de l'activité de coupure endonucléolytique *in vitro* de la RNase P purifiée 81
1.4. Test de la capacité de liaison des pré-ARNt à la RNase P purifiée 81
1.5. Conclusions 82

2. Caractérisation structurale des complexes RNases P et MRP **83**

2.1. Modélisation des complexes RNases P et MRP 83
2.1.1. Modélisation initiale par reconstruction conique aléatoire 83
2.1.1.1. Principe de la reconstruction conique aléatoire (Radermacher *et al.* , 1987) 83
2.1.1.2. Obtention d'une carte initiale de la surface des complexes RNases P et MRP 85
2.1.2. Etablissement de cartes tridimensionnelles des RNases P et MRP 87
2.1.3. Une structure modulaire commune aux deux complexes 87
2.1.3.1. Le module de base 89
2.1.3.2. Le module de tête 89
2.1.3.3. Le module arrière 91
2.2. Localisation structurale des sous-unités protéiques 91
2.2.1. Localisation de différentes sous-unités de la RNase P par marquage aux anticorps. 91
2.2.2. Positionnement des modèles atomiques 93
2.2.2.1. Pop5 et Rpp1 forment le module arrière 96
2.2.2.2. La sous-unité ARN forme la plaque basale du domaine de base 96
2.2.2.3. Pop4 et Rpr2/Snm1 stabilisent le domaine S de la sous-unité ARN 98
2.2.2.4. Pop3 contribue au bras du module de base par liaison à l'hélice P12 99
2.2.2.5. Pop6 et Pop7 forment le domaine au pied du module de base 99
2.3. Conclusions 100

PUBLICATION **101**

1. Préambule **103**

2. Publication **104**

DISCUSSION **105**

1. Purification de complexes RNases P et MRP intacts et fonctionnels **107**

2. Rôle des protéines dans les complexes RNases P/MRP **107**

2.1. Protéines soutenant la structure de la composante ARN chez les eucaryotes 107
2.2. Protéines contribuant à la liaison du substrat 108

3. Implication de l'architecture des complexes RNases P/MRP dans la coordination du substrat **111**

4. Différences architecturales des complexes des RNases P/MRP et spécificité des substrats **114**

5. Corrélation entre l'architecture des complexes P/MRP et les interactions binaires connues **115**

6. Conclusion et perspectives **116**

ANNEXES **119**

Annexe I : Tableau récapitulatif des souches de *S. cerevisiae* utilisées dans cette étude..... 121
Annexe II : Amorces utilisées pour réaliser les intégrations génomiques des cassettes 3HA et yeGFP. 123

ABREVIATIONS **125**

Autres travaux de recherche **127**

Références bibliographiques du sujet de recherche **131**

INTRODUCTION
La famille des RNases P

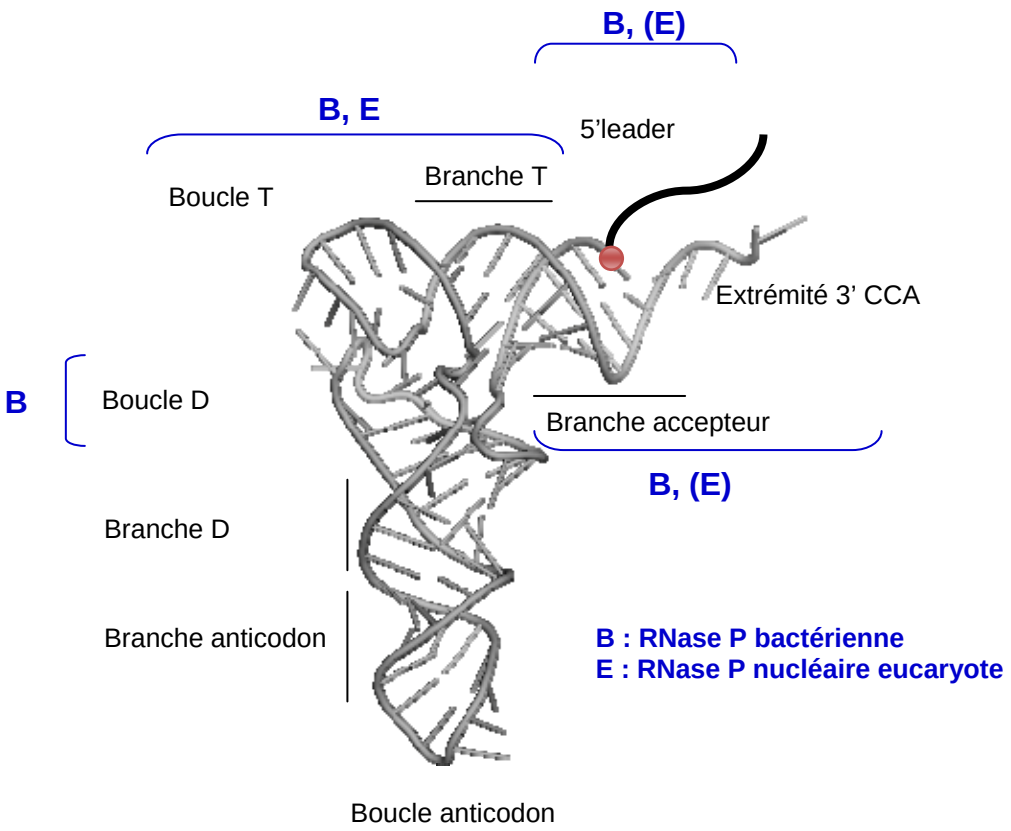
Le monde du vivant est principalement organisé en trois niveaux dans les cellules :

- L'acide désoxyribonucléique (ADN), support de l'information génétique,
- Les protéines, constituant de la cellule et acteurs de son fonctionnement,
- Les acides ribonucléiques (ARN), molécules principalement intermédiaires entre l'ADN et les protéines (ARN messager dit ARNm, ARN de transfert dit ARNt, ARN ribosomique dit ARNr). Ces ARN, transcrits à partir d'ADN, suivent un cycle de vie avec une phase de maturation, d'exercice de leurs activités et enfin de dégradation. Les processus de maturation et de dégradation sont principalement réalisés par des enzymes appelées des ribonucléases (RNases). Si la plupart des RNases sont de nature protéique, il a été mis en évidence en 1983 que lors de la maturation des ARNt, la séquence 5' leader était coupée par une enzyme, la RNase P, dont l'activité catalytique est portée par sa composante ARN (Guerrier-Takada *et al.*, 1983). La découverte de cette activité dite "ribozyme" a contribué à bouleverser le dogme selon lequel seules les protéines pouvaient être porteuses d'une activité enzymatique. D'autres ribozymes ont été découverts depuis. Cependant, ces ribozymes sont presque toujours des complexes ribonucléoprotéiques (RNP) avec une composante ARN catalytique et une composante protéique qui permet d'assurer l'efficacité, la spécificité et la localisation de la réaction enzymatique.

La RNase P a été depuis 30 ans un des complexes ribonucléoprotéiques les mieux étudiés, avec le ribosome. La composante ARN de la RNase P est impliquée dans la maturation des ARNt chez de nombreux organismes appartenant aux trois grands règnes de la vie (bactérien, archée et eucaryote) ainsi que dans certaines mitochondries et chloroplastes (Piccinelli *et al.*, 2005; Rosenblad *et al.*, 2006). Ces ARNt prématures ont une séquence primaire variée mais partagent une structure secondaire en forme de trèfle avec des éléments importants pour la réaction de coupure (**Figure 1**). La forte homologie des composants ARN de la RNase P entre ces espèces suggère que cette ancienne enzyme devait déjà être présente dans le dernier ancêtre commun à tous les organismes connus aujourd'hui. En outre, l'ARN de la RNase P, les ARN d'épissage nucléaire (snRNA) et l'ARN ribosomal sont les seuls ARN décrits comme ayant une activité catalytique capable de réaliser de multiples cycles catalytiques ce qui en fait de véritables enzymes (Kazantsev *et al.*, 2006), les autres ribozymes étant auto-réactifs. Cela suggère que cet ARN pourrait être une réminiscence d'un monde prébiotique principalement à base d'ARN (Hartmann *et al.*, 2003; Sun *et al.*, 2010).

Si la composante ARN de la RNase P est la partie catalytique très conservée, cette holoenzyme se compose aussi d'une partie protéique qui affecte de manière significative

l'activité de la RNase P et est universellement requise *in vivo*. Au cours de l'évolution, l'importance et la complexité de la composante protéique s'est accrue substantiellement, s'adaptant à la complexification de l'environnement cellulaire avec l'ajout d'un nombre variable de sous-unités protéiques : une seule protéine chez les bactéries, quatre à cinq protéines chez les archées et au moins 9 protéines chez les eucaryotes. Outre son rôle dans la maturation des ARNt, la RNase P est aussi impliquée dans la coupure d'autres substrats *in vitro* et *in vivo* selon les organismes.



Adaptée de Marvin *et al*, 2009

Figure 1. Représentation des régions structurales de l'ARNt prémature reconnues par les différentes RNases P.
La structure du pré-ARNt présentée est la structure à haute résolution de l'ARNt de levure (PDB 1EHZ, **(Shi *et al.*, 2000)**). Le site de coupure de la séquence 5' leader est indiqué par un point rouge. Les parenthèses autour d'un type de RNase P indiquent des interactions probables mais manquant d'arguments probant.

1. Les RNases P bactériennes

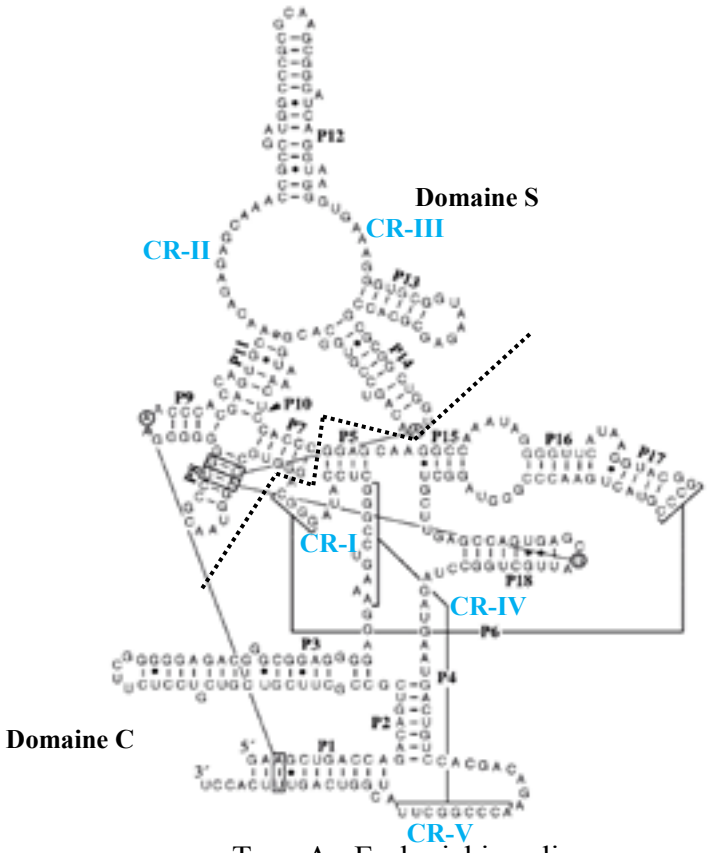
1.1. Structure des RNases P bactériennes

Les RNases P bactériennes possèdent une composante ARN catalytique d'environ 400 nucléotides (nt) codée par le gène *rpnB* et une composante protéique constituée d'une protéine unique (~14 kDa) codée par le gène *rpnA* (Stark *et al.*, 1978; Kole *et al.*, 1979; Kole *et al.*, 1980, Brown *et al*, 1992). L'hydrolyse catalysée d'un pont phosphodiester au niveau d'un substrat ARN a lieu dans un site hautement conservé de la composante ARN. A haute concentration en sel *in vitro*, la composante ARN est capable de couper un substrat sans la composante protéique qui est cependant requise pour l'activité *in vivo* (Kole *et al.*, 1979; Guerrier-Takada *et al.*, 1983).

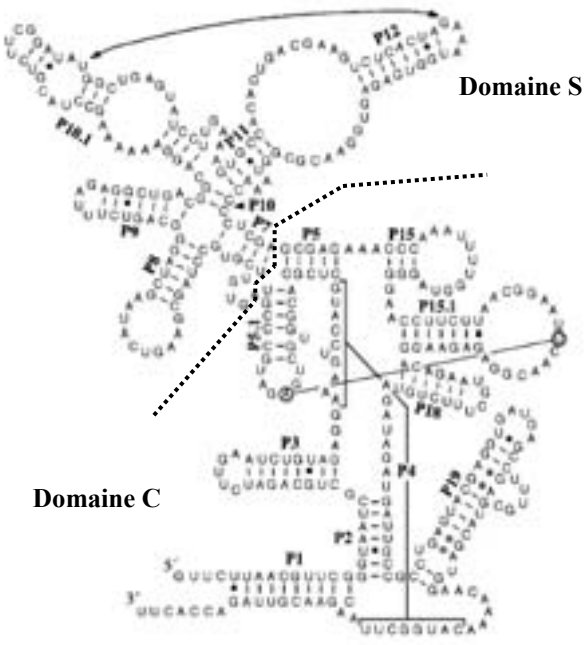
1.1.1. Caractérisation de la composante ARN

1.1.1.1. Les différents types de composantes ARN des RNases P bactériennes

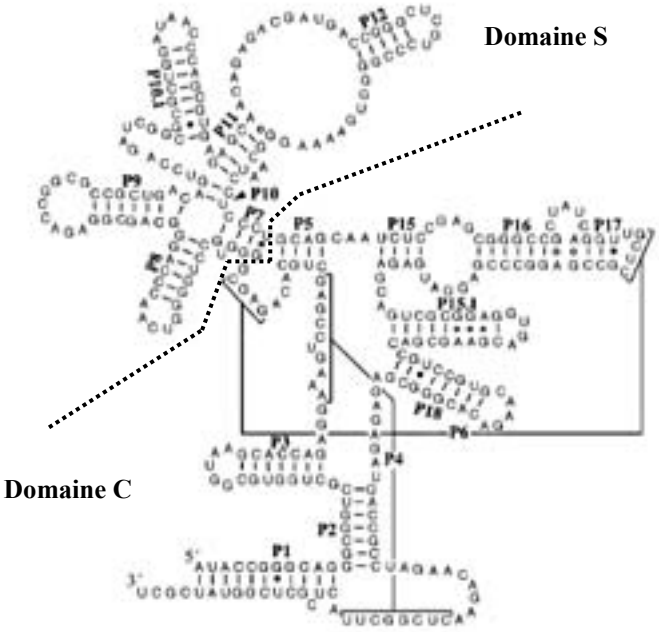
La structure secondaire de la composante ARN a été déterminée par des analyses phylogénétiques comparatives de plusieurs centaines de séquences (James *et al.*, 1988; Harris *et al.*, 2001). Chez les Bactéries, les RNases P bactériennes se divisent principalement en deux types sur la base de la structure secondaire de leur composante ARN: le type A (pour Ancestral), le plus commun avec comme référence-type l'ARN d'*Escherichia coli*, et le type B (pour Bacillus), trouvé uniquement chez les bactéries Gram-positives avec comme référence-type l'ARN de *Bacillus subtilis* (**Figure 2**, type A et B) (Brown *et al.*, 1991; Brown *et al.*, 1993; Brown *et al.*, 1996; Haas *et al.*, 1996). Ces deux types de composantes ARN divergent aussi par leurs voies de biosynthèse. En effet, la composante ARN de type A est synthétisée sous la forme d'un précurseur dont l'extrémité 3' nécessite une maturation par coupure endonucléolytique par la RNase E (Lundberg *et al*, 1995) avant de subir une digestion exoribonucléolytique par la RNase T (Li *et al*, 1998), tandis que les extrémités 3' et 5' de la composante ARN de type B sont maturées par un processus autolytique (Loria *et al*, 2000). Il a cependant été montré que les composantes ARN de ces deux types sont interchangeables *in vitro* (Guerrier-Takada *et al.*, 1983) et *in vivo* (Waugh *et al.*, 1990; Wegscheid *et al.*, 2006) sans affecter l'activité de la RNase P. Cela suggère que les éléments structuraux importants pour l'activité catalytique sont conservés entre ces deux types de RNase P.



Type A : Escherichia coli



Type B : Bacillus subtilis



Type C : Thermomicrobium roseum

Figure 2. Représentation des structures secondaires des différentes classes d'ARN de RNases P bactériennes.

Les hélices sont annotées de P1 à P19 selon la nomenclature de (Haas *et al.*, 1998). Les domaines S et C sont délimités par une ligne pointillée. Les régions conservées CR-I à CR-V sont indiquées sur la structure secondaire de l'ARN d'*E. coli*. Les structures sont issues de *RNase P database* (Brown, 1999).

La plupart des composantes ARN bactériennes sont des variantes des ARN des deux types de RNases P avec des variations dans les éléments périphériques auxiliaires, principalement des différences de longueurs d'hélice ou même d'absence d'hélice dans certains cas (Chen *et al.*, 1997; Brown, 1999). Les structures secondaires se présentent sous la forme d'une suite d'hélices et de boucles numérotées de P1 à P19.

Néanmoins, chez certains Chloroflexi, phototrophes anoxygéniques filamenteux (aussi appelés bactéries vertes non sulfureuses), les composantes ARN partagent une structure intermédiaire entre les types A et B, plus prononcée, à laquelle il est fait référence comme l'ARN de type C (**Figure 2**).

1.1.1.2. L'organisation structurale de la composante ARN des RNases P bactériennes

Les séquences primaires des composantes ARN des RNases P des différentes bactéries sont très variables (Brown, 1999). Néanmoins, cinq régions montrent un haut niveau de conservation et sont appelées CR-I, CR-II, CR-III, CR-IV et CR-V (**Figure 2**, type A) (Chen *et al.*, 1997). Ces régions contiennent des nucléotides qui sont conservés dans les composantes ARN des RNases P de tous les règnes du vivant (Chen *et al.*, 1997; Brown, 1999).

Les informations les plus détaillées sur l'organisation structurale de la composante ARN des RNases bactériennes proviennent des études cristallographiques et particulièrement de la résolution des structures de la composante ARN dans son entier. Les structures résolues sont répertoriées dans le **Tableau 1**. Ces structures cristallines se sont avérées dans leur globalité en accord avec les modèles prédits (Harris *et al.*, 1994; Brown *et al.*, 1996; Chen *et al.*, 1998; Massire *et al.*, 1998). En outre, les composantes ARN s'organisent sous la forme de deux domaines structuraux : le domaine de spécificité (domaine S) et le domaine catalytique (domaine C) (**Figure 2**) (Pan, 1995; Loria *et al.*, 1996; Loria *et al.*, 1999).

Le domaine S est impliqué dans la reconnaissance de la région composée de la branche T, de la boucle T et de la boucle D des ARNt prématures (**Figure 1**). Il permet de lier et de positionner correctement le substrat ce qui confère la spécificité de la RNase P pour les ARNt prématures (Loria *et al.*, 1997; Odell *et al.*, 1998; Mobley *et al.*, 1999; Qin *et al.*, 2001). Des études cristallographiques du domaine S et de la composante ARN entière (Krasilnikov *et al.*, 2003b; Krasilnikov *et al.*, 2004; Kazantsev *et al.*, 2005; Torres-Larios *et al.*, 2005) ont permis de déterminer la structure tertiaire du domaine S dans les ARN des RNases P de types A et B.

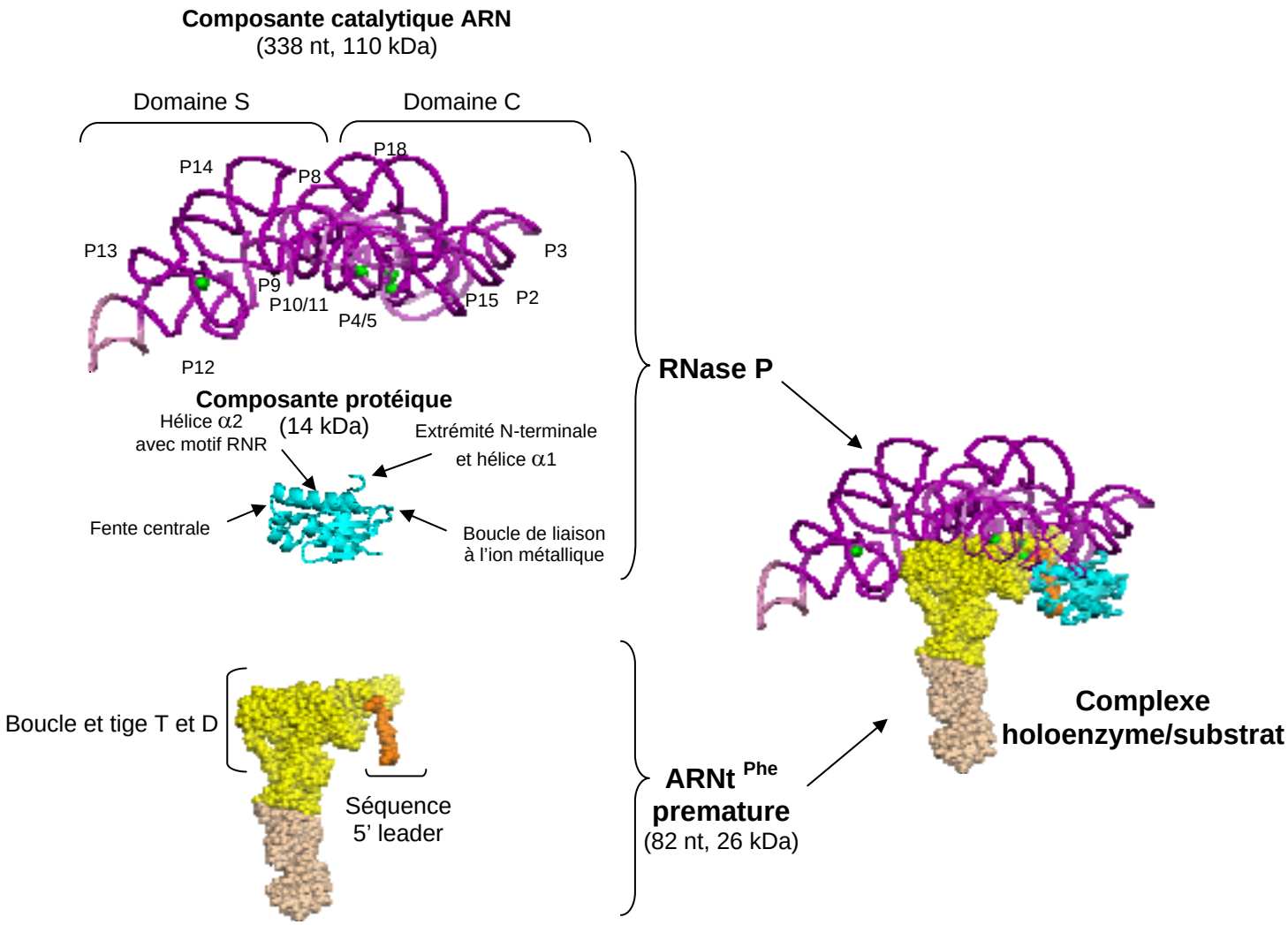


Figure 3. Structure à haute résolution de la RNase P de *T. maritima* en complexe avec un ARNt prémature (Reiter *et al.*, 2010)**.**
Le noyau ARN de type A est représenté en violet avec ses ions magnésiums en vert. La composante protéique est en cyan. L'ARNt est en jaune et la séquence 5' leader en orange. Les nucléotides additionnels ayant permis la cristallisation sont en rose dans la boucle P12 du noyau ARN et couleur grain de blé dans la boucle anticodon de l'ARNt.

Composant	Espèce	Type d'ARN	Résolution	Identifiant PDB	Référence
Domaine S	Thermus thermophilus	A	2.9 Å	1U9S	Krasilnikov *et al.*, 2004
	Bacillus subtilis	B	3.15 Å	1NBS	Krasilnikov *et al.*, 2003b
ARN	Thermotoga maritima	A	3.85 Å	2A2E	Torres-Larios *et al.*, 2005
	Bacillus stearothermophilus	B	3.3 Å	2A64	Kazantsev *et al.*, 2005
Protéine	Thermotoga maritima	A	1.2 Å	1NZ0	Kazantsev *et al.*, 2003
	Bacillus subtilis	B	2.6 Å	1A6F	Stams *et al.*, 1998
	Staphilococcus aureus	B	RMN	1D6T	Spitzfaden *et al.*, 2000
ARN+Protéine +ARNt	Thermotoga maritima	A	3.8-4.2 Å	3OK7 3OK8	Reiter *et al.*, 2010

Tableau 1. Tableau récapitulant les structures résolues des différents composants de la RNase P bactérienne.
RMN : structure résolue par spectrométrie à résonance magnétique nucléaire.

Ce domaine S se présente sous la forme d'une pince dont les branches formées par les hélices P8/P9 et P10/P11 sont reliées par un module compact constitué par la boucle P12. Ces éléments structuraux sont universellement conservés dans toutes les RNases P ribonucléoprotéiques, incluant les RNases P archées et eucaryotes (Chen *et al.*, 1997). Ils contiennent par ailleurs les régions CR-II et CR-III présentes dans la boucle P12. De plus, les nucléotides interagissant avec la région branche T - boucle T des ARNt prématures (LaGrandeur *et al.*, 1994; Loria *et al.*, 1997; Odell *et al.*, 1998) bordent l'ouverture de cette pince de manière similaire pour les deux types d'ARN. La résolution de la structure de la RNase P bactérienne avec son substrat a confirmé l'interaction de ces résidus avec l'ARNt prémature (**Figure 4**) (Reiter *et al.*, 2010). Ces constatations suggèrent que cette structure en forme de pince permet d'identifier le substrat, fonction qui serait conservée dans toutes les RNases P ribonucléoprotéiques des règnes du vivant (Krasilnikov *et al.*, 2003a; Krasilnikov *et al.*, 2004). Pour former cette structure conservée, les éléments formant la pince doivent être correctement positionnés les uns par rapport aux autres. Ce positionnement est mis en œuvre par les éléments auxiliaires du domaine S qui forment un contrefort tridimensionnel fournissant une stabilité structurale à la partie fonctionnellement importante. Les principales différences entre les types A et B de RNase P concernent ces éléments auxiliaires du domaine S. Dans les composantes ARN des RNases P de type A, la rigidité structurale de la région de liaison du substrat est fournie par les hélices P13 et P14 formant des interactions tertiaires avec les hélices P8 et P12 (Krasilnikov *et al.*, 2004). Dans les composantes des RNases P de type B, les hélices P13 et P14 sont absentes, le soutien structural est alors fourni par l'interaction entre l'hélice P10.1 (absente des composantes des RNases P de type A) avec la boucle terminale de l'hélice P12 et les hélices P7/P10 (Krasilnikov *et al.*, 2003b). C'est ainsi que la stabilisation du noyau important du domaine S est réalisée selon deux architectures différentes en fonction du type de RNase P (Krasilnikov *et al.*, 2004; Westhof *et al.*, 2004).

1.1.1.3. Caractérisation fonctionnelle de la composante ARN des RNases P bactériennes

De nombreuses études ont porté sur la caractérisation de l'activité catalytique de la composante ARN des RNases P pour explorer les propriétés de ce ribozyme. Cependant, ce sont encore les études cristallographiques qui ont permis de préciser le mécanisme de coupure endonucléolytique (Reiter *et al.*, 2010). Le site actif de l'holoenzyme se situe dans le domaine C où il est capable de catalyse sans la composante protéique (Guerrier-Takada *et al.*, 1983). Il lie spécifiquement la composante protéique mais n'a par lui-même qu'une faible spécificité et

une faible affinité pour le substrat (Odell *et al.*, 1998; Loria *et al.*, 1999; Mobley *et al.*, 1999). Le domaine C correspond aux hélices et boucles P1 à P6 et P15 à P19. Le site actif se situe dans l'hélice P4 où sont retrouvées les régions conservées CR-I et CR-V (Reiter *et al.*, 2010). La branche accepteur des ARNt prématures s'étend vers le domaine C en croisant les hélices P1/P4/P5 tandis que l'extrémité 3'CCA des ARNt prématures entre dans un conduit formé des hélices P6/P15/P16/P17 et s'apparie avec des nucléotides de la boucle P15 (Kirsebom *et al.*, 1994; Reiter *et al.*, 2010). Par ailleurs, la comparaison du domaine catalytique des deux types d'ARN (Torres-Larios *et al.*, 2006) a montré que la structure tridimensionnelle est très conservée et que comme pour le domaine S, les principales différences concernent l'arrangement des éléments auxiliaires qui semblent servir à stabiliser le domaine C (Kazantsev *et al.*, 2005; Torres-Larios *et al.*, 2005; Torres-Larios *et al.*, 2006).

1.1.2. Caractérisation de la composante protéique

1.1.2.1. Les différentes fonctions de la composante protéique des RNases P bactériennes

La composante protéique des RNases P bactériennes est une protéine basique d'environ 14 kDa (soit 10% de la masse totale de l'holoenzyme). Elle est requise pour l'activité catalytique dans les conditions physiologiques et est essentielle pour la survie des cellules bactériennes (Schedl *et al.*, 1973; Kole *et al.*, 1980; Guerrier-Takada *et al.*, 1983; Kirsebom *et al.*, 1988; Reich *et al.*, 1988; Gossringer *et al.*, 2006). En dépit de sa petite taille, la composante protéique joue de multiples rôles dans l'holoenzyme. Elle semble influencer la fonction de la RNase P à au moins trois niveaux. Tout d'abord, elle améliore la liaison du substrat en réduisant la répulsion électrostatique (Reich *et al.*, 1988) et assouplit l'identification du substrat (Peck-Miller *et al.*, 1991; Liu *et al.*, 1994; Loria *et al.*, 1998; Sun *et al.*, 2010). Elle contribue à stabiliser la conformation active de la composante ARN (Westhof *et al.*, 1996; Gopalan *et al.*, 1997; Kim *et al.*, 1997; Buck *et al.*, 2005a) et allège l'effet de mutations délétères de la composante ARN (Lumelsky *et al.*, 1988). Enfin, elle participe à la catalyse en distinguant le substrat du produit final lors de sa liaison à la séquence 5' leader de l'ARNt prémature (Koutmou *et al.*, 2010; Crary *et al.*, 1998; Kurz *et al.*, 1998; Niranjanakumari *et al.*, 1998; Rueda *et al.*, 2005; Sun *et al.*, 2006; Hsieh *et al.*, 2010) et en augmentant l'affinité de l'ion métallique dans le site actif (Kurz *et al.*, 2002; Sun *et al.*, 2007). La protéine bactérienne est donc la composante de l'holoenzyme qui coordonne le substrat lors des différentes étapes : identification, fixation et coupure du substrat, et enfin libération des produits de réaction.

1.1.2.2. L'organisation structurale de la composante protéique

Les structures des composantes protéiques des RNases P des deux types de composante ARN ont été résolues et sont répertoriées dans le **Tableau 1**. En dépit de séquences primaires divergentes selon les espèces, la structure tridimensionnelle des composantes protéiques est très similaire (Kazantsev *et al.*, 2003). Cette similarité est cohérente avec le fait que les composantes protéiques des deux principaux types de RNases P bactériennes soient interchangeables *in vivo* (Guerrier-Takada *et al.*, 1983; Waugh *et al.*, 1990; Wegscheid *et al.*, 2006).

La composante protéique bactérienne se présente sous la forme d'un sandwich de trois à quatre hélices α et de quatre feuillets β antiparallèles, structure retrouvée dans les protéines liant l'ARN de la machinerie traductionnelle (Stams *et al.*, 1998). Elle s'organise en trois régions : une séquence conservée appelée motif RNR (Arginine-Asparagine-Arginine) (Jovanovic *et al.*, 2002), une fente centrale et une boucle variable contenant un groupe de résidus chargés négativement appelée boucle de liaison à l'ion métallique (**Figure 4**).

1.1.2.3. Interaction entre la composante protéique et la composante ARN

Les interactions entre la composante protéique et la composante ARN ont été très étudiées. Seule une localisation générale a pu être établie (Buck *et al.*, 2005b; Niranjanakumari *et al.*, 2007) jusqu'à la résolution de la structure de l'holoenzyme avec son substrat (Reiter *et al.*, 2010). Cette structure cristalline a mis en évidence que la composante protéique interagit avec le domaine C : son hélice α1 avec l'hélice P15 de l'ARN, son hélice α2 et son extrémité N-terminale contenant le motif RNR avec les régions CR-IV et V et l'interface des hélices P2/P3 (**Figure 4**) mais aussi avec la séquence 5' leader des ARNt au niveau des acides aminés de la fente centrale. Ces constatations sont cohérentes avec les différents fonctions décrites de la composante protéique tant au niveau du site actif qu'au niveau des éléments auxiliaires contribuant à la stabilisation de la composante ARN.

1.2. Substrats de la RNase P bactérienne

La fonction universelle et la mieux caractérisée de la RNase P est la maturation de l'extrémité 5' des ARNt prématures. C'est pourquoi la reconnaissance et le mécanisme de coupure du substrat par la RNase P bactérienne ont été principalement étudiés sur ces ARNt prématures.

1.2.1. Reconnaissance du substrat

Des études par délétion de différents éléments structuraux de l'ARNt prémature ont permis de déterminer la structure minimale nécessaire à l'activité de coupure. Ainsi, la RNase P bactérienne peut couper un substrat constitué uniquement de la branche T, de la boucle T et de la branche accepteur (**Figure 1**). Ces structures forment une hélice reconnue par la RNase P qui la coupe au niveau d'un site spécifique chez la bactérie (Shi *et al.*, 2000; Hansen *et al.*, 2001). De plus, les données biochimiques indiquent qu'*in vitro* à haute force ionique, la composante ARN seule, a des contacts multiples avec l'ARNt prémature : au niveau de la boucle D, la branche T et la boucle T, à l'extrémité 3'-CCA et près du site de coupure (**Figure 1**) (Kirsebom, 2007). La boucle D, la branche T et la boucle T interagissent avec les régions conservées CR-II et CR-III du domaine S de la composante ARN (**Figure 4**) (Reiter *et al.*, 2010) tandis que l'extrémité 3' interagit avec la boucle P15 du domaine C (**Figure 4**) sous la forme d'un motif A-mineur (insertion d'un nucléoside dans le sillon mineur d'un duplex d'ARN) (Kirsebom, 2007). En outre, des études ont montré l'existence d'interactions entre la composante protéique et la séquence 5' leader des ARNt en amont du site de coupure. Les acides aminés de la fente centrale dans la protéine entrent en contact directement avec la séquence leader environ 4 à 7 nt en 5' du site coupure. Cette interaction contribue à normaliser l'identification de substrats variés et à accroître l'efficacité de coupure par la RNase P (Kirsebom *et al.*, 1995; Kirsebom, 2007; Smith *et al.*, 2007).

La normalisation de l'identification du substrat par la composante protéique a permis l'extension de l'usage de la RNase P pour un large éventail de substrats (**Tableau 2**).

Parmi les substrats non-ARNt de la RNase P, un sous-ensemble présente une structure proche de celle des ARNt : les ARNtm prématures d'*E. coli* et de *Cyanelle* (Komine *et al.*, 1994; Gimple *et al.*, 2001), l'ARN viral du virus de la mosaïque jaune du navet (TYMV) (Giegé *et al.*, 1993) et l'ARN ColE1 (Jung *et al.*, 1995). La ressemblance des substrats illustre l'importance de la structure tertiaire du substrat dans l'activité de coupure de la RNase P. L'holoenzyme est aussi impliquée dans la coupure d'autres substrats de structures variées :

l'ARN 4.5S et l'ARN induit du bactériophage Φ80 (Bothwell *et al.*, 1976; Peck-Miller *et al.*, 1991), des ARNm polycistroniques (Alifano *et al.*, 1994), des ARN précurseurs de phages (Forti *et al.*, 1995; Hartmann *et al.*, 1995), des ARN transitoires de type *riboswitch* (Altman *et al.*, 2005), de petits ARN d'au moins 5 nt (Hansen *et al.*, 2001), les ARN OLE (*Ornate Large Extremophilic*) de bactéries extrêmophiles (Ko *et al.*, 2007) et les ARN impliqués dans la réplication de l'ADN mitochondrial (Potuschak *et al.*, 1993).

L'adaptabilité du processus d'identification des substrats a été étudiée *in vitro* au moyen d'une sélection d'ARN pour lesquelles il a été testé leur capacité à être coupé avec ou sans la composante protéique. Il a ainsi été montré qu'en présence uniquement de la composante ARN, les substrats les plus efficacement coupés présentent une structure proche de celle des ARNt. Lorsque la composante ARN est complétée par sa composante protéique, des substrats de structure non ARNt sont alors plus facilement coupés (Liu *et al.*, 1994). Ces résultats montrent l'importance de la composante protéique dans le processus d'identification du substrat dans la RNase P bactérienne. Cela suggère aussi que les RNases P avec une composante protéique plus complexe pourraient avoir significativement plus de substrats que la RNase P bactérienne.

Bactéries		Archées		Eucaryotes			
RNase P	**référence**	**RNase P**	**référence**	**RNase P nucléaire**	**référence**	**RNase MRP**	**référence**
ARNt prémature	Altman, 1971	ARNt prémature	Darr *et al.*, 1990	ARNt prémature	Garber *et al.*, 1979	ARNr prémature (site A3)	Schmitt *et al.*, 1993b
ARNtm précurseurs (ARN 10Sa & Cyanelle)	Komine *et al.*, 1994; Gimple *et al.*, 2001			ARN HRA1 antisens	Yang *et al.*, 2007	Amorce ARN pour la réplication de l'ADN mitochondrial	Chang *et al.*, 1987b
ARN de TYMV	Giege *et al.*, 1993			ARN non codant	Samanta *et al.*, 2006	ARNm CLB2	Gill *et al.*, 2004
ARN ColE1	Jung *et al.*, 1995			Petit ARN nucléolaire à boîtes C/D	Coughlin *et al.*, 2008		
ARN 4.5S prémature	Peck-Miller *et al.*, 1991			ARNr prémature (sites multiples)	Chamberlain *et al.*, 1996		
ARN Φ80 induit	Bothwell *et al.*, 1976			Long ARN non codant	Wilusz *et al.*, 2008		
ARNm polycistronique	Alifano *et al.*, 1994						
ARN précurseurs issus de bactériophages	Forti *et al.*, 1995; Hartmann *et al.*, 1995						
Riboswitch	Altman *et al.*, 2005						
Petit ARN (>5 nt)	Hansen *et al.*, 2001						
ARN OLE	Ko *et al.*, 2007						
Amorce ARN pour la réplication de l'ADN mitochondrial	Potuschak *et al.*, 1993						

Tableau 2. Tableau récapitulatif des substrats pertinents des RNases P et MRP.

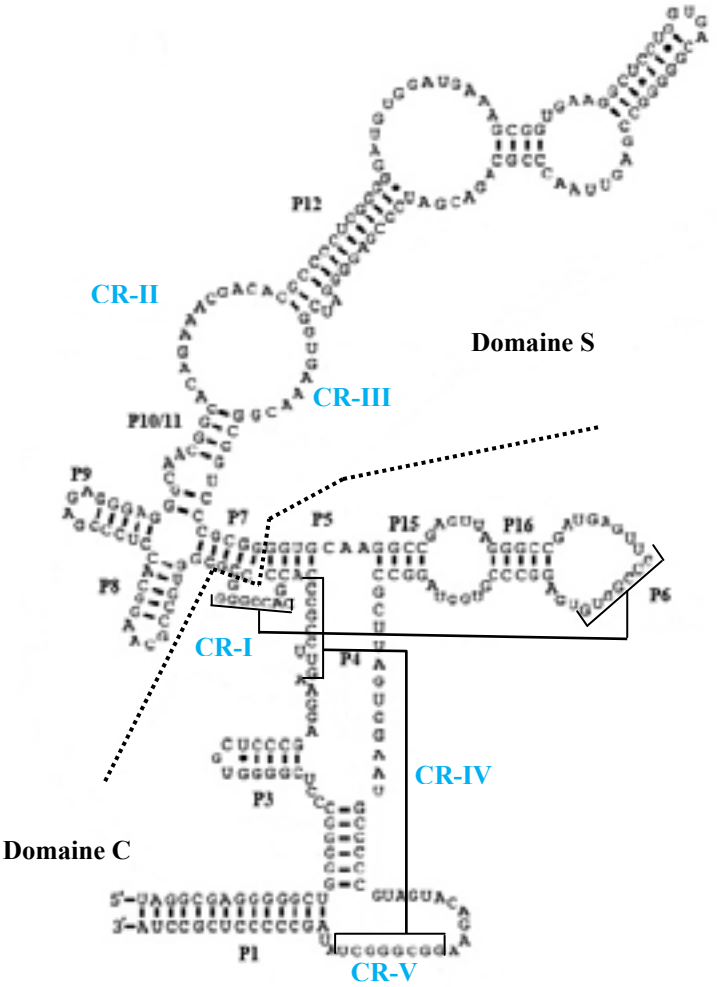
1.2.2. Mécanisme de coupure du substrat

En raison de la nature catalytique de la composante ARN, de nombreuses études ont été réalisées pour élucider son mécanisme catalytique. Ainsi, il a été mis en évidence *in vitro* qu'à haute concentration en sel et en ions métalliques divalents, la composante ARN est capable de catalyser seule la maturation de l'extrémité 5' des ARNt prématures (Guerrier-Takada *et al.*, 1983; Gossringer *et al.*, 2006). Cependant, *in vivo*, la présence de la composante protéique est essentielle à la réalisation de la réaction (Gossringer *et al.*, 2006).

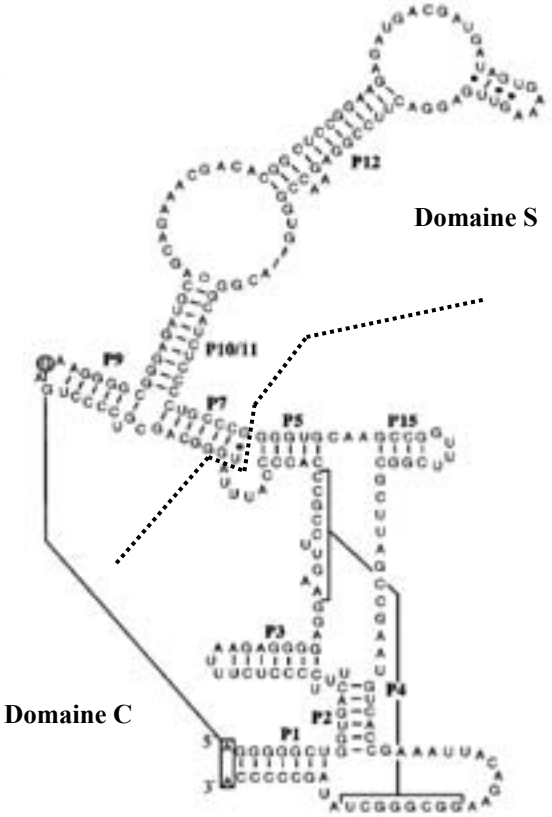
De plus, les études de la cinétique à l'état d'équilibre ont révélées que la coupure de la séquence 5' leader des ARNt prématures est dépendante du pH, de la présence des ions métalliques divalents (Mg^{2+}, Ca^{2+} ou Mn^{2+}), n'accumule pas d'intermédiaires covalents et est activée par de hautes concentrations en sel monovalent (Altman, 1989; Pace *et al.*, 1990; Smith *et al.*, 1992; Warnecke *et al.*, 1999). De plus, ces études ont montré qu'à l'état d'équilibre en concentration saturante en substrat, l'étape limitant la réaction n'est pas la coupure du substrat mais la libération des produits de réaction (Tallsjo *et al.*, 1993).

La RNase P est une endoribonucléase qui catalyse la coupure d'un pont phosphodiester spécifique sur un ARN pour générer des extrémités 5' monophosphate et 3' hydroxyle (Doudna *et al.*, 2005). Bien que le mécanisme chimique reste mal défini, la résolution de la structure de la RNase P de *T. maritima* avec son substrat (Reiter *et al.*, 2010) a permis d'identifier les principaux composants du site actif dans l'hélice P4 au niveau d'un uracile universellement conservé. Par ailleurs, la structure met en évidence qu'au moins deux ions métalliques divalents jouent un rôle direct. Reiter et ses collègues ont proposé qu'un des deux ions oriente directement les atomes d'oxygène du phosphate au niveau du site de coupure et permette à un ion hydroxyle de réaliser une substitution nucléophile de type SN_2 (Cassano *et al.*, 2004). Le second ion métallique stabiliserait alors l'état de transition et faciliterait le transfert du proton de l'atome d'oxygène en 3' au niveau du site de coupure au cours de la libération du substrat. La RNase P utiliserait donc un mécanisme catalytique basé sur l'action d'ions métalliques, similaires aux mécanismes d'autres ribozymes comme les introns auto-épisseurs des Groupes I et II (Stahley *et al.*, 2005; Toor *et al.*, 2008).

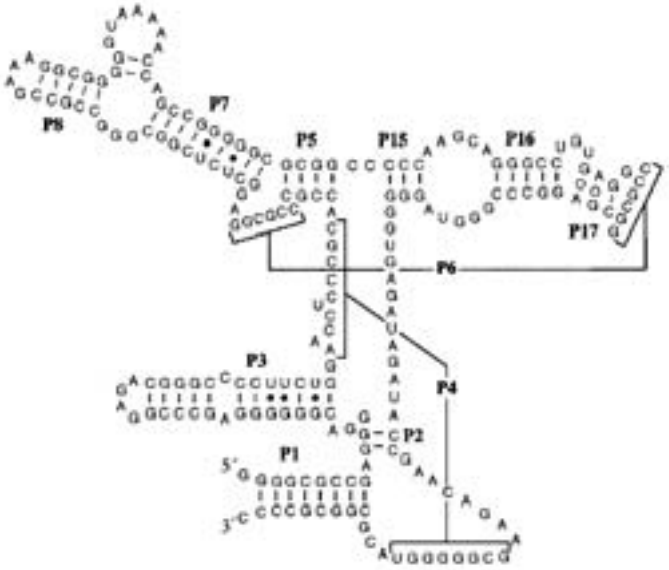
Par ailleurs, ces ions métalliques divalents sont aussi requis pour le repliement correct de la composante ARN de la RNase P (Warnecke *et al.*, 1999; Baird *et al.*, 2007; Qu *et al.*, 2008; Kazantsev *et al.*, 2009; Baird *et al.*, 2010; Hsieh *et al.*, 2010). Cependant, bien qu'ils accroissent l'affinité de la RNase P pour son substrat, ils ne semblent pas être absolument requis pour la fixation du substrat sur l'holoenzyme (Smith *et al.*, 1992; Beebe *et al.*, 1996).



Type A : Pyrococcus horikoshii



Type M : Methanocaldococcus jannaschii



Type P : Pyrobaculum aerophilum

Figure 4. Représentation des structures secondaires des différentes classes d'ARN de RNase P archée.

Les hélices sont annotées de P1 à P17 selon la nomenclature de (Haas *et al.*, 1998). Les domaines S et C sont délimités par une ligne pointillée. Les régions conservées CR-I à CR-V sont indiquées sur la structure secondaire de l'ARN de *P. horikoshii*. Les structures sont issues de *RNase P database* (Brown, 1999).

2. Les RNases P archées

Les RNases P archées sont apparentées tant d'un point de vue phylogénétique que structural aux RNases P bactériennes et eucaryotes (Brown, 1999; Hall *et al.*, 2001). Elles sont constituées d'une composante ARN et d'une composante protéique de quatre à cinq sous-unités (Hall *et al.*, 2002; Hartmann *et al.*, 2003; Kouzuma *et al.*, 2003; Fukuhara *et al.*, 2006; Pulukkunat *et al.*, 2008). Avec cet accroissement de la complexité de leur composante protéique, les RNases P archées sont un excellent substitut par rapport à leurs homologues eucaryotes difficilement étudiables expérimentalement (Tsai *et al.*, 2006). En dépit des importantes différences de la composante ARN, le potentiel catalytique, l'organisation de l'holoenzyme en deux composantes dans les tous les règnes du vivant et la constante catalytique des holoenzymes sont comparables (Andrews *et al.*, 2001; Xiao *et al.*, 2005; Sun *et al.*, 2006). La haute conservation des séquences entre les sous-unités protéiques archées et eucaryotes (Hall *et al.*, 2002) permet d'étendre la compréhension des fonctions des différentes sous-unités protéiques archées aux RNases P eucaryotes dont la reconstitution *in vitro* est difficile. De plus, avec seulement quatre à cinq protéines de petit poids moléculaire, les RNases P archées sont des complexes ribonucléoprotéiques plus simples qui pourraient correspondre au noyau ribonucléoprotéique minimale nécessaire à la catalyse dans le complexe eucaryote. Enfin, les sous-unités protéiques des RNases P archées sont fortement basiques et thermostables ce qui facilitent leur purification, les études structurales et la reconstitution du complexe *in vitro*. De nombreuses études se sont donc focalisées sur la composante protéique des RNases P tandis que la composante ARN, structurellement très proche de celle des RNases P bactériennes est peu caractérisée.

2.1. Caractérisation de la composante ARN

Comme chez les bactéries, la composante ARN archée montre une forte conservation de sa structure secondaire et un noyau catalytique conservé avec des éléments périphériques variables (Brown, 1999). Sur la base de l'étude des structures secondaires de leur ARN, les RNases P archées ont été principalement divisées en deux types : le type A (Ancestral) qui rassemble la majorité des RNases P archées et le type M représenté par les RNases P des *Methanococcales* et des *Archeoglobales* (**Figure 4**, type A et M) (Harris *et al.*, 2001).

Les composantes ARN des RNases P de type A ont une structure très similaire à celle des RNases P de type A bactérienne. Les principales différences entre les composantes archées et bactériennes sont (1) l'absence des hélices P13, P14 dans le domaine S, essentielles pour le pontage avec le domaine C chez les bactéries, (2) la présence d'éléments structuraux

complexes dans l'hélice P12, caractéristiques très conservés chez les archées et les eucaryotes (Brown, 1999; Harris *et al.*, 2001) et (3) l'absence de l'hélice P18 dans le domaine C. Les composantes ARN des RNases P de type M, ont de plus perdu les hélices P6 et P16/17 du domaine C, hélices impliquées dans la stabilisation de la structure tertiaire des ARN bactériens ainsi que l'hélice P8 du domaine S, requise pour l'identification de la boucle T des ARNt prématures. En outre, ils possèdent une hélice P15 atypique sans la boucle qui ancre la séquence 3'-CCA des ARNt prématures chez les bactéries (Kirsebom *et al.*, 1994; Oh *et al.*, 1994). La façon dont l'holoenzyme de type M compense pour l'absence de ces éléments fonctionnels essentiels n'est pas connue.

Ces différences structurales sont corrélées à des différences d'activité catalytique. Ainsi, les ARN des RNases P de type A peuvent être catalytiquement actifs en l'absence de la composante protéique *in vitro*. Cependant, cette activité fortement réduite par rapport à celle de la composante bactérienne requière une force ionique extrêmement élevée (avec une activité maximale dans 2 à 3 M d'acétate d'ammonium avec 200 à 500 mM de chlorure de magnésium) (Pannucci *et al.*, 1999; Kouzuma *et al.*, 2003; Tsai *et al.*, 2006). Cependant, le remplacement des domaines S de l'ARN de type A par le domaine S bactérien dans les composantes ARN archées conduit à une augmentation conséquente de l'activité catalytique pour les ARN de type A (Pannucci *et al.*, 1999). La composante ARN des RNases P de type M est, quant elle, incapable d'activité catalytique *in vitro* en l'absence de sa composante protéique (Pannucci *et al.*, 1999). Le remplacement du domaine S de l'ARN de type M par son homologue bactérien ou une fusion de la composante ARN archée avec son substrat conduit à une restauration de cette activité (Pulukkunat *et al.*, 2008; Li *et al.*, 2009). Il semble donc que, bien que les composantes ARN archées possèdent tous les éléments structuraux nécessaires à une activité catalytique, elles dépendent de leur composante protéique pour leur stabilité structurale.

Il existe par ailleurs des espèces d'archées où les composantes ARN sont encore plus compactes et qui définissent une troisième classe de RNase P dit de type P (**Figure 4**, type P). C'est le cas des espèces du genre *Pyrobaculum* où les ARN se composent essentiellement du domaine C, les hélices P9, P11 et P12 incluant les régions conservées CR-II et CR-III étant absentes, la boucle P15 et les hélices P6/P16/P17 stabilisant le noyau catalytique (Lai *et al.*, 2010). Comme cette découverte est récente, peu de données sont disponibles quant aux propriétés biochimiques de l'holoenzyme et à sa composition exacte.

De nombreuses questions liées à l'organisation structurale de ces composantes ARN archée et à leur activité catalytique restent ouvertes. Aucune structure tridimensionnelle à haute résolution de ces composantes ARN n'a pour le moment été résolue.

2.2. Caractérisation de la composante protéique archée

2.2.1. Caractérisation fonctionnelle des sous-unités de la composante protéique archée

Comme dans le cas de la RNase P bactérienne, la composante protéique des RNases P archées est nécessaire à l'activité de l'holoenzyme *in vivo* et à faible force ionique *in vitro* (Tsai *et al.*, 2006). L'analyse initiale des composants protéiques de la RNase P archée a permis d'identifier quatre protéines aPop4, aPop5, aRpp1 et aRpr2 par des outils bioinformatiques grâce à leur homologies avec des sous-unités protéiques des RNases P eucaryotes (**Tableau 4**) (Hall *et al.*, 2002). Leur implication dans l'activité de la RNase P a été confirmée par des tests d'activité de coupure des ARNt prématures après co-immunoprécipitation des protéines (Hall *et al.*, 2002). Plusieurs études ont rapporté la reconstitution *in vitro* d'une RNase P archée de type A hautement active à partir de la composante ARN et des quatre protéines initialement identifiées (Boomershine *et al.*, 2003; Kouzuma *et al.*, 2003; Terada *et al.*, 2006; Tsai *et al.*, 2006). L'analyse de l'influence des sous-unités protéiques sur l'activité de la RNase P partiellement assemblée montre que les quatre protéines agissent par paire : aPop4-aRpr2 et aPop5-aRpp1 (Tsai *et al.*, 2006), ce qui est cohérent avec les résultats des expériences de double-hybride réalisées sur les protéines de *Pyrococcus horikoshii* (Kifusa *et al.*, 2005).

Les deux sous-complexes agissent de façon collaborative : la paire aPop5-aRpp1 interagit avec le domaine C de la composante ARN et le stabilise (Tsai *et al.*, 2006; Pulukkunat *et al.*, 2008; Honda *et al.*, 2010) tandis que la paire aPop4-aRpr2 est impliquée dans la stabilisation du domaine S (Xu *et al.*, 2009; Honda *et al.*, 2010). Les résultats des études de localisation des sites de liaison des protéines sur la composante ARN sont cohérents avec le rôle des deux sous-complexes (Fukuhara *et al.*, 2006; Xu *et al.*, 2009). Les études cinétiques réalisées à partir de complexes RNases P partiellement reconstitués *in vitro* (Tsai *et al.*, 2006) montrent que les deux paires de protéines contribuent à l'activité catalytique différemment. Seule la paire aPop5-aRpp1 augmente la constante catalytique (k_{cat}) ce qui suggère que cette paire de protéines joue un rôle significatif dans la coupure du substrat et/ou la libération des produits de réaction. La paire aPop4-aRpr2, quant à elle, contribue à augmenter l'affinité de la RNase

P pour ses substrats (Chen *et al.*, 2010). Indépendamment l'une de l'autre, ces deux sous-complexes réduisent la constante de Michaelis (Km) de l'enzyme d'une manière similaire tandis que leur présence conjointe diminue la Km de façon synergique. De plus, chaque sous-complexe peut diminuer les concentrations optimales de substrats, d'ions magnésium et d'ions ammonium requis par la composante ARN pour la catalyse de la coupure des ARNt prématures (Tsai *et al.*, 2006). Par ailleurs, une étude a montré que ces protéines sont capables de compenser les défauts structuraux de leur composante ARN et *vice-versa*, mettant en évidence la nature coopérative des composantes protéiques et ARN pour conduire la RNase P archée vers sa conformation fonctionnelle (Chen *et al.*, 2011).

En outre, des études ont mis en évidence que certaines RNases P archées pourraient posséder une 5ème sous-unité protéique : la protéine ribosomale L7Ae, apparentée à la protéine eucaryote Pop3 (Fukuhara *et al.*, 2006). L'ajout de L7Ae à la RNase P de *Pyroccocus horikoshii* reconstituée *in vitro* (un ARN avec les quatre protéines individualisées ci-dessus) élève la température optimale d'activité de 55°C à 70°C et multiplie par cinq la vitesse maximale (Vmax) de la réaction enzymatique (Fukuhara *et al.*, 2006; Terada *et al.*, 2006). Une étude récente a mis en évidence par analyse du dichroïsme circulaire et du spectre d'absorbance aux UV de la composante ARN de *P. horikoshii* que la fixation des quatre sous-unités aPop5, aRpp1, aPop4 et aRpr2 déstabilise la composante ARN tandis que la fixation de L7Ae favorise l'empilement des bases (Kosaka *et al.*, 2010). Ces données suggèrent un mécanisme d'activation de la composante ARN où la fixation de aPop5-aRpp1 sur le domaine C et de aPop4-aRpr2 sur le domaine S déstabilise l'empilement des bases de l'ARN pour le replier dans une conformation catalytiquement active instable tandis que la fixation de L7Ae sur les hélices terminales des domaines C et S stabilise cette conformation active en favorisant le ré-empilement des bases (Fukuhara *et al.*, 2006; Kawano *et al.*, 2006; Honda *et al.*, 2008; Kosaka *et al.*, 2010).

Enfin, une autre protéine, la protéine Alba apparentée aux protéines eucaryotes Pop6 et Pop7, pourrait participer à la composante protéique archée (Aravind *et al.*, 2003). Cependant, la présence de cette protéine n'affecte ni le taux de coupure, ni la température optimale de la réaction de catalyse *in vitro* de la RNase P de *P. horikoshii*. (Hada *et al.*, 2008). De plus, elle n'est pas co-purifiée avec la RNase P de *Methanothermobacterium thermoautotrophicus* et l'application d'antisérums polyclonaux dirigés contre elle ne permet pas d'immunoprécipiter une fraction contenant cette protéine et capable de réaliser l'activité de la RNase P sur les ARNt prématures chez *M. thermoautotrophicus* (Ellis *et al.*, 2007).

2.2.2. Caractérisation structurale de la composante protéique archée

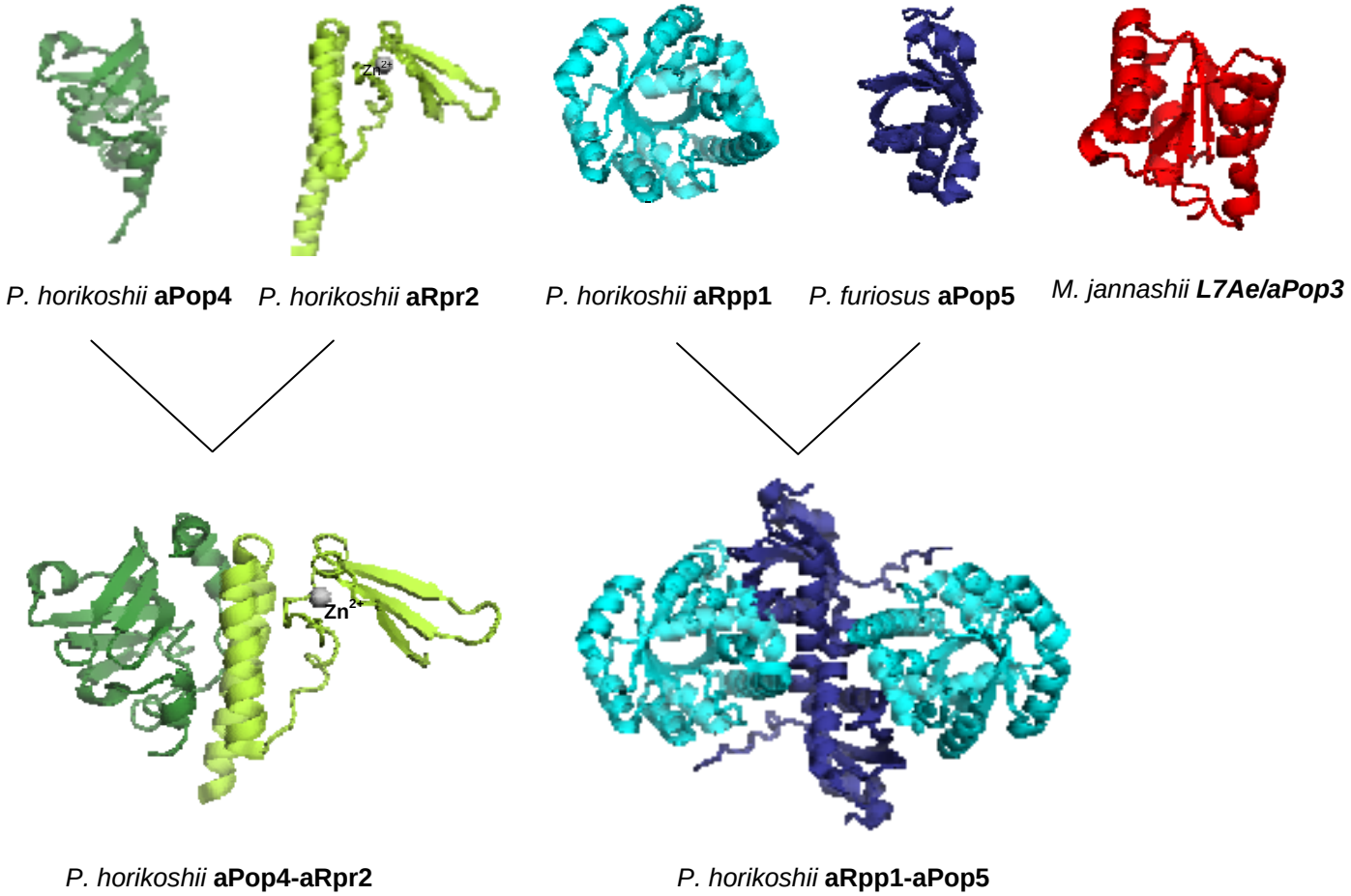
En raison de leur haute solubilité, de leur thermo-stabilité et de la structure simplifiée de leur domaine, la structure des cinq sous-unités protéiques aPop4, aRpr2, aRpp1, aPop5 et L7Ae a été résolue par spectroscopie par résonance magnétique nucléaire (RMN) et/ou par cristallographie aux rayons X sur une période de quelques années. La structure tridimensionnelle des différentes protéines archées (**Figure 5**) est distincte de celle de la protéine de la composante protéique bactérienne (**Figure 4**). Le **Tableau 3** récapitule les différentes structures disponibles pour les différentes sous-unités protéiques.

La protéine aPop4 se présente sous la forme de six feuillets β antiparallèles torsadés entourant un noyau hydrophobe conservé (Boomershine *et al.*, 2003; Sidote *et al.*, 2003; Numata *et al.*, 2004; Sidote *et al.*, 2004).

La protéine aRpr2 est constituée d'un doigt de zinc inclus dans une structure en forme de L. Un bras comprend deux hélices α dans le domaine N-terminal tandis que l'autre bras comprend trois feuillets β antiparallèles dans le domaine C-terminal. Une boucle connecte les deux bras. Ces deux bras fournissent la structure de base pour le motif ordonné des quatre cystéines que coordonne un unique ion zinc. La mutation d'un des résidus cystéines conduit à la perte de la structure et de son activité (Kakuta *et al.*, 2005; Amero *et al.*, 2008).

La protéine aRpp1 adopte une structure en tonneau α/β similaire à celle de la triose phosphate isomérase (TIM) (Takagi *et al.*, 2004). Elle se compose de dix hélices α et de sept feuillets β qui constituent une forme ellipsoïde avec une hélice C-terminale servant de couvercle au tonneau.

La protéine aPop5 présente une structure en sandwich avec quatre feuillets β antiparallèles centraux entourés de quatre hélices α (Kawano *et al.*, 2006; Wilson *et al.*, 2006). Une face des feuillets β se tasse contre les hélices α1-α3 pour former un noyau hydrophobe. L'autre face des feuillets β se tasse contre l'hélice α4 et présente quelques résidus apolaires qui pourraient s'engager dans les interactions avec la composante ARN et les ARNt prématures. Bien que cette protéine ne partage aucune homologie de séquence avec la protéine de la composante protéique bactérienne (Stams *et al.*, 1998), la structure tridimensionnelle des deux protéines est similaire malgré des connectivités différentes de la structure secondaire (Wilson *et al.*, 2006). En outre, en association avec aRpp1, aPop5 semble avoir des fonctions similaires à la protéine bactérienne (Tsai *et al.*, 2006; Pulukkunat *et al.*, 2008; Honda *et al.*, 2010).



Adaptée de Mondragon, 2010

Figure 5. Structures à haute résolution des composants protéiques de la RNase P archée.
Le panel du haut montre la structure des protéines individuelles tandis que le panel du bas montre la structure de l'hétérodimère aRpr2-aPop4 [2ZAE, (Honda *et al.*, 2008)] à gauche et de l'hétérotétramère aPop5-aRpp1 [2CZV (Kawano *et al.*, 2006)] à droite.
aPop4 [1V76, (Numata *et al.*, 2004)] est représentée en vert foncé, aRpr2 [1X0T, (Kakuta *et al.*, 2005)] en vert clair, aRpp1 [1V77, (Takagi *et al.*, 2004)] en bleu clair, aPop5 [2AV5, (Wilson *et al.*, 2006)] en bleu foncé et aPop3 [2CZW, (Fukuhara *et al.*, 2006)] en rouge.

Enfin, la protéine L7Ae/aPop3 adopte une structure globulaire en sandwich avec quatre feuillets β antiparallèles centraux entourés de cinq hélices α (Hamma *et al.*, 2004; Fukuhara *et al.*, 2006), proche de celle de aPop5. Les feuillets β sont pris en sandwich d'un côté par les hélices α1, α4 et α5 et de l'autre par les hélices α2 et α3.

En raison de l'organisation en paire de quatre des sous-unités, la structure des sous-complexes aPop4-aRpr2 et aRpp1-aPop5 a aussi été étudiée et résolue (**Tableau 3**).

La résolution de la structure du sous-complexe aPop4-aRpr2 a mis en évidence que les hélices α de aRpr2 interagissent avec le segment non structuré N-terminal, le feuillet β2 et l'hélice α3 C-terminale de aPop4 (Honda *et al.*, 2008). L'interface protéine-protéine est majoritairement dominée par plusieurs liaisons hydrogène et des ponts ioniques. La surface du sous-complexe aPop4-aRpr2 présente des aires avec des résidus basiques phylogénétiquement conservés qui pourraient constituer une surface de liaison à l'ARN (Xu *et al.*, 2009). Cependant, le détail des interactions ARN-protéines n'est pas connu actuellement.

Lors de la résolution de la structure du sous-complexe aRpp1-aPop5, il a été mis en évidence une organisation de ce sous-complexe en un hétérotétramère où un homodimère de protéines aPop5 est entouré de part et d'autre de deux copies de la protéine aRpp1 (Kawano *et al.*, 2006). En plus de l'homodimérisation des protéines aPop5 par un réseau de liaisons hydrogène, chaque molécule aPop5 interagit simultanément avec deux molécules aRpp1 par des liaisons ioniques et des interactions hydrophobes. Les hélices α6-α8 et les feuillets β6 et β7 des protéines aRpp1 forment un ensemble hydrophobe qui participe à l'interaction avec aPop5. Des expériences de mutagenèse indiquent que la formation d'un hétérodimère aRpp1-aPop5 peut être découplée de la formation de l'hétérotétramère par l'abolition de la capacité de la protéine aPop5 à homodimériser (Kawano *et al.*, 2006). Cependant, si les données de chromatographie d'exclusion indique que l'hétérotétramère se forme en solution, la pertinence fonctionnelle de cette oligomérisation *in vivo* n'est pas certaine. La distribution du potentiel électrostatique de surface de l'hétérotétramère aRpp1-aPop5 révèle des aires chargées positivement probablement impliquées dans l'interaction avec la composante ARN et/ou des substrats. Cependant, les détails de ces interactions ne sont, eux non plus, pas connus actuellement.

Une structure à haute résolution d'une RNase P archée est nécessaire pour fournir des informations détaillées sur les interactions ARN-protéine et protéine-protéine dans cette

holoenzyme. Une telle structure permettrait de fournir des indices importants pour la compréhension des holoenzymes eucaryotes plus complexes. L'homologie de séquence entre les sous-unités protéiques archées et eucaryotes pourrait permettre d'extrapoler la structure tridimensionnelle de certaines sous-unités protéiques eucaryotes grâce aux structures à haute résolution de ces protéines archées.

Composant protéique	Espèce	Type d'ARN	Résolution	Identifiant PDB	Référence
aPop4	Archaeoglobus fulgibus	A	RMN	1PC0	Sidote *et al.*, 2003
	Methanothermobacterium thermoautotrophicus	M	RMN	1OQK	Boomershine *et al.*, 2003
	Pyrococcus horikoshii OT3	A	2.0 Å	1V76	Numata *et al.*, 2004
	Archaeoglobus fulgibus	A	1.7 Å	1TSF, 1TS9	Sidote *et al.*, 2004
aRpr2	Pyrococcus horikoshii OT3	A	1.6 Å	1X0T	Kakuta *et al.*, 2005
	Pyrococcus furiosus	A	RMN	2K3R	Amero *et al.*, 2008
aRpp1	Pyrococcus horikoshii OT3	A	1.8 Å	1V77	Takagi *et al.*, 2004
aPop5	Pyrococcus furiosus	A	2.5 Å	2AV5	Wilson *et al.*, 2006
aPop3/L7Ae	Methanococcus jannashii	M	1.8 Å	1SDS	Hamma *et al.*, 2004
	Pyrococcus horikoshii OT3	A	1.9 Å	2CZW	Fukuhara *et al.*, 2006
aPop4+aRpr2	Pyrococcus horikoshii OT3	A	2.2 Å	2ZAE	Honda *et al.*, 2008
	Pyrococcus furiosus	A	RMN	2K17	Xu *et al.*, 2009
aPop5 + aRpp1	Pyrococcus horikoshii OT3	A	2.0 Å	2CZV	Kawano *et al.*, 2006

Tableau 3. Tableau récapitulant les structures résolues des différents composants de la RNase P archée.
RMN : structure résolue par spectrométrie à résonance magnétique nucléaire.

3. Les RNases P eucaryotes

A l'exemple d'autres processus de biosynthèse chez les eucaryotes, l'activité RNase P eucaryote a été divisée et localisée spécifiquement selon les compartiments subcellulaires pour permettre l'accomplissement de fonctions supplémentaires et une régulation plus complexe. Ce partitionnement a été accompagné par l'apparition de plusieurs enzymes distinctes constituées d'un nombre variable de protéines et dont la composante ARN peut être absente. Chez les eucaryotes, l'activité RNase P nucléaire est portée soit par un complexe ribonucléoprotéique (RNP) comme chez la levure et les animaux, soit par des enzymes uniquement protéiques notamment chez les plantes. Dans les organelles, l'activité RNase P est selon les organismes portée par une ribonucléoprotéine chez la levure, un complexe protéique chez l'homme ou une protéine seule chez les plantes.

Les RNases P bactériennes et archées, de nature ribonucléoprotéique, ont aussi évolué vers une holoenzyme paralogue, la RNase MRP. Bien que la RNase MRP ne soit pas une RNase P au sens strict, de par son incapacité à couper les ARNt prématures, elle est cependant clairement dérivée de la RNase P ribonucléoprotéique dont elle a gardé des éléments significatifs de la composante ARN et avec qui elle partage une grande partie des sous-unités de la composante protéique.

3.1. Les RNases P ribonucléoprotéiques

3.1.1. Les RNases P ribonucléoprotéiques nucléaires

Les RNases P nucléaires, holoenzymes sous la forme d'une composante ARN et d'une composante protéique sont retrouvées chez presque tous les eucaryotes analysés à l'exception principalement des plantes (Hartmann *et al.*, 2003; Piccinelli *et al.*, 2005; Rosenblad *et al.*, 2006). A cause de la nature complexe des eucaryotes et de la faible abondance de complexes macromoléculaires (Srisawat *et al.*, 2001), les RNases P sont difficiles à étudier. C'est pourquoi de nombreuses questions relatives à leur fonction, leur assemblage et leur organisation structurale sont encore sans réponse. Les deux exemples de RNases P nucléaires les mieux étudiées sont celles retrouvées chez *S. cerevisiae* et de *H. sapiens*.

3.1.1.1. Structure des RNases P nucléaires

3.1.1.1.1. Composition des RNases P nucléaires

La RNase P nucléaire de *S. cerevisiae* se compose d'une composante ARN, appelée RPR1 (*RNase P RNA 1*) de 369 nt et d'une composante protéique de neuf protéines (Pop1

pour *Processing for precursor RNAs 1*, Pop3-Pop8, Rpp1 pour *RNase P protein 1* et Rpr2 pour *RNase P ribonucleoprotein 2*) (**Tableau 4**) (Chamberlain *et al.*, 1998). Ces sous-unités protéiques et ribonucléiques sont essentielles pour l'activité RNase P et la survie cellulaire (Lee *et al.*, 1991b; Lygerou *et al.*, 1994; Chu *et al.*, 1997; Dichtl *et al.*, 1997; Stolc *et al.*, 1997; Chamberlain *et al.*, 1998). La plupart de ces protéines est de taille modérée (15-33 kDa) à l'exception de la protéine Pop1 comparativement plus grande (101 kDa). En considérant une composition équimolaire des sous-unités dans ce complexe, les protéines représenteraient plus de la moitié de la particule ribonucléoprotéique qui serait alors d'une masse approximative de 410 kDa. L'activité de la RNase P nucléaire de *S. cerevisiae* a été localisée principalement dans le nucléole où sont également retrouvés les gènes des ARNt ainsi que les ARNt prématures (Bertrand *et al.*, 1998; Thompson *et al.*, 2003; Wang *et al.*, 2005; Haeusler *et al.*, 2008). La RNase P nucléaire semble localisée directement au niveau du site de transcription des ARNt prématures par l'ARN polymérase III suggérant une maturation co-transcriptionnelle des ARNt.

La RNase P nucléaire humaine a également été très étudiée. Elle possède une composante ARN de 339 nt, appelée H1 (Bartkiewicz *et al.*, 1989) et une composante protéique constituée de dix protéines (Rpp14, Rpp20, Rpp21, Rpp25, Rpp29, Rpp30,Rpp38, Rpp40, hPop1 et hPop5) (Eder *et al.*, 1997; Jarrous *et al.*, 2001a) dont neuf sont homologues aux protéines de la RNase P de levure (**Tableau 4**) (Lygerou *et al.*, 1996b; Jarrous *et al.*, 1999a; van Eenennaam *et al.*, 1999; van Eenennaam *et al.*, 2001; Jarrous, 2002). Dans les cellules humaines, l'ARN H1 semble seulement entrer dans le nucléole transitoirement avant de diffuser dans le nucléoplasme. Il est d'ailleurs aussi détecté dans le compartiment périnucléaire et dans le cytoplasme (Wolin *et al.*, 1999). De plus, si certaines protéines sont spécifiquement localisées dans le nucléole, d'autres sont aussi retrouvées dans les particules cytoplasmiques de Cajal (Lygerou *et al.*, 1996b; Jarrous *et al.*, 1999b; Savino *et al.*, 1999; van Eenennaam *et al.*, 2001). La RNase P semble par ailleurs être impliquée directement dans la machinerie transcriptionnelle aussi bien auprès de l'ARN polymérase III (Reiner *et al.*, 2006) que de l'ARN polymérase I (Reiner *et al.*, 2008). Actuellement, la relation entre la localisation de la RNase P et son implication dans la transcription des ARNr et des ARNt est mal connue, tandis que l'extrapolation de ce modèle à d'autres organismes comme la levure reste sujet à discussion.

Quatre des sous-unités de la composante protéique sont homologues aux protéines archées aPop4, aPop5, aRpp1 et aRpr2 (**Tableau 4**, protéines surlignées en orange) ce qui montre que les RNases P eucaryotes et archées partagent un ancêtre commun (Hall *et al.*,

2002; Hartmann *et al.*, 2003). Actuellement, il n'est pas bien compris pourquoi les RNases P eucaryotes ont acquis autant de sous-unités protéiques additionnelles, ni quelles sont leurs fonctions spécifiques dans l'activité de la RNase P. Elles pourraient potentiellement jouer un rôle en définissant la localisation subcellulaire, en stabilisant la structure de la composante ARN, en interagissant avec le substrat ou en facilitant un changement conformationnel pour optimiser la catalyse.

S. cerevisiae						*H. sapiens*						Archée *P. horikoshii*			
Gène	Référence	Poids moléculaire (kDa)	pI	RNase P	RNase MRP	Gène	Référence	Poids moléculaire (kDa)	pI	RNase P	RNase MRP	Gène	Référence	Poids moléculaire (kDa)	pI
Composante ARN						**Composante ARN**						**Composante ARN**			
RPR1	Lee *et al.*, 1991b	120	-	+	-	H1	Bartkiewicz *et al.*, 1989	112	-	+	-	Ph rnpB	Kouzuma *et al.*, 2003	107	-
NME1	Schmitt *et al.*, 1992; Chu *et al.*, 1994	112	-	-	+	7-2	Chang *et al.*, 1989	88	-	-	+				
Composante protéique						**Composante protéique**						**Composante protéique**			
Pop1	Lygerou *et al.*, 1994	100.5	9.8	+	+	hPop1	Lygerou *et al.*, 1996b	115	9.9	+	+				
Pop3	Dichtl *et al.*, 1997	22.6	9.6	+	+	hPop3 (Rpp38)	Eder *et al.*, 1997	32	9.3	+	+	aPop3 /L7Ae (Ph1496)	Fukuhara *et al.*, 2006	13.5	5.35
Pop4	Chu *et al.*, 1997	32.9	9.3	+	+	hPop4 (Rpp29)	Jarrous *et al.*, 1999a; van Eenennaam *et al.*, 1999	25	10.1	+	+	aPop4 (Ph1771)	Hall *et al.*, 2002	15.1	11.2
Pop5	Chamberlain *et al.*, 1998	19.6	7.8	+	+	hPop5	van Eenennaam *et al.*, 2001	19	7.9	+	+	aPop5 (Ph1481)	Hall *et al.*, 2002	14.0	10.0
Pop6	Chamberlain *et al.*, 1998	18.2	9.3	+	+	hPop6 (Rpp25)	Guerrier-Takada *et al.*, 2002	21	9.7	+	+				
Pop7	Chamberlain *et al.*, 1998; Stolc *et al.*, 1998	15.8	9.3	+	+	hPop7 (Rpp20)	Jarrous *et al.*, 1998	16	8.7	+	+				
Pop8	Chamberlain *et al.*, 1998	15.5	4.6	+	+	hPop8 (Rpp14)	Jarrous *et al.*, 1999a	14	7.6	+	+?				
Rpp1	Stolc *et al.*, 1997	32.2	9.8	+	+	hRpp1 (Rpp30)	Eder *et al.*, 1997	30	9.4	+	+	aRpp1 (Ph1877)	Hall *et al.*, 2002	24.7	9.6
Rpr2	Chamberlain *et al.*, 1998	16.3	10.0	+	-	hRpr2 (Rpp21)	Jarrous *et al.*, 2001b	18	9.6	+	-	aRpr2 (Ph1601)	Hall *et al.*, 2002	14.6	10.5
Snm1	Schmitt *et al.*, 1994	22.5	9.8	-	+				-	-	-				
Rmp1	Salinas *et al.*, 2005	23.6	9.8	-	+				-	-	-				
						Rpp40	Jarrous *et al.*, 1998	35	5.3	+	+				

Tableau 4. Composants des RNases P et MRP eucaryotes et de l'homologue archée de la RNase P.
Les composants des RNases P et MRP de *S. cerevisiae* et de *H. sapiens* sont détaillés pour illustrer les composants des complexes eucaryotes. Les composants de la RNase P de *P. horikoshii* sont présentés pour illustrer la RNase P archée. Les protéines ayant une homologue archée sont surlignées en orange. Le signe "+?" est indiqué lorsque la présence de la protéine dans le complexe est incertaine.

3.1.1.1.2. Caractérisation de la composante ARN des RNases P nucléaires

3.1.1.1.2.1. Organisation structurale de la composante ARN

Les structures secondaires des composantes ARN des RNases nucléaires de *S. cerevisiae* et d'*Homo sapiens* (**Figure 6**, panel du haut) ressemblent à celles des RNases P bactériennes et archées avec une plus grande proximité pour la composante ARN des RNases P archées de type M (**Figure 4**, type M). Les régions conservées décrites chez les bactéries sont présentes (Chen *et al.*, 1997) tandis que les éléments périphériques divergent fortement. Les éléments impliqués dans la stabilisation de la structure tertiaire des RNases P bactériennes sont, quant à eux, absents.

Les ARN des RNases P nucléaires partagent ainsi un certain nombre d'éléments structuraux avec les ARN des RNases P des autres règnes du vivant, tels que les hélices P1, P2, P3, P4, P7, P10/11, P12 (Frank *et al.*, 2000). La structure minimale consensus contient aussi chacune des régions conservées (CR-I à CR-V) trouvées dans les composantes ARN bactériennes et archées. Un sous-ensemble d'hélices est dit eucaryote (eP8, eP9, eP15 et eP19) du fait que ces hélices occupent la même position que les hélices bactériennes bien que le manque d'homologie dans la séquence ne permette pas de conclure si elles ont des fonctions analogues chez les eucaryotes (**Figure 6**, panel du haut). Les données phylogénétiques ont mis en évidence que l'hélice eP19 est présente dans presque toutes les composantes ARN des eucaryotes tandis que l'hélice eP15 n'est présente que dans les lignées fongiques. Les composantes ARN eucaryotes ne possèdent pas les nombreux éléments périphériques présents dans les ARN des RNases P bactériennes de type A (hélices P6, P13, P14, P18 et boucle P15) et de type B (P5.1, P10.1, P15.1 et P18) (Haas *et al.*, 1996). La comparaison des structures à haute résolution des composantes ARN des RNases P bactériennes de type A et B montrent que ces hélices sont localisées en périphérie de la structure où elles contribuent à stabiliser et orienter correctement le noyau de la structure (Krasilnikov *et al.*, 2003b; Krasilnikov *et al.*, 2004; Kazantsev *et al.*, 2005; Torres-Larios *et al.*, 2005). L'absence de structures comparables dans les ARN eucaryotes est vraisemblablement compensée par l'augmentation de la complexité de la composante protéique.

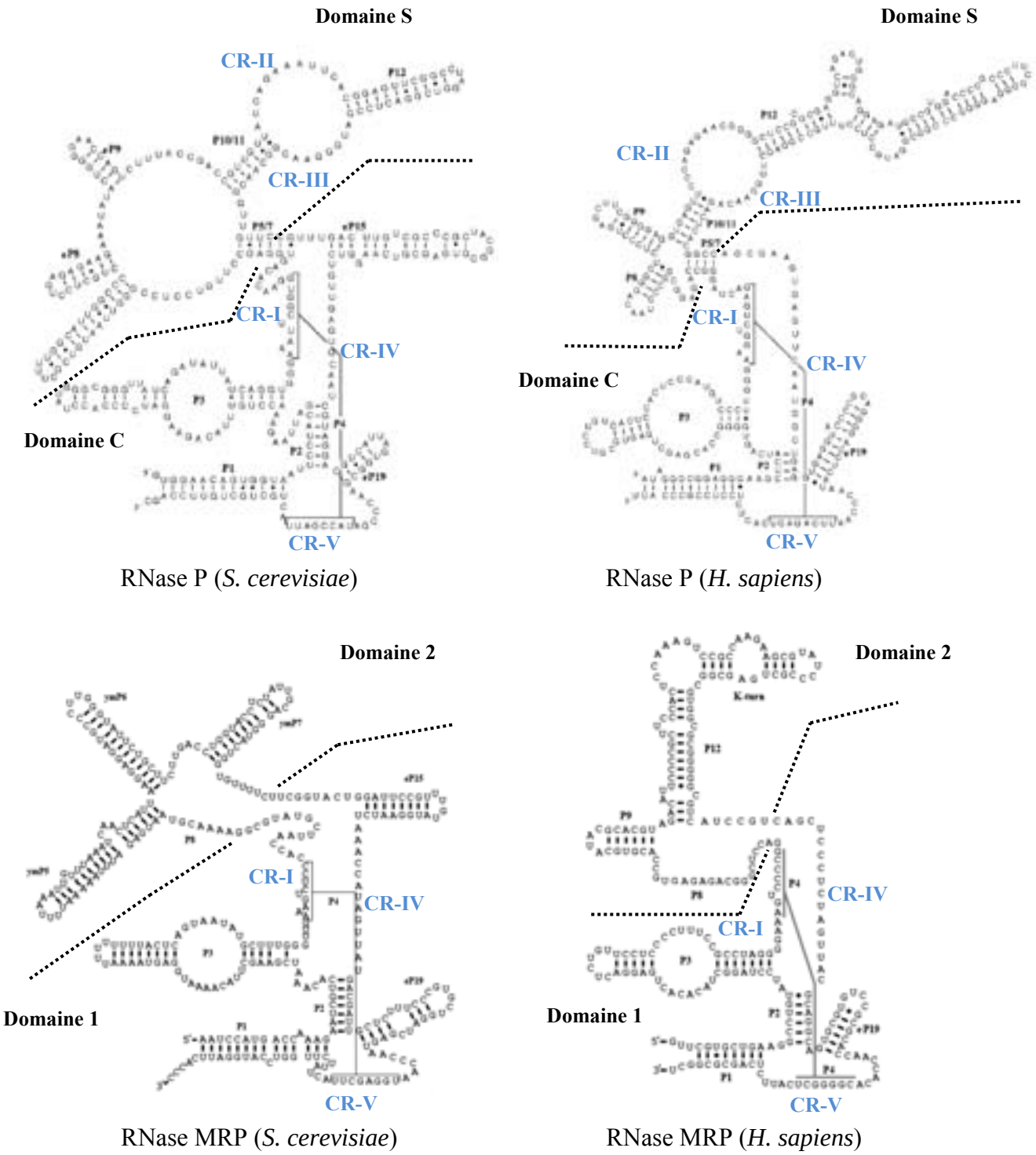


Figure 6. Comparaison des structures secondaires des ARN de RNases P eucaryotes et des RNases MRP.

Les structures sont celles des ARN de *S. cerevisiae* et *H. sapiens*. Les structures des ARN de RNases P sont issues de *RNase P database* (Brown, 1999). Les structures des ARN de RNases MRP sont issues de (Esakova *et al.*, 2010).

Les hélices sont annotées de P1 à eP19 selon la nomenclature de (Haas *et al.*, 1998). Les domaines S et C des ARN de RNase P et les domaines 1 et 2 des ARN de RNase MRP sont délimités par une ligne pointillée. Les régions conservées CR-I à CR-V sont indiquées le cas échéant.

Ainsi, chez les ARN des RNases P eucaryotes, l'hélice P3 a évolué vers une structure en hélice-boucle-hélice qui remplace la longue hélice P3 des ARN bactériens et archées (Lindahl *et al.*, 2000; Ziehler *et al.*, 2001; Piccinelli *et al.*, 2005). Cette région sert de site de liaison pour l'hétérodimère protéique Pop6-Pop7 (Pluk *et al.*, 1999; Perederina *et al.*, 2007; Welting *et al.*, 2007; Perederina *et al.*, 2010a) et probablement pour la protéine Pop1 (Ziehler *et al.*, 2001). La délétion de la région P3 de la composante ARN de la RNase P humaine conduit à une localisation erronée de cet ARN (Jacobson *et al.*, 1997). En outre, la purification de la RNase P humaine avec un ARN dont l'hélice P3 est absente ne permet pas de l'obtention d'une holoenzyme active (Li *et al.*, 2002). En effet, la région de l'hélice-boucle-hélice P3 avec ses protéines semble être impliquée dans le recrutement d'autres protéines essentielles pour nucléer l'assemblage de l'holoenzyme chez la levure et chez l'homme (Ziehler *et al.*, 2001; Li *et al.*, 2002; Perederina *et al.*, 2010b). En outre, une structure de la région P3 avec les protéines Pop6 et Pop7 a été résolue récemment (Perederina *et al.*, 2010a). Elle met en évidence que l'hétérodimère Pop6-Pop7 se lie directement ou indirectement sur l'hélice-boucle-hélice P3 selon une position qui suggère que ces protéines contribuent à compenser la perte des éléments structuraux impliqués dans la stabilisation de la structure tertiaire de l'ARN bactérien (**Figure 7**). Cette constatation est aussi soutenue par la modélisation de la composante ARN de *S. pombe* (Marquez *et al.*, 2006). Bien que les hélices P3 et P12 n'aient pas pu être modélisées, il a été mis en évidence que la zone d'interaction des hélices P5.1-P15.1 dans l'ARN bactérien de type B est adjacente à la position probable de l'hélice P3 eucaryote, site de liaison de Pop6-Pop7 qui pourraient donc assumer un rôle structural équivalent à l'interaction des hélices P5.1-P15.1 bactériennes (**Figure 8**).

L'autre spécificité des composantes ARN eucaryotes est la variation en longueur, en séquence et en structure des hélices P12 et P15 (**Figure 6**) (Brown, 1999). De plus, il est intéressant de noter que de nombreuses espèces de vertébrés ont conservé un motif structural dans l'hélice P12, le motif kink-turn (K-turn) (Klein *et al.*, 2001; Rosenblad *et al.*, 2006). Chez les archées, les structures ARN K-turn sont connues pour lier la protéine L7Ae, homologue de la protéine Pop3 chez la levure (Rozhdestvensky *et al.*, 2003; Hamma *et al.*, 2004; Dlakic, 2005).

Les composantes ARN des RNases P nucléaires contiennent également des parties similaires au domaine C et au domaine S des composantes bactériennes dont les régions conservées CR-I à CR-V (Pan, 1995; Loria *et al.*, 1996; Loria *et al.*, 1999).

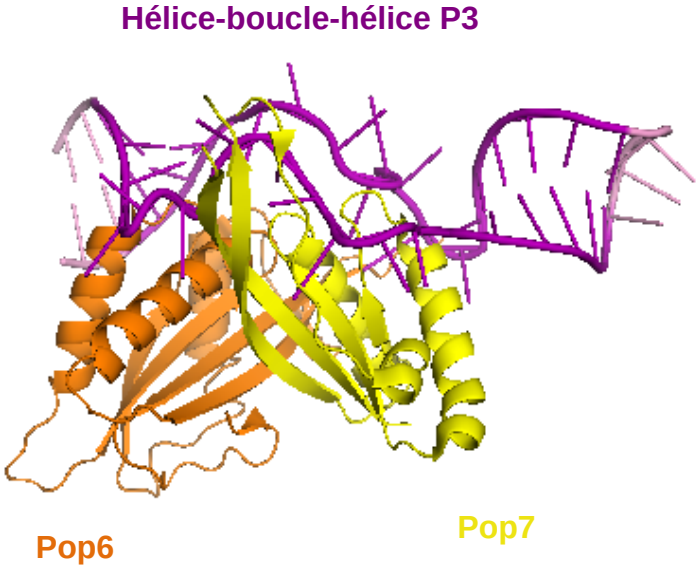


Figure 7. Structure à haute résolution du complexe Pop6-Pop7-hélice P3 issu de la RNase P de *S. cerevisiae*.
L'hélice P3 est représentée en violet Les nucléotides additionnels ayant permis la cristallisation sont en rose.
La protéine Pop6 est représentée en orange et la protéine Pop7 en jaune.

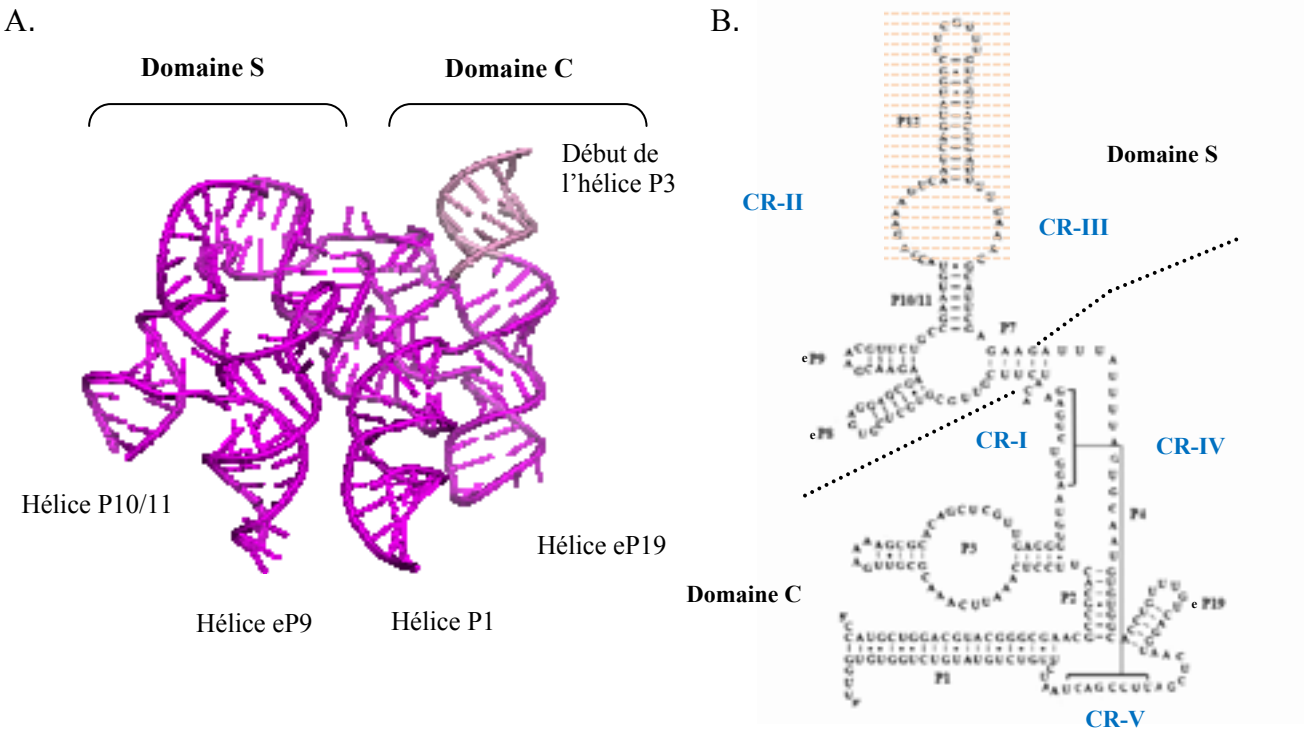


Figure 8. Structure de la composante ARN de la RNase P de *S. pombe*.
(A) Structure modélisée de la composante ARN de la RNase P (en magenta) à partir de la structure à haute résolution de la composante ARN de la RNase P bactérienne issue de *B. stearothermophilus* (Marquez *et al.*, 2006). Le début de l'hélice P3 modélisée est représenté en rose.
(B) Structure secondaire de l'ARN de RNase P. Les hélices sont annotées de P1 à P19 selon la nomenclature de (Haas *et al.*, 1998). Les domaines S et C des ARN de RNase P sont délimités par une ligne pointillée. Les régions conservées CR-I à CR-V sont indiquées. Le rectangle orange indique les régions manquantes dans la structure tridimensionnelle modélisée : 30 nt de la boucle-hélice P3 et 50 nt de la boucle-hélice P12.

Ces régions conservées CR-I, CR-II, CR-IV et CR-V ont été étudiées par mutations aléatoires des nucléotides conservés. Les mutations dans les régions CR-I, CR-IV et CR-V conduisent à une réduction de l'efficacité catalytique (kcat/Km) de l'holoenzyme *in vitro*, avec une plus grande influence des mutations sur la catalyse (kcat) que sur l'identification du substrat (Km). Ces résultats confirment la localisation de ces régions conservées à proximité du site catalytique du domaine C comme dans les RNases P bactériennes (Reiter *et al.*, 2010). Les mutations dans la région CR-II, participant à la boucle P12 avec la région CR-III, ont montré une dépendance au magnésium pour la croissance et une nécessité accrue en magnésium pour l'activité catalytique *in vitro* (Pagan-Ramos *et al.*, 1996; Ziehler *et al.*, 1998).

Chez la levure, il existe plusieurs autres régions dont les séquences semblent conservées. Tout d'abord, la séquence entre les hélices P4 et P7 (5'-GAAC-3') montre une certaine conservation adjacente à l'hélice P4 des composantes ARN des trois règnes du vivant. La mutation de ces nucléotides affecte la croissance, l'assemblage de l'holoenzyme et sa localisation chez la levure. Les enzymes avec ces mutations montrent un défaut sévère dans la maturation des ARNt prématures *in vitro* et une réduction (10 fois) de l'efficacité catalytique (kcat/Km) (Xiao *et al.*, 2005). Comme ces nucléotides conservés sont adjacents à l'hélice P4 et à proximité de la région catalytique, déterminée par homologie avec la structure de la composante ARN bactérienne (Reiter *et al.*, 2010), il est possible qu'ils soient utilisés pour une fonction spécifique chez les eucaryotes. Des séquences hautement conservées ont aussi été mis en évidence dans les boucles des hélices eP8 (5'-NUGA-3') et eP9 (5'-GNAA-3'). Les mutations dans la séquence de ces boucles affectent la croissance, la maturation des ARNt *in vivo* et l'assemblage de la RNase P. Les mutations dans la boucle de l'hélice eP9 ont un effet plus sévère que celles dans l'hélice eP8 et affectent *in vivo* la localisation cellulaire de la RNase P. De plus, les enzymes mutantes montrent des cinétiques aberrantes dans les tests de coupure endonucléolytique des ARNt prématures *in vitro* : si les mutations dans l'hélice eP9 conduisent à une réduction de l'efficacité catalytique, les mutation dans l'hélice eP8 conduisent à une augmentation inattendue de cette efficacité catalytique. Cette augmentation semble due à la perte de la sélectivité du substrat de type ARNt prémature ce qui rend l'enzyme mutante plus sensible à l'inhibition compétitive par les ARNt totaux de levure. Ainsi, les hélices eP8 et eP9 jouent un rôle clé dans la maturation et l'activité de l'holoenzyme. La conservation de la séquence dans les structures boucle-hélice suggère que les boucles seraient impliquées dans les interactions ARN-ARN et ARN–protéine. Cela est cohérent avec l'observation selon laquelle les boucles sont protégées contre les nucléases dans les études d'empreinte de la RNase P de *S. cerevisiae* (Tranguch *et al.*, 1994).

3.1.1.1.2.2. Caractérisation fonctionnelle de la composante ARN

D'un point de vue fonctionnel, la composante ARN des RNases P nucléaires reste la composante qui contient le domaine catalytique de l'holoenzyme (Thomas *et al.*, 2000; Kikovska *et al.*, 2007). Cependant, la constante catalytique observée pour la composante ARN de la RNase P nucléaire humaine en l'absence de sa composante protéique est très faible (0.000026 min-1) comparée à celle de la composante ARN des RNases P bactériennes (8 min-1) soit seulement environ 100 fois supérieure à la vitesse de réaction de l'hydrolyse spontanée de l'ARN dans les mêmes conditions expérimentales (Kikovska *et al.*, 2007). Cette diminution importante de l'activité catalytique est probablement due à l'absence d'éléments structuraux impliqués dans la stabilisation de la conformation catalytique active de la composante ARN, présents chez les bactéries mais absents des holoenzymes eucaryotes comme des RNases P archées (Gopalan, 2007).

3.1.1.1.3. Caractérisation structurale des RNases P nucléaires

Le but final des études sur les RNases P eucaryotes est de permettre la compréhension de l'architecture de l'holoenzyme et la façon dont sont liées la sélection des substrats et le mécanisme de catalyse. Actuellement, il est suggéré que les protéines eucaryotes Pop4, Rpr2, Pop5 et Rpp1, homologues aux quatre protéines archées pourraient former le noyau protéique de l'holoenzyme eucaryote (Walker *et al.*, 2010). Bien que des progrès aient été réalisés vers la reconstitution *in vitro* des RNases P nucléaires (Mann *et al.*, 2003), le principal obstacle à cette reconstitution concerne la production et la purification des protéines individuellement, étant donné que certaines sont difficiles à exprimer, insolubles et uniquement stables transitoirement (Houser-Scott *et al.*, 2002; Aspinall *et al.*, 2007; Perederina *et al.*, 2007). Cependant, des progrès significatifs ont été réalisés dans la reconstitution *in vitro* des RNases P archées (Tsai *et al.*, 2006; Honda *et al.*, 2008; Xu *et al.*, 2009; Honda *et al.*, 2010) qui, en raison de l'homologie entre les sous-unités protéiques archées et eucaryotes, devraient permettre d'éclairer les études sur les holoenzymes eucaryotes. Néanmoins, au regard de la complexité des systèmes eucaryotes comparés aux systèmes bactériens et archées, les connaissances sur l'architecture des RNases P eucaryotes restent encore peu étendues.

Tandis que la composante ARN est la sous-unité catalytique de la RNase P ribonucléoprotéique nucléaire, il semble très probable que la composante protéique soit requise pour assister la composante ARN et faciliter la fonction enzymatique. Chez la levure, les sous-unités peuvent individuellement être manipulées génétiquement ce qui a permis des études sur l'assemblage de l'holoenzyme *in vivo* (Ziehler *et al.*, 2001). De plus, les sous-unités protéiques

et ribonucléique doivent s'assembler les unes aux autres pour former le ribozyme actif. Des études d'interactions protéine-ARN et protéine-protéine ont ainsi été réalisées *in vitro* aussi bien sur les complexes RNases P que RNases MRP, complexes qui partagent sept protéines sur 8-10 selon les espèces et dont la composante ARN est fortement apparentée (voir paragraphe 3.3).

3.1.1.1.3.1. L'assemblage in vivo de la RNase P nucléaire

La maturation et l'assemblage de la RNase P nucléaire *in vivo* ont été les mieux caractérisés dans le modèle de levure. Chez *S. cerevisiae*, l'ARN RPR1 est initialement transcrit par l'ARN polymérase III sous la forme d'un long transcrit primaire allongé de 84 nt à l'extrémité 5' et de 16 à 30 nt à l'extrémité 3' (Lee *et al.*, 1991b). Aucun ARN précurseur n'ayant été mis en évidence chez *H. sapiens*, cette voie de maturation semble donc être spécifique des *Saccharomycètes*. Le transcrit prémature chez *S. cerevisiae* subirait une coupure unique au niveau de son extrémité 5' et de multiples coupures à son extrémité 3' pour obtenir un transcrit mature de 369 nt (Lee *et al.*, 1991a), cependant les différentes étapes de cette maturation ne sont pas connues. La RNase P avec la composante ARN prémature a été purifiée grâce à une étiquette d'affinité insérée dans sa séquence 5' leader (Srisawat *et al.*, 2002). Cette RNase P précurseur est capable de réaliser la coupure de l'extrémité 5' des ARNt prématures bien que les protéines Pop3 et Rpr2 n'aient pas ou peu été détectées dans l'holoenzyme. La protéine Pop3 est aussi une sous-unité de la RNase MRP et la pertinence de son absence dans la RNase P précurseur purifiée est difficile à interpréter. La protéine Rpr2, quant à elle, est à la fois essentielle et spécifique de la RNase P, ce qui semble donc indiquer qu'elle est requise pour d'autres fonctions indépendantes de la coupure des ARNt prématures.

Au niveau de la composante ARN, la structure en hélice-boucle-hélice P3 s'est révélée importante non seulement pour la localisation nucléolaire des RNases P nucléaires mais aussi pour la maturation de l'ARN RPR1 (Ziehler *et al.*, 2001). Une étude avec des composantes ARN mutées aléatoirement sur 4 nt conservés de l'hélice P3 a montré que les composantes mutées ont une maturation réduite de l'ARN RPR1 précurseur qui coïncide avec une perte d'efficacité de la maturation des ARNt. Par ailleurs, cette étude a mis en évidence que l'hélice P3 est le site de fixation préférentiel de la protéine Pop1, mais que cette fixation est abolie en présence de mutations au niveau de résidus conservés de l'hélice P3. De plus, l'hélice P3 interagit aussi avec les protéines Pop6-Pop7 ce qui provoque un réarrangement local de la boucle P3 (Perederina *et al.*, 2007). Perederina et ses collègues ont alors suggéré que l'interaction directe entre la région P3 et l'hétérodimère Pop6-Pop7 pourrait contribuer au

recrutement des autres protéines de l'holoenzyme lors de son assemblage. L'hélice P3 eucaryote semble donc être une région dont l'interaction avec les sous-unités protéiques est critique pour la maturation de la RNase P nucléaire (Ziehler *et al.*, 2001).

Par ailleurs, l'ARN RPR1 précurseur semble être associé avec un complexe de protéines Sm-like (Lsm), le complexe Lsm 2-8, ce qui n'est pas le cas de l'ARN RPR1 mature (Salgado-Garrido *et al.*, 1999; Pannone *et al.*, 2001; Kufel *et al.*, 2002). Le complexe Lsm 2-8 interagirait *in vitro* avec les transcrits RPR1 précurseurs naissants de l'ARN polymérase III *via* l'extrémité 3' polyuracilée. Les homologues archées des protéines Sm ont aussi été mis en évidence en association avec la composante ARN des RNases P archées (Toro *et al.*, 2001). De plus, la baisse du niveau d'expression des protéines Lsm2, Lsm5 et Lsm8 conduit à l'accumulation d'ARNt prématures *in vivo* ce qui suggère un rôle essentiel de ce complexe dans la maturation des ARNt (Kufel *et al.*, 2002).

Il semble donc que le processus d'assemblage des RNases P nucléaires soit un processus complexe dont toutes les étapes ne sont pas encore explorées.

3.1.1.1.3.2. Les interactions ARN-protéine au sein de la RNase P nucléaire

Le système triple hybride en levure a été utilisé pour démontrer les interactions individuelles entre la composante ARN de *S. cerevisiae*, RPR1 et les protéines Pop1 et Pop4 (Houser-Scott *et al.*, 2002). Il a ainsi été montré que l'interaction RPR1-Pop1 est spécifique de l'hélice P3 et qu'elle peut être abolie par des mutations dans la boucle interne de l'hélice P3 (Ziehler *et al.*, 2001).

Une approche similaire a été utilisée pour mettre en évidence l'interaction entre la composante ARN humaine H1 et les protéines hPop3 (Rpp38), hPop4 (Rpp29), hRpp1 (Rpp30) et hRpr2 (Rpp21) (Jiang *et al.*, 2001b). Bien que cette étude n'ait pas mis en évidence d'interaction ARN H1-hPop1, une autre étude utilisant des co-purifications avec l'étiquette GST a révélé cette interaction en impliquant l'hélice P3 comme site probable de fixation (Welting *et al.*, 2004). L'hélice P3 a aussi été décrite comme le site de liaison de l'hétérodimère Pop6-Pop7 qui s'y lie avec une affinité de l'ordre du nanomolaire (Perederina *et al.*, 2007). La protéine Pop6 ne semble pas, quant à elle, se lier individuellement à la sous-unité ARN, mais sans elle, la protéine Pop7 est peu soluble et instable.

D'autres études ont utilisé des approches par co-purifications avec l'étiquette GST et fixation aux rayons ultra-violets pour explorer la liaison des protéines individuelles avec les composantes ARN des RNases P et des RNases MRP de *S. cerevisiae* et de *H. sapiens* (Pluk *et al.*, 1999; Jiang *et al.*, 2001b; Van Eenennaam *et al.*, 2002; Welting *et al.*, 2004; Aspinall *et al.*,

2007; Perederina *et al.*, 2007). La protéine Pop4 de levure et son homologue humaine, Rpp29 ont toutes deux été décrites comme essentielles à leur holoenzyme respective notamment pour se lier à leur composante ARN (Jiang *et al.*, 2001b; Houser-Scott *et al.*, 2002; Welting *et al.*, 2004; Aspinall *et al.*, 2007).

La protéine hPop3 (Rpp38) semble lier l'ARN de la RNase MRP humaine dans l'hélice P12 qui porte un motif K-turn (Welting *et al.*, 2004). Or, il a été prédit que hPop3 est un homologue de la protéine L7Ae archée connue pour se lier aux motifs K-turn des petits ARN nucléolaires archées (Rozhdestvensky *et al.*, 2003; Hamma *et al.*, 2004). Des motifs K-turn ont été identifiés dans les hélices P12 des composantes ARN des RNases P nucléaires et des RNases MRP (Rosenblad *et al.*, 2006). Cependant, malgré la présence de la protéine Pop3 dans les levures, le motif K-turn est absent leurs composantes ARN. Il semble donc que cette particularité soit spécifique de cette lignée d'eucaryotes (Rosenblad *et al.*, 2006). Une interaction forte entre Pop3 et le substrat des RNases P, les ARNt prématures a également été rapportée (Brusca *et al.*, 2001).

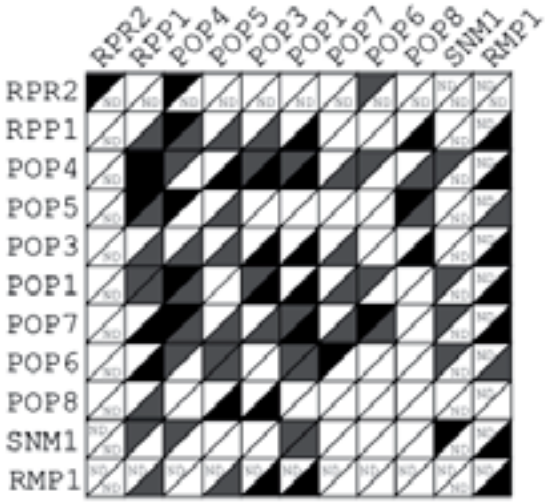
Bien qu'un grand nombre d'interaction ARN-protéine ait été rapporté dans les complexes RNases P et MRP, en utilisant des approches variées, seules quelques interactions sont confirmées par les différentes études: entre Pop1 et l'hélice P3, entre Pop6-Pop7 et l'hélice P3, entre Pop4 et l'ARN et entre Pop3 et le motif k-turn de l'hélice P12.

3.1.1.1.3.3. Les interactions protéine-protéine au sein de la RNase P nucléaire

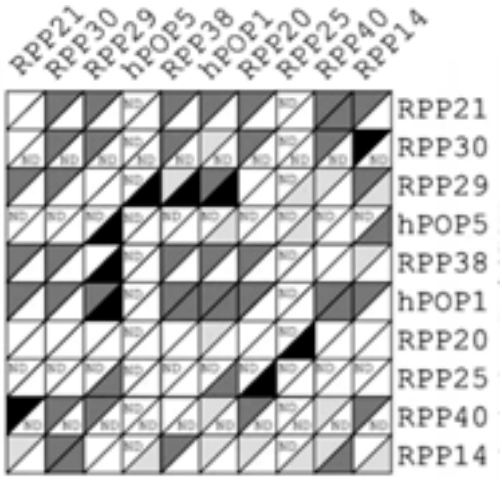
Les interactions protéine-protéine ont-elles aussi été très étudiées chez la levure et chez l'homme avec le système double hybride en levure (Y2H) (Jiang *et al.*, 2001a; Houser-Scott *et al.*, 2002) et par des techniques de copurification avec une étiquette GST (Welting *et al.*, 2004; Aspinall *et al.*, 2007). Ces dernières n'ont malheureusement pas inclus de version étiquetée des protéines hRpp1 (Rpp30) et Rpp40 chez l'homme et de la protéine Rpr2 chez la levure tandis que les études de double hybride ne couvrent pas les protéines hPop5 et hPop6 (Rpp25) chez l'homme. Les résultats de ces études sont récapitulés dans la **Figure 9**.

Selon les études, de nombreuses interactions ont été décrites avec une cohérence variable. Cependant, étant donné l'homologie structurale et la conservation des acides aminés essentiels entre les protéines Pop4, Rpr2, Rpp1 et Pop5 avec leurs homologues archées, il est raisonnable de penser que les interactions fortes entre aPop4 et aRpr2 et entre aRpp1 et aPop5 (Kawano *et al.*, 2006; Wilson *et al.*, 2006; Honda *et al.*, 2008; Honda *et al.*, 2010) sont conservées chez les eucaryotes bien que les résultats selon les études ne soutiennent pas aussi fortement ces interactions.

A. Chez *S. cerevisiae*



B. Chez *H. sapiens*



Légende

Adaptée de Walker *et al.*, 2006 et de Walker *et al.*, 2010

Figure 9. Récapitulatif des interactions protéine-protéine mises en évidence dans les RNases P et MRP de *S. cerevisiae* et de *H. sapiens.*
(Jiang *et al.* , 2001a; Houser-Scott *et al.* , 2002; Welting *et al.* , 2004; Aspinall *et al.* , 2007).
A. Interactions mises en évidence chez la levure *S. cerevisiae*. Les interactions mises en évidence par double hybride (Y2H) sont indiquées comme décrit dans le texte original (Houser-Scott *et al.*, 2002). Les interactions mises en évidence par GST-pulldown (GST) présentées ci-dessus sont les interactions correspondant aux résultats obtenus dans les conditions les plus stringentes (300 mM NaCl) (Aspinall *et al.*, 2007).
B. Interactions mises en évidence chez *H. sapiens*. Les interactions sont indiquées comme décrit dans le texte original (Jiang *et al.*, 2001a; Welting *et al.*, 2004).
Les protéines appâts sont représentées en ordonnée et les protéines cibles en abscisse.

Les interactions les mieux soutenues par les quatre différentes études concernent l'interaction entre les protéines Pop1 (hPop1) et Pop4 (Rpp29) et entre Pop1 et Pop3 (Rpp38). La protéine Pop4 (Rpp29) semble aussi interagir avec les protéines Pop3 (Rpp38), Rpp1 (Rpp30) et Pop5. La dernière interaction mise en évidence concerne l'interaction entre les protéines Pop6 (Rpp25) et Pop7 (Rpp20) dont le complexe a été caractérisé et cristallisé en présence de l'hélice P3 (Perederina *et al.*, 2007; Perederina *et al.*, 2010a). De nombreuses autres interactions ont été mises en évidence mais sont peu soutenues selon les études, c'est pourquoi leur pertinence est plus discutable.

Si ces études décrivent les interactions entre deux protéines, il est important de remarquer qu'elles n'informent pas sur la stœchiométrie des sous-unités protéiques dans la RNase P nucléaire. Une analyse préliminaire des sous-unités protéiques de la RNase MRP de *S. cerevisiae* avec du SYPRO-Ruby, un fluorophore se fixant au squelette polypeptidique des protéines de manière indépendante de leur composition en acide aminé, suggère que la composition en protéine des holoenzymes eucaryotes pourrait être très complexe avec de nombreuses sous-unités présentes en plusieurs exemplaires (Salinas *et al.*, 2005).

Des modèles de l'architecture des holoenzymes eucaryotes ont été réalisés à partir de ces interactions ARN-protéine et protéine-protéine (Jiang *et al.*, 2001a; Jiang *et al.*, 2001b; Houser-Scott *et al.*, 2002; Welting *et al.*, 2004). Cependant, des connaissances plus détaillées des sous-complexes eucaryotes seraient utiles pour affiner les modèles actuels des RNases P. Une complication dans l'étude des RNases P eucaryotes provient de l'existence de la RNase MRP constituée d'un ARN fortement apparenté et d'une composante protéique quasi identique. Les RNases P nucléaires et les RNases MRP étant très similaires, les études de ces deux holoenzymes seront utiles pour la détermination des contributions relatives aux composantes ARN et protéiques dans la localisation, la sélection des substrats et la catalyse.

3.1.1.1.4. Etude des substrats des RNases P nucléaires

3.1.1.1.4.1. Particularités dans la reconnaissance du substrat

Comme pour la RNase P bactérienne, les études sur la RNase P nucléaire se sont surtout portées sur son rôle dans la maturation des ARNt prématures. Il existe de subtiles différences entre les ARNt prématures bactériens et eucaryotes qui pourraient potentiellement influencer l'identification du substrat par la RNase P.

Les ARNt prématures eucaryotes sont transcrits par l'ARN polymérase III et se terminent par un oligomère d'uraciles (U_{4-6}) en 3', juste après la fin de l'hélice aminoacyle. Cet

oligomère forme avec la séquence 5' leader une courte hélice commençant par une boucle. La présence de la séquence polyuracile en 3' augmente légèrement le taux de coupure du substrat tandis que la boucle au début de l'hélice 5' leader-3' terminale laisse non apparié le nucléotide en position -1 du site de coupure de la RNase P ce qui semble être un pré-requis pour une coupure efficace (Lee *et al.*, 1997).

Chez les bactéries, les ARNt prématures sont souvent transcrits avec une séquence terminale 3'-CCA, séquence reconnue par la boucle P15 de l'holoenzyme bactérienne (Kirsebom, 2007). Or, chez les eucaryotes, la séquence 3'-CCA n'est pas codée par la séquence ADN des ARNt et est ajoutée à l'ARNt post-transcriptionnellement (Hopper *et al.*, 2003). C'est probablement en réponse à l'absence de cette séquence 3'-CCA que l'hélice P15 de type bactérien a été perdue au cours de l'évolution (Evans *et al.*, 2006). Actuellement, il n'a pas été déterminé clairement quelle partie de l'holoenzyme eucaryote interagit avec les séquences 5' leader et 3' terminateur des ARNt prématures, puisque le retrait de ces séquences n'a qu'un effet mineur sur la liaison du substrat sur la RNase P (Ziehler *et al.*, 2000).

En dépit de ces différences dans la séquence des transcrits primaires, les structures des ARNt sont extrêmement bien conservées et les RNases P bactériennes, archées et eucaryotes sont chacune capable de reconnaitre et de couper les ARNt prématures de chacun des trois règnes du vivant.

En plus des différences liées aux substrats, il a été observé que les holoenzymes bactériennes et eucaryotes réagissent différemment envers des substrats non-ARNt. Ainsi, la RNase P humaine ne coupe pas les mêmes substrats minimaux que la RNase P bactérienne. Pour son activité endonucléolytique, elle requière la présence d'une boucle additionnelle entre la branche accepteur et le domaine branche T-boucle T. Cette boucle peut n'être composée que d'un nucléotide mais la flexibilité qu'elle apporte semble favoriser la réaction de coupure (Yuan *et al.*, 1995).

Pour pallier la perte de certains des contacts bactériens RNase P-ARNt prématures, les RNases P nucléaires ont acquis de nouveaux contacts spécifiques des eucaryotes avec l'ARN simple brin. En effet, les RNases P eucaryotes lient plus fortement les ARN simple brin que les RNases P bactériennes, ces contacts pouvant par compétition même inhiber la coupure des ARNt prématures chez la RNase P de *S. cerevisiae* (Ziehler *et al.*, 2000). La RNase P de levure est ainsi capable de lier et d'être fortement inhibée par des homoribopolymères simple brin (polyG>polyU>>polyA>>polyC) tandis que des ARN hautement structurés comme l'ARN 5S n'ont pas ou peu montré de compétition avec les ARNt prématures (Ziehler *et al.*, 2000). Les

holoenzymes bactériennes et eucaryotes montrent aussi des motifs de coupure différents lors de la coupure *in vitro* de précurseurs d'ARN ribosomiques (Chamberlain *et al.*, 1996) ce qui suggère un mécanisme de reconnaissance différent du substrat entre les RNases P bactériennes et eucaryotes.

Les interactions entre les RNases P eucaryotes et les ARNt prématures semblent être influencées par la protéine La (La-like chez la levure) liées aux ARNt prématures (Yoo *et al.*, 1997). La protéine La se lie à leur extrémité 3' polyuracylée et les protège de la digestion exonucléolytique (Maraia *et al.*, 2001; Wolin *et al.*, 2002). Dans les cellules humaines, la protéine La atténue le processus de maturation de l'extrémité 5' des ARNt prématures. Mais cet effet est réduit par sa phosphorylation bien qu'elle reste liée aux ARNt prématures (Fan *et al.*, 1998; Intine *et al.*, 2000). Il semble que la protéine La phosphorylée facilite la maturation des ARNt par la RNase P en dénaturant la structure secondaire des séquences terminales 3' du substrat (Xiao *et al.*, 2002).

Plusieurs études ont mis en évidence l'existence de substrats non ARNt variés pour la RNase P. Chez *S. cerevisiae*, l'ARN antisens HRA1 (*hidden-in-reading-frame antisense 1 RNA*) a été identifié comme substrat potentiel ce qui a été vérifié *in vitro* (Samanta *et al.*, 2006; Yang *et al.*, 2007). Un rôle de la RNase P dans la maturation des petits ARN nucléolaires (snoARN) à boîtes C/D a également été suggéré. Ces petits ARN nucléolaires sont en effet copurifés avec la RNase P et leurs précurseurs s'accumulent dans les mutants de la RNase P sensibles à la température. Cependant, les sites de coupures dans les snoARN par la RNase P n'ont pas pu être déterminés dans les réactions *in vitro* (Coughlin *et al.*, 2008). Par ailleurs, un mutant de l'ARN RPR1 sensible à la température montre une accumulation d'une forme d'ARN 5.8S allongée d'environ 35 nt à température non permissive (Chamberlain *et al.*, 1996). Cependant, il n'a été pas établi s'il s'agit d'un effet direct ou indirect de la RNase P, même si la colocalisation de la RNase P et de la biogenèse ribosomale dans le nucléole rendent une implication directe possible.

Dans les cellules humaines, la RNase P est impliquée dans la maturation de l'ARN MALAT1 (*metastasis-associated lung adenocarcinoma transcript 1*), un ARN non codant qui est dérégulé dans de nombreux cancers. La RNase P coupe un segment de structure proche de celle des ARNt dans la région 3' des ARN MALAT1 naissants, générant une extrémité 3' mature d'un long ARN MALAT1 et une extrémité 5' d'un petit ARN cytoplasmique appelé *MALAT1-associated small cytoplasmic RNA* de 61 nt qui ressemble à un ARNt (Wilusz *et al.*, 2008).

Chez la levure, d'autres substrats potentiels ont été identifiés lors de l'étude de la copurification d'ARN avec la RNase P affectée par un défaut soit dans la composante ARN, soit dans la sous-unité protéique Rpr2 spécifique de la RNase P, soit dans la grande sous-unité protéique, Pop1 (Coughlin *et al.*, 2008). Parmi les ARN purifiés se trouvent notamment des ARNm qui codent des sous-unités protéiques des ribosomes, des sous-unités protéiques des ARN polymérases I, II et III, des facteurs d'initiation de traduction, des protéines associées aux snoARN ainsi que des transcrits de six régions intergéniques. Cependant, la méthodologie employée ne permet pas de distinguer le sens des brins copurifiés (sens versus antisens) alors que parmi ces ARN, certains sont connus pour avoir des transcrits antisens. Enfin, un autre ensemble de substrats potentiels a été identifié (Samanta *et al.*, 2006) et se compose d'ARN non codants adjacents ou antisens à des gènes codant des protéines.

L'ensemble de ces données suggère l'existence d'un large pool de substrats potentiels pour la RNase P, mais de plus amples études sont nécessaires pour étudier la pertinence physiologique de chacun de ces candidats.

3.1.1.1.4.2. Particularités dans la coupure du substrat

En dépit de l'accroissement de la complexité des RNases P eucaryotes par rapport aux RNases P bactériennes, il semble que la coupure des ARNt prématures soit réalisée selon un mécanisme de même nature également localisé dans la composante ARN.

En effet, des expériences de substitutions avec du phosphorothioate dans les substrats ont montré que les RNases P eucaryotes ont la même sorte de dépendance aux ions métalliques que les RNases P bactériennes (Pfeiffer *et al.*, 2000). Il a aussi été montré que la composante ARN de la RNase P humaine est capable de couper seule les ARNt prématures mais avec une activité extrêmement faible et à haute force ionique (Kikovska *et al.*, 2007). Le déroulement de la réaction de coupure des ARNt prématures semble être conservé avec une formation initiale rapide d'ARNt suivie d'une étape où l'activité est limitée par le processus de libération des produits de réaction (Hsieh *et al.*, 2009). Enfin, une analyse exploratoire de la RNase P nucléaire de *S. cerevisiae* a montré que la constante catalytique à l'état d'équilibre est très largement dépendante de la température, cette dépendance énergétique suggérant qu'un changement conformationnel important cinétiquement aurait lieu lors de la catalyse de manière semblable à celui de la RNase P bactérienne (Hsieh *et al.*, 2009).

Toutes ces constatations suggèrent donc l'existence d'un mécanisme conservé propre à la nature ribonucléique du noyau catalytique des RNases P nucléaires.

3.1.2. La RNase P RNP mitochondriale de *S. cerevisiae*

Chez *S. cerevisiae*, l'holoenzyme mitochondriale se compose d'une sous-unité ARN essentielle, RPM1 d'environ 490 nt (Morales *et al.*, 1989), qui est codée par le génome mitochondrial et d'une sous-unité protéique nucléaire Rpm2 de 105 kDa (Morales *et al.*, 1992).

L'ARN RPM1 a perdu de la complexité structurale comparé à l'ARN RPR1 de la RNase P nucléaire mais partage encore quelques éléments conservés, principalement les hélices P1, P4 et P18 essentielles à l'activité endonucléolytique (Morales *et al.*, 1989; Seif *et al.*, 2003). Etant donné la taille relativement large de la protéine Rpm2, la perte en complexité de la composante ARN pourrait être compensée par la composante protéique de la RNase P mitochondriale. Contrairement à la composante ARN RPM1, la sous-unité protéique Rpm2 chez *S. cerevisiae* montre peu de similarité avec la séquence des autres protéines des RNases P ribonucléoprotéiques. Elle est localisée aussi bien dans le noyau que dans les mitochondries. La protéine Rpm2 est une protéine multifonctionnelle qui semble agir à plusieurs niveaux dans les cellules en plus de sa fonction dans la maturation des ARNt. Ainsi, Rpm2 agit en tant que suppresseur d'un allèle mutant sensible à la température du composant protéique de canaux d'import mitochondriaux, Isp42, et est requis pour la croissance (Kassenbrock *et al.*, 1995). Elle est impliquée dans la biogenèse mitochondriale (Stribinskis *et al.*, 2001) et dans la maturation de la composante ARN de la RNase P mitochondriale (Stribinskis *et al.*, 1996). Une partie des protéines Rmp2 est localisée dans le noyau où elle agit en tant qu'activateur transcriptionnel de gènes codant des protéines mitochondriales notamment des protéines liées au complexe d'import (Stribinskis *et al.*, 2005). En outre, elle intervient dans la coordination de la transcription, la dégradation des ARNm et le stockage dans des structures cytoplasmiques appelées P-bodies (*Processing bodies*) (Stribinskis *et al.*, 2007)

3.2. Les RNases P protéiques

Bien que l'activité RNase P soit réalisée dans la majorité des espèces du règne du vivant par un complexe ribonucléoprotéique, il a été mis en évidence que cette activité pouvait avoir lieu sans la présence d'un ribozyme dans les chloroplastes d'épinard (Wang *et al.*, 1988) ainsi que dans les mitochondries humaines (Holzmann *et al.*, 2008). Par ailleurs, selon les études phylogénétiques, le gène codant la composante ARN des RNases P semble absent chez la majorité des végétaux et chez certains protistes (Hartmann *et al.*, 2003; Piccinelli *et al.*, 2005; Rosenblad *et al.*, 2006). Il a donc été suggéré que ces organismes pourraient avoir complètement remplacé les RNases P ribonucléoprotéiques par une autre sorte de RNase P strictement protéique (Gutmann *et al.*, 2012).

3.2.1. La RNase P mitochondriale multiprotéique de *H. sapiens*

Malgré la présence d'une composante ARN dans la RNase P mitochondriale de la levure, il semble que les organelles d'autres espèces aient développé, au travers d'une évolution convergente, une stratégie à base uniquement de protéines pour réaliser les activités de la RNase P. Bien que la nature de ces enzymes "non ribozymes" ait été controversée initialement, il a été montré que la RNase P mitochondriale humaine ne contient pas de composante ARN. Un complexe composé de trois protéines suffit à couper *in vitro* la séquence 5' leader des ARNt prématures au niveau du site attendu (Holzmann *et al.*, 2008). Ce complexe se compose d'une ARNt méthyltransferase (MRPP1 pour *Mitochondrial RNase P Protein 1*) de 42 kDa, d'un membre de la famille déshydrogénase/réductase (MRPP2) de 26 kDa et d'une métallonucléase MRPP3 de 60 kDa responsable de l'hydrolyse du pont phosphodiester (Holzmann *et al.*, 2009).

La protéine MRPP1 est une protéine dont l'homologue chez la levure est responsable d'une modification des ARNt en position 9 (Jackman *et al*, 2003) et pourrait donc lier l'ARNt. La protéine MRPP2 possède un domaine "Rossmann fold" et pourrait donc être impliquée dans la liaison à l'ARN (Holzmann *et al.*, 2008). La protéine MRPP3, quant à elle, est composée d'un domaine métallonucléase NYN conservé (Anantharaman *et al.*, 2006), d'un motif doigt de zinc à deux cystéines et de plusieurs motifs répétés pentatricopeptide (PPR) (Rossmanith, 2012) suggérant une implication dans les mécanismes post-transcriptionnels.

Enfin, ces sous-unités protéiques ne présentent aucune similarité avec les protéines des holoenzymes RNases P ribonucléoprotéiques des différents règnes.

3.2.2. Les RNases P chez les végétaux

Récemment, il a été mis en évidence l'existence de plusieurs protéines chez *Arabidopsis thaliana*, homologues à la protéine MRPP3 de la RNase P mitochondriale humaine : PRORP1, 2 et 3 (*PROteinaceous RNase P*) d'environ 60 kDa (Gobert *et al.*, 2010). Des études de dérégulation des gènes codant ces protéines par ARN interférents introduits par infection virale ont mis en évidence que l'absence de PRORP1 conduit à l'accumulation d'ARNt prématures dans les organelles tandis que l'absence de PRORP2 et 3 se traduit par une accumulation nucléaire. (Gutmann *et al.*, 2012). De plus, ces études ont mis en évidence le caractère essentiel de ces protéines pour la survie cellulaire chez *A. thaliana*.

La protéine PRORP1, localisée dans les plastides et les mitochondries, est capable de réaliser seule la coupure endonucléolytique des ARNt prématures et est impliquée dans la maturation d'ARNm mitochondriaux ayant une structure de type ARNt (Gobert *et al.*, 2010;

Gutmann *et al.*, 2012). Elle est, de plus, capable de remplacer la RNase P bactérienne dans *E. coli* (Gobert *et al.*, 2010). Les protéines PRORP2 et 3, localisées dans le noyau, sont non seulement impliquées dans la maturation des ARNt prématures mais aussi dans la maturation des petits ARN nucléolaires, comme la RNase P nucléaire de *S. cerevisiae* (Coughlin *et al.*, 2008; Gutmann *et al.*, 2012).

Il semble donc que l'activité RNase P soit portée uniquement par des protéines chez *A. thaliana*.

3.2.3. Les RNases P de *Trypanosoma brucei*

Très récemment, deux protéines homologues des protéines PRORP ont été identifiées chez *T. brucei*. Ces protéines d'environ 65-70 kDa présentent la particularité de contenir une insertion de 35 à 41 acides aminés dans le domaine métallonucléase NYN, contrairement aux protéines végétales et humaines. La protéine PRORP2 de *T. brucei* contient de plus une autre insertion de 13 à 18 acides aminés dans la région C-terminale de ce même domaine.

La protéine PRORP1 est localisée dans le nucléole tandis que PRORP2 se trouve dans les mitochondries. Il a été mis en évidence que ces deux protéines sont les seules formes de RNase P chez le trypanosome. De plus, la protéine PRORP1 est capable de suppléer à l'absence de la composante ARN de la RNase P nucléaire chez la levure *in vivo*. Cette protéine, seule, est donc capable de remplir toutes les fonctions d'une RNase P ribonucléoprotéique.

Aucun homologue de protéines PRORP n'a été identifié chez les bactéries et les archées. Néanmoins, ces protéines semblent être présentes dans un large éventail d'espèces eucaryotes suggérant une origine à la base de la phylogénie des eucaryotes et contribuant à expliquer que les études phylogénétiques n'aient pas mis en évidence de gènes de la composante ARN de la RNase P ribonucléoprotéique chez tous les organismes vivants. La caractérisation de ces protéines et de leur mode d'action devrait permettre dans le futur de les comparer aux RNases P ribonucléoprotéiques pour en apprécier les particularités tant au niveau de l'identification du substrat que du mécanisme de coupure endonucléolytique.

Dans les parties suivantes de ce manuscrit, il sera fait référence à la RNase P nucléaire ribonucléoprotéique sous le terme simplifié "RNase P".

3.3. Les RNases MRP

Chez les eucaryotes, la lignée des RNases P s'est divisée en deux, donnant naissance à une autre endoribonucléase site-spécifique à base d'ARN, la RNase MRP (*Mitochondrial RNA Processing*) (Chang *et al.*, 1987b; Chang *et al.*, 1987a; Karwan *et al.*, 1991). Comme la RNase P, la RNase MRP est essentielle pour la viabilité des cellules eucaryotes (Schmitt *et al.*, 1992). Malgré sa dénomination, la RNase MRP est localisée principalement dans le nucléole (Chang *et al.*, 1987b; Reimer *et al.*, 1988; Yuan *et al.*, 1989; Topper *et al.*, 1990; Kiss *et al.*, 1992a) et transitoirement dans le cytoplasme (Gill *et al.*, 2006). Une RNase MRP mitochondriale existe et semble être impliquée dans la maturation des ARN mitochondriaux en générant les amorces ARN nécessaire à la réplication de l'ADN mitochondrial (Chang *et al.*, 1987b; Stohl *et al.*, 1992; Topper *et al.*, 1992). La RNase MRP mitochondriale partage la même composante ARN que la RNase MRP nucléaire mais possède une composante protéique et une spécificité distinctes (Lu *et al.*, 2010) qui n'ont pas encore été totalement caractérisées. Dans la suite de ce manuscrit, il ne sera fait référence qu'à la RNase MRP nucléaire.

La RNase MRP semble être une enzyme eucaryote universelle pour laquelle la composante ARN a été mise en évidence chez tous les eucaryotes étudiées à l'exception de certains protistes (Piccinelli *et al.*, 2005; Rosenblad *et al.*, 2006). De plus, la similarité entre les ARN des RNases P et MRP, leur conservation au sein des eucaryotes et le fait que chez certains protistes, les gènes des ARN des RNases P et MRP soient en tandem suggèrent une relation phylogénétique forte entre les RNases P et MRP qui remontent aux origines des eucaryotes (Piccinelli *et al.*, 2005; Rosenblad *et al.*, 2006; Zhu *et al.*, 2006; Woodhams *et al.*, 2007). La RNase MRP a été plus particulièrement caractérisée chez *S. cerevisiae* et *H. sapiens* et commence à être étudiée chez les plantes (Kiss *et al.*, 1992b; Krehan *et al.*, 2012).

3.3.1. Structure des RNases MRP

3.3.1.1. Composition des RNases MRP

La RNase MRP est constituée d'une composante ARN apparentée à celle de la RNase P et d'une composante protéique dont la majorité des sous-unités sont partagées avec la RNase P (**Tableau 4**). La RNase MRP est donc fortement similaire à la RNase P au sein d'un organisme donné.

Chez *S. cerevisiae*, la RNase MRP se compose d'un ARN de 339 nt, NME1 (*Nuclear Mitochondrial processing Endoribonuclease 1*) (Schmitt *et al.*, 1992) et de dix protéines dont huit sont partagées avec la RNase P (Pop1, Pop3-8, Rpp1) (Lygerou *et al.*, 1994; Dichtl *et al.*,

1997; Stolc *et al.*, 1997; Chamberlain *et al.*, 1998; Salinas *et al.*, 2005). Tandis que la RNase P contient spécifiquement la protéine Rpr2, la RNase MRP possède deux protéines qui lui sont propres, Snm1 (*Suppressor of nuclear mitochondrial endoribonuclease 1*) (Schmitt *et al.*, 1994) et Rmp 1 (*RNase MRP protein 1*) (Salinas *et al.*, 2005). Comme les protéines Rpr2 et Snm1 possèdent toutes deux un motif doigt de zinc, il a été suggéré que ces protéines seraient paralogues (Schmitt *et al.*, 1994; Hartmann *et al.*, 2003). Trois substrats de la RNase MRP ont été bien caractérisés: un ARNr prémature, un ARNm et des ARN mitochondriaux.

Chez l'homme, la composante ARN est constituée par l'ARN 7-2 (Hashimoto *et al.*, 1983) et coprécipite avec au moins 7 sous-unités protéiques de la RNase P (Rpp20, Rpp25, Rpp29, Rpp30, Rpp38, hPop5 et hPop1) suggérant leur appartenance à la composante protéique de la RNase MRP (Stolc *et al.*, 1997; Pluk *et al.*, 1999; van Eenennaam *et al.*, 1999; van Eenennaam *et al.*, 2001; Welting *et al.*, 2004). De plus, des données soutiennent l'existence de différents ensembles d'enzyme avec des compositions protéiques différentes (Welting *et al.*, 2006). Cependant, la composition protéique complète de la RNase MRP humaine n'a pas encore été déterminée par une purification biochimique. Il est particulièrement intéressant de remarquer qu'aucune protéine homologue à Snm1 et Rmp1 spécifiques de la RNase MRP de levure n'est prédite chez *H. sapiens* (Rosenblad *et al.*, 2006). Par ailleurs, la composante ARN de la RNase MRP humaine a été décrite comme formant un complexe ribonucléoprotéique avec une activité ARN polymérase ARN-dépendante (Maida *et al.*, 2009). Utilisant l'ARN de la RNase MRP comme matrice, ce complexe produit des ARN double brins qui sont coupés par l'enzyme Dicer en petits ARN interférents de 22 nt qui éteignent spécifiquement l'expression du gène de la composante ARN de la RNase MRP. Ce mécanisme suggère une régulation de l'expression de la composante ARN de la RNase MRP selon une boucle de rétroactivation.

La RNase MRP est majoritairement localisée dans le nucléole où son rôle dans la maturation des ARNr a été mis en évidence (Schmitt *et al.*, 1993b; Lygerou *et al.*, 1996a). Comme pour la RNase P, cette localisation correspond au site de transcription de son substrat principal. Bien qu'ayant une localisation similaire et une composition voisine, les RNases P et MRP ont des substrats distincts et les isolations par affinité ne suggèrent pas que ces holoenzymes soient associées l'une avec l'autre de façon stable dans les cellules (Srisawat *et al.*, 2001; Li *et al.*, 2002). Cependant avec une telle similarité entre les composantes ARN et protéiques, les deux holoenzymes ont probablement un mécanisme catalytique apparenté. Les différences dans la spécificité du substrat entre la RNase P et la RNase MRP pourraient provenir des sous-unités protéiques spécifiques à chaque complexe et/ou aux différences entre les structures des composantes ARN.

3.3.1.2. Caractérisation de la composante ARN des RNases MRP

3.3.1.2.1. Structure de la composante ARN

Les composantes ARN de la RNase MRP possèdent de nombreuses caractéristiques structurales communes aux composantes ARN des RNases P bactériennes, archées et eucaryotes. Les composantes ARN des RNases MRP de *S. cerevisiae* et *H. sapiens* sont présentées en **Figure 6**.

La composante ARN possède une partie, appelée domaine 1, ressemblant fortement au domaine C de la composante ARN de la RNase P où les hélices P1, P2, P3, P4 et eP19 (eP15 chez certaines levures) sont structuralement homologues (Forster *et al.*, 1990; Schmitt *et al.*, 1993a; Paluh *et al.*, 1995). Les ARN des RNases MRP ont aussi conservés certaines des régions conservées des ARN de RNase P : CR-I, CR-V et CR-IV (Piccinelli *et al.*, 2005). La conservation du domaine catalytique suggère fortement que dans la RNase MRP, la composante ARN soit aussi celle qui porte le noyau catalytique. Cependant, aucune confirmation expérimentale de l'activité catalytique de la composante ARN des RNase MRP en l'absence de sa composante protéique n'a été rapportée.

De plus, certaines régions semblent avoir plus particulièrement co-évoluées au sein des espèces. Chez les levures, l'hélice P3 de la RNase P peut ainsi remplacer celle de la RNase MRP de la même espèce de manière fonctionnelle, mais ne peut remplacer celle de la même enzyme d'une autre espèce de levure, même apparentée (Lindahl *et al.*, 2000). Il apparait en outre que la séquence de cette hélice P3 soit fortement conservée entre les ARN des RNases P et MRP au sein d'une même espèce et plus variablement lors d'une comparaison inter-espèce. (Perederina *et al.*, 2007). Or, cette hélice P3 avec une structure en hélice-boucle-hélice est spécifique des eucaryotes et est une région importante de liaison pour au moins trois protéines, Pop1, Pop6 et Pop7 (Ziehler *et al.*, 2001; Perederina *et al.*, 2007; Perederina *et al.*, 2010a).

Une autre structure conservée semble avoir co-évoluée : l'hélice eP15 des levures, absente chez les vertébrés (Li *et al.*, 2002; Marquez *et al.*, 2005). Bien qu'aucun rôle de cette hélice n'ait été mis en évidence, il semble que chez les levures, cette structure soit nécessaire à l'assemblage des RNases P et MRP (Li *et al.*, 2004).

Ainsi, la forte similarité entre le domaine 1 de la composante ARN de la RNase MRP et le domaine catalytique de l'ARN de la RNase P suggère que le mécanisme de coupure du substrat soit apparenté dans ces deux holoenzymes mais peu de données existent principalement à cause de la difficulté à étudier des structures eucaryotes aussi complexes.

La composante ARN ne possède pas de domaine homologue au domaine S mais un domaine fortement divergent, appelé domaine 2 qui divise les séquences ARN étudiées en deux groupes : d'une part *S. cerevisiae* et d'autre part *Mus musculus* et *H. sapiens* (Schmitt *et al.*, 1993a; Li *et al.*, 2002). Des études avec des sondes biochimiques ont confirmé la variabilité des séquences secondaires de ce domaine entre les espèces (Topper *et al.*, 1990; Li *et al.*, 2002; Piccinelli *et al.*, 2005; Walker *et al.*, 2005; Woodhams *et al.*, 2007). L'organisation la plus communément mise en évidence se compose de trois hélices P8, P9 et P12 comme chez *H. sapiens* (**Figure 6**). L'hélice P8 contient une séquence particulièrement conservée qui forme une boucle de 4 à 5 résidus (5'-GARA(R)-3'). Les autres hélices sont plus variables comme chez la levure (ymP5, ymP6 et ymP7) et la façon dont elles sont apparentées entre espèces et dont elles interviennent dans l'activité de la RNase MRP reste actuellement mal comprise.

3.3.1.2.2. Mutation de la composante ARN associées à des maladies humaines

Chez l'homme, les mutations de la composante ARN de la RNase MRP ou l'interruption de la transcription de cet ARN à cause de mutations (insertion, duplication) dans la région promotrice du gène *RMRP* (*RNase MRP*) conduisent à une variété de maladies pléiotropiques récessives généralement caractérisées par des désordres du développement. Ces maladies incluent notamment l'hypoplasie des cartilages et des cheveux (OMIM CHH) (Ridanpaa *et al.*, 2001) et la dysplasie anauxétique (Thiel *et al.*, 2005).

Les mutations dans la région promotrice du gène *RMRP* réduisent fortement la transcription du gène *RMRP in vitro* et conduisent à une réduction de l'expression de la RNase MRP chez les patients (Hermanns *et al.*, 2005; Nakashima *et al.*, 2007). Les mutations dans la région promotrice n'ont jamais été trouvées sous une forme homozygote ou associées à une autre mutation dans cette région (Bonafe *et al.*, 2005) suggérant une probable létalité d'un tel génotype, ce qui est cohérent avec le rôle essentiel de la RNase MRP dans les cellules. D'autre part, ces mutations dans la région promotrice sont toujours associées à des mutations dans le gène *RMRP* (Bonafe *et al.*, 2005).

Les mutations dans la composante ARN de la RNase MRP sont retrouvées dans un large éventail de positions sous la forme de substitutions, duplications, insertions et délétions, généralement dans des régions conservées phylogénétiquement (Bonafe *et al.*, 2005; Hermanns *et al.*, 2005; Hirose *et al.*, 2006; Martin *et al.*, 2007). Le type et la sévérité des manifestations cliniques semblent définis par le type et la localisation des mutations dans les deux allèles affectés. Une corrélation entre la localisation des mutations dans la composante ARN de la RNase MRP et le phénotype a été rapporté dans le cas de l'hypoplasie des cartilages et des

cheveux et de la dysplasie anauxétique (Thiel *et al.*, 2007). Des mutations dans la composante ARN ont été rapportées pour interférer lors des interactions entre la composante ARN et les sous-unités de la composante protéique. Ces mutations conduisent à une stabilité réduite de l'ARN et/ou changent l'activité ou la spécificité de l'holoenzyme (Bonafe *et al.*, 2005; Hermanns *et al.*, 2005; Thiel *et al.*, 2005; Nakashima *et al.*, 2007; Welting *et al.*, 2008) conduisant probablement à un retard du cycle cellulaire comme chez la levure (Cai *et al.*, 2002). Cependant les mécanismes spécifiques responsables de ces maladies pléiotropiques causées par des mutations de la RNase MRP chez l'homme ne sont pas actuellement connus.

3.3.1.3. Caractérisation structurale des RNases MRP

La structure des RNases MRP et P a été étudiée par des expériences d'empreinte à partir de particules purifiées chez *S. cerevisiae* (Tranguch *et al.*, 1994; Esakova *et al.*, 2008). Ces études ont montré des motifs similaires de nucléotides exposés aux solvants ou protégés par une protéine dans ces holoenzymes, suggérant des architectures comparables entre le domaine 1 de l'ARN de la RNase MRP et le domaine C de l'ARN de la RNase P. Dans les deux holoenzymes, les hélices P3 et eP15 sont exposées aux solvants tandis que les hélices P1, P2, P4 et eP19 sont protégées par la présence de protéines. Dans le domaine S, les hélices P8 et P9 sont protégées dans la RNase P ce qui est cohérent avec un rôle actif de ces structures mis en évidence par les études de mutagenèse (Xiao *et al.*, 2005). De manière similaire, dans le domaine 2 de l'ARN de RNase MRP, les hélices P8 et ymP5 sont protégées par des protéines ce qui est cohérent avec la conservation des séquences et les études de mutagenèse qui suggèrent un rôle important de ces structures (Shadel *et al.*, 2000; Li *et al.*, 2004; Piccinelli *et al.*, 2005; Esakova *et al.*, 2008).

Chez *S. cerevisiae*, une étude de co-purification de la composante ARN avec les différentes sous-unités protéiques a mis en évidence une liaison entre l'ARN NME1 et les protéines Pop3, Pop4, Pop5, Rmp1 et Snm1 et suggère une préférence pour la liaison à des ARN simple brin (Aspinall *et al.*, 2007). De plus, cette étude a mis en évidence une interaction d'un ARNr prémature avec les protéines Pop4, Pop6, Rmp1 et Snm1.

En ce qui concerne les interactions des sous-unités protéiques entre elles, en raison de la composition similaire des RNases P et MRP, les interactions mis en évidence lors des études de la RNase P sont largement transposables à l'étude de la RNase MRP et sont résumées dans la **Figure 9**. Les interactions les mieux supportées concernent Pop6 avec Pop7 (Perederina *et al.*, 2007; Perederina *et al.*, 2010a) ainsi que les interactions héritées des archées entre Rpp1 et Pop5, Pop4 et Snm1/Rmp1 (Houser-Scott *et al.*, 2002; Aspinall *et al.*, 2007). De nombreuses

autres interactions ont été mises en évidence mais leur nature est variable selon les études (Houser-Scott *et al.*, 2002; Welting *et al.*, 2004; Aspinall *et al.*, 2007). En outre, une étude préliminaire de la stœchiométrie des protéines dans la RNase MRP a mis en évidence une composition protéique qui pourrait être assez complexe (Salinas *et al.*, 2005).

Ces différentes études ont permis de dessiner une ébauche de la structure de la RNase MRP comme de la RNase P eucaryote sans que la différence de spécificité de substrat entre ces deux holoenzymes ne soit mécaniquement et structuralement comprise.

3.3.2. Etude des substrats des RNases MRP

3.3.2.1. Les différents substrats connus des RNases MRP

Actuellement, trois types de substrats de la RNase MRP ont été bien caractérisés : des ARN mitochondriaux amorces de la réplication de l'ADN mitochondrial, les ARNr 5.8S prématures et l'ARNm CLB2.

La RNase MRP fut initialement décrite comme ayant une activité de maturation des ARN mitochondriaux. Elle est en effet capable de couper certains ARN mitochondriaux qui seront alors utilisés comme amorces ARN lors de la réplication de l'ADN mitochondrial (Chang *et al.*, 1987b; Lee *et al.*, 1998) comme la RNase P bactérienne (Potuschak *et al.*, 1993). L'activité de maturation dans les mitochondries est liée à la composante ARN de la RNase MRP mais la RNase MRP mitochondriale n'a pas encore été purifiée de manière homogène. Il est probable que la composition protéique de la RNase MRP mitochondriale diffère de celle de la RNase MRP nucléaire (Lu *et al.*, 2010).

La visualisation de la composante ARN de la RNase MRP a mis en évidence que cet ARN est localisé principalement dans le nucléole et un rôle dans la maturation des ARNr a été décrit chez la levure. A l'exception de l'ARNr 5S, les ARNr eucaryotes sont transcrits par l'ARN polymérase I en tant que précurseurs uniques qui doivent subir de multiples étapes de coupure endo- et exo-nucléolytique pour former des ARNr matures. De nombreux détails de cette maturation sont encore peu connus (Lindahl *et al.*, 2009). Chez *S. cerevisiae*, la RNase MRP est connue pour couper au niveau du site A3 entre les ARNr 18S et 5.8S du précurseur ARN 35S (Schmitt *et al.*, 1993a; Chu *et al.*, 1994; Lygerou *et al.*, 1996a). Cette coupure endonucléolytique conduit à l'obtention d'une extrémité 5' mature de l'ARN 5.8S (Henry *et al.*, 1994). Il existe deux formes d'ARN 5.8S mature (Rubin, 1974) : une forme courte (ARN $5.8S_S$) et une forme plus longue de 7 nt à l'extrémité 5' (ARN $5.8S_L$). La forme courte de la ARN 5.8S représente normalement 80 à 90% des ARN 5.8S chez la levure. La signification biologique de l'existence de ces deux formes d'ARN 5.8S n'est actuellement pas connue mais

la présence des ces deux formes semble être commune chez les eucaryotes (Bowman *et al.*, 1983; Smith *et al.*, 1984; Henry *et al.*, 1994). La diminution de l'expression de la RNase MRP ou la délétion du site A3 dans l'ARN 35S conduit à l'accumulation de la forme longue de l'ARN $5.8S_L$ tandis la quantité d'ARN $5.8S_S$ se trouve très réduite, indiquant l'existence de deux voies alternatives de maturation de l'extrémité 5' des ARN 5.8S : une voie principale impliquant une coupure du site A3 par la RNase MRP conduisant à la production d'ARN $5.8S_S$ et une autre voie indépendante du site A3 conduisant à la production d'ARN $5.8S_L$ (Schmitt *et al.*, 1993b; Henry *et al.*, 1994). Les récentes données obtenues avec des mutants sensibles à la température (Lindahl *et al.*, 2009) ont montré que l'inactivation de la RNase MRP conduit à une sévère réduction de l'abondance des précurseurs intermédiaires dans la voie de maturation canonique des ARNr tandis que la transcription des ARNr précurseurs n'est, quant à elle, pas affectée, ce qui suggère que la RNase MRP joue un rôle majeur dans la maturation de l'ensemble des ARNr en plus de la coupure du site A3 de l'ARN précurseur 35S.

Outre son rôle dans la maturation des ARNr, la RNase MRP joue aussi un rôle dans la régulation du cycle cellulaire chez la levure. Chez *S. cerevisiae*, des mutations de la protéine Snm1 conduisent à une mauvaise ségrégation des plasmides (Cai *et al.*, 1999) et causent un retard du cycle cellulaire à la fin de la mitose accompagné d'un accroissement de la quantité de cycline B2 (Clb2) résultant d'une augmentation de la concentration d'ARNm CLB2 (Cai *et al.*, 2002). Il a été démontré que la RNase MRP coupe spécifiquement l'extrémité 5' non traduite (5'UTR) de l'ARNm CLB2 ce qui permet sa dégradation rapide par l'exoribonucléase Xrn1 (Gill *et al.*, 2004). La diminution de la protéine Clb2 est une étape requise pour la progression du cycle cellulaire de l'anaphase à la phase G1 (Charles *et al.*, 1998). En outre, chez la levure, la RNase MRP a été localisée dans le nucléole après l'initiation de la mitose et sous la forme d'un point unique dans le cytoplasme de la cellule fille qui disparait après la télophase (Gill *et al.*, 2006). Cet élément cytoplasmique est appelé corps MRP asymétrique temporel (TAM) et semble être une version distincte des corps cytoplasmiques appelés P-bodies lesquels seraient des sites privilégiés de la machinerie de dégradation des ARNm (Parker *et al.*, 2007). De plus, il a été mis en évidence que l'exonucléase Xrn1 colocalise avec les corps TAM (Gill *et al.*, 2006).

Récemment, un quatrième substrat de la RNase MRP a été mis en évidence chez *S. cerevisiae* et *Nicotiana benthamiana* : l'ARN du virus du "rabougrissement buissonneux" de la tomate (TBSV). En effet, dans les cellules infectées de levure, cet ARN est soumis à un processus de coupure par des endoribonucléases pour obtenir un ARN avec une extrémité 5' mature soit utilisé pour la recombinaison virale soit dégradé par l'exonucléase Xrn1. Or, la

diminution de l'expression de Snm1 ou une mutation sensible à la température dans l'ARN NME1 conduisent à une réduction de la quantité d'ARN TBSV coupés, dans les souches dépourvues d'enzyme Xrn1. De plus, il a été mis en évidence que la RNase MRP de levure est capable de couper cet ARN *in vitro* conduisant à la formation de produits de dégradation de l'ARN TBSV de taille similaire à ceux mis en évidence *in vivo*. L'inactivation du gène homologue à NME1 chez *N. benthamiana* avec des ARN interférents introduits par infection virale conduit également à la diminution de la quantité d'ARN TBSV matures suggérant que chez les plantes, la RNase MRP soit aussi impliquée dans la dégradation de l'ARN TBSV (Jaag *et al.*, 2011).

En raison du degré d'implication de la RNase MRP dans la maturation des ARNr, qui n'est pas encore complètement comprise et de sa présence transitoire dans le cytoplasme, il est attendu que d'autres substrats pour la RNase MRP soient identifiés dans le future. D'ailleurs, très récemment, quatre nouveaux substrats de la RNase MRP ont été identifiés chez la levure lors d'une étude d'identification des ARN associés à la RNase MRP (Aulds *et al*, 2012) : l'ARNm CTS1 codant une endochitinase requise lors de la cytodiérèse, l'ARNm SIC1, homologue chez la levure de l'ARNm mammalien p27 (Kip1) codant un inhibiteur des couples cycline/CDK, le composant ARN U2 du spliceosome et l'ARN transposon TY1.

3.3.2.2. Particularité dans la reconnaissance du substrat

Bien que la RNase MRP soit de composition voisine à celle de la RNase P, la reconnaissance du substrat par la RNase MRP semble cependant différer de la reconnaissance des ARNt prématures par la RNase P. Celle-ci ne reconnait pas de séquences spécifiques dans ses substrats mais utilise plutôt des éléments structuraux conservés du substrat (Esakova *et al.*, 2010) notamment les boucles T et D des ARNt prématures (Torres-Larios *et al.*, 2005; Torres-Larios *et al.*, 2006; Reiter *et al.*, 2010) éloignées du site de coupure. C'est pourquoi le domaine S, responsable de la reconnaissance du substrat sur la composante ARN, est aussi éloigné du site catalytique (Reiter *et al.*, 2010). Au contraire, une étude de sélection de substrats *in vitro* de la RNase MRP chez *S. cerevisiae* a récemment mis en évidence que la RNase MRP coupe les ARN simple brin de manière dépendante aux séquences au voisinage immédiat du site coupure, requérant une cytosine en position +4 par rapport au site de coupure (Esakova *et al.*, 2011). Ces résultats suggèrent que les éléments structuraux de la RNase MRP responsables de l'identification du substrat (certainement dans le domaine 2) seraient localisés dans le voisinage immédiat du site de coupure. Cela impliquerait un repliement particulier de l'ARN du domaine 2 vers le site catalytique dans la composante ARN et formerait ainsi une structure plus

compacte que dans la RNase P. Or, le seul élément phylogénétiquement conservé dans le domaine 2 de la RNase MRP, la séquence GARAR de l'hélice P8, est positionnée suffisamment proche du site catalytique supposé pour remplir ce rôle d'élément se repliant vers le domaine catalytique et interagissant avec le substrat. Des études complémentaires devront dans l'avenir tenter d'explorer cette voie pour appréhender les différences mécanistiques entre les RNases MRP et P.

4. Description du sujet de recherche

Les protéines composent la plupart de la masse d'une cellule. Cependant, au sein d'une cellule, les protéines n'agissent pas seules : il a été mis en évidence que la plupart des fonctions cellulaires dans la cellule sont réalisées par des complexes macromoléculaires contenant de multiples sous-unités avec des fonctions spécifiques (Alberts, 1998). Chacun de ces assemblages protéiques interagit avec plusieurs autres complexes de protéines. Les analyses génomiques de complexes multi-protéiques ont révélé que plus de 80 % des protéines de levure sont assemblées en complexes qui sont hautement connectés au travers du partage de certains de leurs composants (Gavin *et al.*, 2002) comme c'est le cas de la RNase P et de la RNase MRP. Ces complexes sont souvent décrits comme "des machines moléculaires", appellation qui prends en compte l'ensemble des aspects caractérisant ces complexes biologiques : modularité, complexité, fonction cyclique et souvent consommation d'énergie (Nogales *et al.*, 2001). La machinerie transcriptionnelle, le spliceosome, le ribosome et la RNase P en sont des exemples. Pour comprendre les mécanismes d'une machine macromoléculaire tels que l'assemblage du complexe ou l'orchestration de ses activités, il est nécessaire de déterminer sa structure tridimensionnelle.

La cristallographie aux rayons X a été très utilisée pour résoudre la structure de protéines seules et de domaines fonctionnels. Depuis quelques années, l'accent a été mis sur l'automatisation des techniques pour réaliser à haut débit l'expression des protéines et les essais de cristallisation. Néanmoins, la taille des complexes macromoléculaires rends souvent inaccessible les études structurales par la cristallographie aux rayons X à cause de la difficulté à obtenir des cristaux exploitables. Il existe quelques exceptions dont le ribosome (Ban *et al*, 2000; Yusupov *et al*, 2001, Ben-Shem *et al*, 2010), l'ARN polymérase II chez la levure (Cramer *et al.*, 2001) et la F-ATPase (Abrahams *et al.*, 1994). En outre, dans la plupart des cas, ces complexes sont peu abondants dans les cellules, par exemple environ 400 particules de RNase P chez la levure (Srisawat *et al.*, 2001). La purification ne fournit alors qu'une faible quantité de complexe d'une homogénéité variable qui est le plus souvent insuffisante pour les tests de cristallisation. La reconstitution *in vitro* de grands complexes à partir de protéines recombinantes crée d'autres difficultés telles que le maintien correct de la stœchiométrie et l'ordre de l'assemblage. De plus, la composition exacte d'un complexe n'est souvent pas connue. Ainsi dans la communauté scientifique, de nombreux efforts se sont portés sur la détermination de la structure des sous-unités individuelles et de domaines ou de sous-complexes.

La microscopie électronique à transmission s'est révélée être une excellente alternative pour l'étude de ces macromolécules. A l'inverse de la cristallographie aux rayons X, les techniques d'analyse des particules uniques (*single particle*) ne nécessitent pas la formation de cristaux. En outre, seule une petite quantité d'échantillon (de l'ordre de quelques pico moles) est suffisante pour l'analyse rendant ces méthodes adaptées à l'étude des complexes biologiques. En contrepartie, la résolution des structures obtenues est bien moindre que celle des structures cristallographiques (10-15 Å versus 1-2 Å). A cette résolution, la forme générale d'un complexe peut néanmoins être déterminée et les différents domaines sont visibles. Les structures à haute résolution des sous-unités peuvent alors être utilisées pour créer un modèle pseudo-atomique. La technique la plus prometteuse pour la caractérisation structurale de complexes macromoléculaires est donc une approche hybride combinant la microscopie électronique à transmission et le traitement des images avec des méthodes permettant de résoudre la structure à haute résolution de protéines ou de sous-complexes comme la cristallographie aux rayons X et la RMN.

Dans cette étude, la RNase P et la RNase MRP ont été plus particulièrement étudiées chez la levure *S. cerevisiae.* Nous avons développé une méthode de purification spécifique pour séparer chacun des complexes de façon à obtenir des particules actives et homogènes. Puis, nous avons utilisé une méthode combinant la microscopie électronique, le traitement d'image, le marquage par anticorps des sous-unités et l'utilisation des structures à haute résolution pour caractériser la forme et l'organisation structurale des RNases P et MRP. En ce qui concerne l'organisation macromoléculaire et la biocatalyse, la RNase P illustre un rare exemple soulignant les différences frappantes au niveau (1) des compositions des sous-unités de ses variantes naturelles dans les trois domaines du vivant et (2) de ses nombreux substrats qui varient en séquence et en structure. La résolution d'une structure pseudo-atomique de la RNase P nous a permis ensuite par comparaison avec la structure de la RNase MRP de discuter de la coordination et la spécificité du substrat dans ces deux complexes.

MATERIELS ET METHODES

1. Description des souches de levure *S. saccharomyces*

Pour créer les protéines de fusion Rpr2-HA (étiquette 3HA à l'extrémité C-terminale) et HA-Snm1 (étiquette 3HA à l'extrémité N-terminale), des intégrations génomiques de l'étiquette 3HA (associée au marqueur HIS3MX6 (Longtine *et al.*, 1998)) sont réalisées par recombinaison homologue dans la souche YSW1 (MATa POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1::LEU2 sep1::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112) (Gill *et al.*, 2004) sur le gène *YIR015w* pour Rpr2 et le gène *YDR478w* pour Snm1 selon la méthode adaptée de LiAc/SS DNA carrier/PEG (Schiestl *et al.*, 1989). Les transformants sont sélectionnés sur milieu dépourvu d'histidine. Pour identifier les souches modifiées, la localisation de l'intégration de l'étiquette est vérifiée par PCR et la présence de l'étiquette HA associée aux protéines recombinantes est recherchée par Western-blot avec un anticorps anti-HA (Covance) sur un extrait cellulaire total.

Dans la suite de ce manuscrit, la souche YSW1 dont la protéine Rpr2 porte l'étiquette HA sera nommée YSW1-Rpr2-HA tandis que celle dont la protéine Snm1 porte l'étiquette HA sera nommée YSW1-HA-Snm1.

Pour créer des protéines de fusion avec l'étiquette GFP dans la souche YSW1-Rpr2-HA, nous avons intégrés par recombinaison homologue, une cassette yeGFP::*hph*NT1 à l'extrémité C-terminale de Pop5, Pop6, Pop8 et Rpp1, et une cassette natNT2::PGal1::yeGFP à l'extrémité N-terminale de Rpr2-HA et Pop1 (Janke *et al.*, 2004). Les transformants sont sélectionnés en présence d'hygromycine B à 300 mg/l (gène *hphNT1*) ou nourséothricine à 100 mg/l (gène *natNT2*). La présence de l'étiquette GFP associées aux protéines recombinantes est confirmée sur extrait cellulaire total par Western-blot avec un anticorps anti-GFP (Roche).

La liste des souches créées pour cette étude se trouve en annexe I.

1.1. Préparation des cassettes à intégrer

Chaque amorce se compose de 50 à 55 nucléotides homologues à la séquence en 5' (amorce sens) ou en 3' (amorce anti-sens) du point d'insertion (codon ATG ou codon stop) sur le gène concerné et d'environ 20 nucléotides spécifiques de la cassette à amplifier. La liste des amorces utilisées dans cette étude se trouve en annexe II.

Les cassettes 3HA::His3MX6 (1742 pb) et His3MX6::PGal1::HA (2045 pb) sont amplifiées à partir respectivement des plasmides pFA6a-3HA-His3MX6 et pFA6a-His3MX6-PGAL1-3HA dans les conditions suivantes : 250 ng d'ADN plasmidique sont mis en présence

de 1 µM d'amorces sens et anti-sens, de 0.2 mM de dNTPs, 1.5 mM MgCl2 et 2.6 U d'un mélange d'ADN polymérases du kit Expand High Fidelity PCR system (Roche) dilué dans du tampon 2 dans un volume final de 100 µl. Après une étape de dénaturation de 5 minutes à 94°C, l'ADN est amplifié par 35 cycles d'une minute de dénaturation à 94°C, une minute d'hybridation à 55°C, 3 minutes d'élongation à 68°C. Une étape d'élongation de 10 minutes à 68°C termine la réaction.

Les cassettes yeGFP::*hph*NT1 (2550 pb) et natNT2::PGal1::yeGFP (2578 pb) sont amplifiées à partir respectivement des plasmides pYM25 et pYM-N25 dans les conditions suivantes : 200 ng d'ADN plasmidique sont mis en présence de 0.64 µM d'amorces sens et anti-sens, de 0.35 mM de dNTPs, 1.5 mM MgCl2, de 2% de DMSO et 4.9 U d'un mélange d'ADN polymérases du kit Expand High Fidelity PCR system (Roche) dilué dans le tampon 2 dans un volume final de 100 µl. Après une étape de dénaturation de 5 minutes à 94°C, l'ADN est amplifié par 35 cycles d'une minute de dénaturation à 94°C, 30 secondes d'hybridation à 54°C, 3 minutes d'élongation à 68°C. Une étape d'élongation de 10 minutes à 68°C termine la réaction.

L'ADN amplifié est ensuite visualisé sous rayonnements UV sur un gel d'agarose à 1% coloré avec du bromure d'éthidium (0.25 µg/ml) après électrophorèse dans du tampon TBE (89 mM Tris, 89 mM acide borique; 2 mM EDTA) pendant 30 minutes à 120 volt. Les produits d'amplification, issus de 3 réactions sont mélangés et purifiés sur colonnes Qiagen puis élués dans 30 µl d'eau ultrapure à pH 8.

1.2. Transformation par la méthode LiAc/SS DNA/PEG

A partir d'une préculture sur la nuit, la souche d'intérêt est cultivée dans 50 ml de milieu YPD (*Yeast extract Peptone Dextrose*) jusqu'à une densité optique (DO) entre 0.6 et 1.0 à 30°C sous agitation. Les cellules sont collectées par centrifugation à 3000 g pendant 10 minutes. Elles sont ensuite lavées une première fois avec 25 ml d'eau ultrapure puis avec 1 ml de 100 mM d'acétate de lithium / 10 mM Tris-HCl pH 7.4 et sont enfin remises en suspension avec ce dernier tampon dans un volume final de 0.5 ml.

100 µl de suspension cellulaire sont alors mis en présence de 270 µl de 50 % PEG4000 / 100 mM d'acétate de lithium / 10 mM Tris-HCl pH 7.4, 10 µl d'ADN de sperme de saumon à 10 mg/ml préalablement dénaturé pendant 5 minutes à 95°C, de 13 µl de cassette amplifiée et de 30 µl d'eau ultrapure. Après mélange, la suspension cellulaire est incubée pendant 25-30 minutes à température ambiante sous agitation douce. Un choc thermique est ensuite réalisé par incubation à 42°C pendant 20 minutes. Puis, les cellules culotées sont remises en

suspension dans 500 µl de milieu YPD et incubées à 30°C sous agitation douce pendant 0 à 3 heures selon les marqueurs de sélections, avant d'être étalées sur un milieu gélosé sélectif. Il est incubé à 30°C jusqu'à l'apparition des transformants (2 à 4 jours).

1.3. Test de l'intégration génomique d'une étiquette à la localisation souhaitée

A partir de colonies fraîchement étalées sur milieu gélosé sélectif, les levures sont mises en solution dans 34 µl d'eau ultrapure et bouillies pendant 15 minutes à 95°C. Une réaction de PCR est ensuite réalisée sur cet extrait total par ajout de 5.8 µl de mélange réactionnel (1 µM d'amorces sens et anti-sens, 0.2 mM de dNTPs, 1.5 mM MgCl2 et 2.6 U d'un mélange d'ADN polymérases du kit Expand High Fidelity PCR system (Roche) dilué dans du tampon 2). L'ADN est amplifié par 35 cycles de 30 secondes de dénaturation à 94°C, 30 secondes d'hybridation à 52°C (Rpr2) ou 55°C (Snm1), 1 minute d'élongation à 68°C. Une étape d'élongation de 10 minutes à 68°C termine la réaction.

L'ADN est ensuite visualisé après électrophorèse sous rayonnements UV sur un gel d'agarose à 2%.

Nom de l'amorce	Séquence	Taille de l'amplicon
End 5' to 3' Rpr2 (sens)	CCAAATATCCGCTTACCAAAC	403 pb
3' to5' HA (anti-sens)	GAACGTCATATGGATAGGATCC	
5' to 3' HA (sens)	CCCATACGATGTTCCTGACTATG	403 bp
Start 3' to 5' Snm1 (anti-sens)	GGTCCTTCCATACTCTCGCCGC	

1.4. Test de l'expression des protéines recombinantes

A partir d'une culture sur la nuit de la souche d'intérêt en milieu liquide non sélectif (milieu YPD ou YPG), un volume de culture correspondant à 20 DO de cellules est centrifugé pendant 3 minutes à 3000 g. Les cellules sont ensuite lavées avec de l'eau ultrapure. Puis, 200 µl de billes de verre et 200 µl de tampon de charge (62 mM Tris-HCl pH6.8, 2.3% SDS, 10% glycérol, 0.05% bleu de bromophénol, 25 mM DTT) sont mis en présence des cellules. Une lyse mécanique est réalisée en chauffant le mélange pendant 5 minutes à 95°C puis en agitant fortement au vortex pendant 1 minute. Cette étape est répétée une fois.

L'extrait cellulaire est alors chargé sur un gel de polyacrylamide 10% pour réaliser une SDS-PAGE dans du tampon Tris-Glycine (25 mM Tris, 192 mM glycine, 0.1% SDS) en

conditions dénaturantes (Towbin *et al.*, 1979) pendant 45 minutes à 180 volt. Après transfert sur une membrane de nitrocellulose, les protéines sont révélées par Western-blot.

2. Purification des complexes RNases P et MRP

2.1. Purification des complexes natifs

Les cellules sont cultivées dans 12 L de milieu YPD (souche YSWI-Rpr2-HA) ou YPG (souche YSWI-HA-Snm1) et collectées à une DO entre 3.0 et 3.5 par centrifugation pendant 5 min à 5000 g. Puis, elles sont lysées mécaniquement avec des billes de verres (0.40/0.60 mm, Sartorius) dans du tampon de lyse (50 mM Tris HCl pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, protease inhibitor FY (Serva), 0.0375% NP40) en utilisant une Pulverisette 5 (Fritsch, 3 cycles de broyage, puissance 350 tr/min pendant 4 min, 1 min d'arrêt entre les cycles). Les extraits sont ensuite centrifugés à 3220 g pendant 10 minutes pour retirer les billes de verre puis à 50 000 g pendant 60 minutes pour retirer les débris cellulaires.

Le lysat est alors appliqué sur une matrice IgG *Sepharose 6 Fast Flow slurry* (Amersham, 0.4 ml/L de culture cellulaire) pendant 90 minutes à 4°C sous agitation douce. Puis, la matrice est lavée avec du tampon de lyse contenant 0.5 mM DTT pendant 15-20 minutes à 4°C sous agitation douce. Les protéines sont éluées par coupure de l'étiquette TAP sur la nuit à 4°C avec 10 µM de protéase TEV sous agitation douce, l'éluat étant collecté par gravité. Cet éluat est alors appliqué sur matrice d'anticorps anti-HA (HA affinity matrix, Roche) pendant 5 heures à 4°C sous agitation douce. La fraction liée est éluée par compétition avec un peptide HA à 250 mg/L (Roche) par incubation pendant 30 minutes à 30°C. L'éluat est concentré dans un concentrateur Millipore (membrane Biomax avec une limite d'exclusion de 30 kDa) par centrifugation pendant 5 min à 9279 g à 4°C jusqu'à ce que l'échantillon soit contenu dans 50 µl.

Après concentration, les protéines purifiées sont séparés par SDS-PAGE sur un gel avec un gradient de polyacrylamide 4-12% NuPAGE (Invitrogen) dans du tampon MOPS par électrophorèse pendant 45 min à 180 volt. Les protéines sont révélées soit par coloration au bleu de Coomassie colloïdale soit par Western-blot avec anticorps anti-HA.

2.2. Purification de complexes marqués avec un anticorps

Pour la purification de complexes marqués avec un anticorps, l'éluat issu de la coupure par la protéase TEV est partagé en 2 parties, dont une suit la procédure décrite ci-dessus.

La seconde partie de l'éluat est mise en présence de la matrice anti-HA pendant une heure. Puis, des anticorps monoclonaux anti-GFP (Roche, 40 µg/ml concentration finale) sont

ajoutés et laissés incuber pendant 4 heures. L'élution est réalisée dans les conditions décrites ci-dessus.

3. Méthodes de caractérisations biochimiques

3.1. Western-blot

Après SDS-PAGE, les protéines sont transférées sur une membrane de nitrocellulose avec un *Trans-Blot Semi-dry Electrophoretic Transfer Cell* (BioRad) par application d'un courant de 10 volt pendant 30 minutes dans du tampon de transfert (48 mM Tris, 39 mM glycine pH 9.2, 20% méthanol, 0.0375% SDS) (Bjerrum, 1986). Après blocage des sites libres de la membrane de nitrocellulose avec une solution de PBS + 5% de lait, une solution d'anticorps anti-HA (anticorps monoclonal de souris HA.11 MMS-101R, Covance, dilué au 1:1000) ou anti-GFP (anticorps monoclonal de souris, Roche, dilué au 1:1000) est incubée pendant une heure à température ambiante. Après lavage dans une solution de PBS + 5% lait, une solution avec un anticorps anti-IgG de souris couplé à la peroxydase (anticorps de Chèvre, SIGMA, dilué au 1:5000) est appliquée pendant 45-60 minutes. Après lavage dans du PBS, une solution de chimioluminescence *Western Lightning® Western Blot Chemiluminescence Reagent Plus* (Perkin Elmer) permet de révéler la présence des anticorps sur film autoradiographique (ECL Amersham).

3.2. Spectrométrie de masse

Après séparation par SDS PAGE sur un gel 4-12% NuPAGE (Invitrogen), les protéines sont colorées au bleu de Coomassie colloïdale. Les bandes contenant les protéines sont découpées et digérées à la trypsine (Shevchenko *et al.*, 1996). Le mélange de peptides obtenu est analysé par spectrométrie de masse en tandem. Les protéines sont identifiées grâce au logiciel Mascot (Matrix Science) et à la banque de données MSDB.

3.3. RT-PCR quantitative

Les éluats HA issus des souches YSW1-Rpr2-HA et YSW1-HA-Snm1 sont traités avec la protéinase K à 10 µg/ml pendant 30 minutes à 37°C pour digérer les protéines puis à la DNase I pendant 20 minutes à 37°C pour digérer l'ADN résiduel contaminant. Les noyaux ARN sont alors purifiés par une extraction classique au phénol/chloroforme (Sambrook, 2001).

Les noyaux ARN NME1 (RNase MRP) et RPR1 (RNase P) sont reverse-transcrits en utilisant le kit *Transcriptor First Strand cDNA Synthesis* (Roche) pendant 30 minutes à 60°C,

pour améliorer l'efficacité de la réaction en déployant l'ARN. Les amorces utilisées sont 5'-TCCATTGGGTTACTCGATCC-3' pour l'ARN NME1 et 5'-TGGAACAGCAGCAGTAAT CG-3' pour l'ARN RPR1.

La PCR quantitative (qPCR) est réalisée en utilisant le mélange SYBR®GREEN PCR Master Mix (Applied Biosystem) selon les instructions du fabriquant. Les réactions sont réalisées en triplicatas à partir de 3 différentes purifications pour chacune des souches en utilisant un thermocycleur ABI 7500 dans les conditions suivantes: 50°C pendant 2 minutes, dénaturation à 95°C pendant 10 minutes, puis 40 cycles d'amplification (95°C pendant 15 s et 60°C pendant 1 minute). Une étape de dissociation est réalisée pour vérifier la spécificité des réactions. Les paires d'amorces (5'-CGGTACTGGATTCCGTTTGT-3' and 5'-AGCA CGGGAAAGAGCAATC-3' pour l'ADNc NME1 et 5'-CCGGGTTAATGTCGCTTTT-3' et 5'-CGGTCGGTAAAGACTGGTTC-3' pour l'ADNc RPR1) ont été définies par le logiciel Primer3. Toutes les amorces ont été choisies en prenant soin qu'elles ne s'hybrident pas dans les régions avec les structures secondaires des molécules ciblées, prédites par le logiciel Mfold (Zuker *et al.*, 1999).

La quantité relative d'un complexe par rapport à l'autre est déterminée en calculant le rapport entre la quantité relative d'ARN du complexe contaminant et celle du complexe attendu dans l'échantillon lors de la purification selon la formule :

$$\textbf{[complexe contaminant]/[complexe majoritaire]} = 2^{(\text{Ct complexe contaminant} - \text{Ct complexe majoritaire})} = 2^{\Delta Ct}$$

où Ct correspond à la valeur du cycle seuil (*threshold cycle*) obtenu par qPCR.

3.4. Synthèse *in vitro* des pré-ARNt

L'ADN génomique de *S. cerevisiae* est utilisé comme matrice initiale pour amplifier le gène de l'ARNt^{Tyr} (GenBank J01380) avec les amorces ptRNA-Tyr_F (5'-GTAATACGACTCACTATAGGACAAAAACGACTCTCGGTAG-3') et ptRNA-Tyr_R (5'-AAAAATCTCCCGGGGGCGA-3') précédemment décrites (Ziehler *et al.*, 2000) et le gène de l'ARNt^{Glu} (tE(CUC)D GenBank X06132.1) avec les amorces ptRNA-Glu_F (5'-GTAATACGACTCACTATAGAGGTCAAACTATCACCAAGGTACA-3') et ptRNA-Glu_R (5'-AAAAAGAAACTCCGAAGCGG-3'). Les réactions d'amplification sont réalisées avec 100 ng d'ADN génomique en présence de 1 µM d'amorces sens et anti-sens, 0.2 mM de dNTPs, 1.5 mM MgCl2 et 2.6 U d'un mélange d'ADN polymérases du kit Expand High Fidelity PCR system (Roche) dilué dans du tampon 2. L'ADN est amplifié par 30 cycles de 30 secondes de dénaturation à 94°C, 30 secondes d'hybridation à 56°C, 1 minute d'élongation à 68°C. Une étape d'élongation de 10 minutes à 68°C termine la réaction. Les

amplicons de 127 pb sont visualisés sous rayonnements UV après électrophorèse sur un gel d'agarose à 4 %. Les produits d'amplification, issus de trois réactions, sont mélangés et purifiés sur colonnes Qiagen puis élués dans 30 µl d'eau ultrapure à pH 8.

Grâce à la présence du promoteur de la T7 RNA polymérase sur l'amorce sens, les produits d'amplification génèrent une matrice ADN permettant de réaliser une transcription *in vitro* en utilisant le kit *Transcript Aid T7 High Yield Transcription* (Fermentas) pour synthétiser les pré-ARNt^{Tyr} et pré-ARNt^{Glu}. 2 µg de produit d'amplification purifiés sont incubés en présence de 4 µl de tampon *Transcriptor* 5X, 2 µl du mélange d'enzyme *Transcript Aid*, de 8 µl de rNTP à 10 mM dans un volume final de 24 µl pendant 8 heures à 37°C. Puis, le mélange est traité à la DNase I pendant 15 minutes à 4°C. Les pré-ARNt sont enfin purifiés par extraction phénol/chloroforme (Sambrook, 2001).

3.5. Test de l'activité endonucléolytique *in vitro* de la RNase P purifiée

6 µg de pré-ARNt sont incubés avec 3 µg de RNase P purifiée (ratio molaire approximatif 20:1) dans le tampon PA modifié (50 mM Tris HCl pH 8.0, 100 mM NH_4Cl, 10 mM $MgCl_2$ ou 2 mM EDTA) pendant 30 minutes à 37°C (Yang *et al.*, 2007). Les réactions sont arrêtées par ajout d'une solution à 9 M urée / 1 mM EDTA / 1 mg/ml bleu de bromophénol / tampon TBE 1X. Les mélanges sont dénaturés une minute à 95°C et séparés sur un gel de 12% polyacrylamide / 7M urée. Le gel est enfin coloré avec une solution de bleu de toluidine 0.1%.

3.6. Test de la capacité de liaison des pré-ARNt à la RNase P purifiée

Une gamme de 0 à 13.5 µg de pré-ARNt est incubée avec 3 µg de RNase P purifiée dans le tampon PA modifié avec 10 mM $CaCl_2$ (Yang *et al.*, 2007) pendant 30 minutes à température ambiante. Les réactions sont arrêtées par ajout de tampon de charge (50 mM Tris HCl pH 8.0, 100 mM NH_4Cl, 10 mM $CaCl_2$, 1 mg/ml bleu de bromophénol, 10% glycérol). Les mélanges sont séparés sur un gel de 4.5% polyacrylamide dans du tampon TBE par électrophorèse pendant 40 minutes à 150 volt à 4°C. Le gel est ensuite coloré avec du *SYBR® Green II RNA Gel Stain* pendant 15 minutes pour visualiser les ARN puis une coloration à l'argent est réalisée pour visualiser les complexes ribonucléoprotéiques.

4. Microscopie électronique à transmission

4.1. Préparation des échantillons

Les grilles sont préparées selon la méthode de coloration négative en sandwich (Golas *et al.*, 2003). Les échantillons concentrés sont mis en présence d'acétate d'uranyle 2% entre 2 films de carbone fins (2-3 nm) dont un recouvre une grille de cuivre.

Pour la reconstruction conique aléatoire, des paires de champs avec et sans inclinaison (-55°et 0°) sont acquises sur un microscope électronique à transmission Philips CM200 FEG avec une caméra CCD 2kx2k (TVIPS GmbH) à grossissement nominal de 27 500 (5.2Å/pixel) à température ambiante. Les micrographes sont acquis sous faible dose d'électrons ($10e^-$/Å) et l'alignement des aires d'intérêt a été fait manuellement.

Pour l'analyse des images à haute résolution, les grilles sont préparées comme décrites ci-dessus et congelées dans de l'azote liquide immédiatement après l'étape finale de séchage. Les grilles sont ensuite conservées dans de l'azote liquide.

Les grilles sont transférés dans le microscope avec un support Gatan 626, adapté à la conservation des échantillons à -96°C. Les micrographes sont acquis sur la caméra à grossissement nominal de 66 000 (2.2Å /pixel) à 200 kV à faible dose d'électron. 2 006 micrographes ont été enregistrés pour la RNase MRP et 1 861 micrographes pour la RNase P.

4.2. Traitement des images des complexes RNase P et MRP

Les cartes initiales de la RNase P et de la RNase MRP ont été calculées par reconstruction conique aléatoire en utilisant le logiciel Spider (Frank *et al.*, 1996) pour la sélection des particules et le fenêtrage et le logiciel Imagic 5 (van Heel *et al.*, 1996) pour l'alignement, la classification et la reconstruction.

Ces cartes affinées des RNases P et MRP ont été utilisées pour générer les références initiales pour l'affinement itératif par appariement de projections avec Spider d'un plus grand ensemble de données de particules colorées négativement et congelés de RNase P (42 009 images de particules) et de RNase MRP (59 400 images de particules).

A ce stade, une correction de la fonction de transfert de contraste a été appliquée. De plus, pour chaque complexe, les données ont été traitées dans 20 groupes de même focale ; pour chaque groupe, une carte a été calculée par rétroprojection pondérée en utilisant Spider. Les cartes ont été additionnées pour donner la carte finale.

En outre, l'hétérogénéité des populations de particules de RNase P et RNase MRP a été estimée par les deux approches suivantes : 1) en utilisant Xmipp Maximum-likelihood 3D-

sorting (Scheres *et al.*, 2007) pour classer les particules soit de RNase P soit de RNase MRP en 5 classes. 2) en utilisant IMAGIC 4D, pour générer 10 cartes à partir des particules aléatoirement assignées à 10 groupes puis affinées itérativement par appariement de projections et classification supervisée.

La résolution des cartes a été estimée par corrélation des enveloppes de Fourier (FSC, (Harauz *et al.*, 1986)).

4.3. Traitement des images des complexes RNases P marqués avec un anticorps

Pour l'analyse d'images dans les expériences de marquage par anticorps, les images des particules ont été sélectionnées en utilisant Boxer, une partie du logiciel EMAN (Ludtke *et al.*, 1999). Les alignements et la classification ont été réalisés avec IMAGIC5 (van Heel *et al.*, 1996) et ont été répétés jusqu'à ce que les classes restent stables. Pour localiser les densités additionnelles à la périphérie des particules résultant des anticorps liés, les moyennes des classes ont été sous-classifiées en utilisant un masque se focalisant sur la périphérie des particules.

4.4. Positionnement des modèles atomiques dans la carte de microscopie électronique

Le logiciel Chimera (Pettersen *et al.*, 2004) a été utilisé pour représenter graphiquement les résultats et pour positionner les modèles atomiques dans la carte. Certaines structures (le sous-complexe Pop6-Pop7-P3 (Perederina *et al.*, 2010a) et le sous-complexe aPop5-aRpp1 2:2 d'archée (Kawano *et al.*, 2006)) ont été placées automatiquement sur la carte segmentée en domaines grâce à l'option Segger de Chimera. Les autres composants (le modèle de l'ARN de *S. pombe* (Marquez *et al.*, 2006), le sous-complexe aRpr2-aPop4 d'archée (Honda *et al.*, 2008) et la protéine L7Ae, homologue archée de Pop3 (Hamma *et al.*, 2004)) ont été placés manuellement en tenant compte des études d'interactions publiées (Jiang *et al.*, 2001a; Jiang *et al.*, 2001b; Houser-Scott *et al.*, 2002; Welting *et al.*, 2004; Aspinall *et al.*, 2007) et du résultat du marquage par les anticorps des différentes sous-unités protéiques. Les positionnements manuels ont été optimisés avec l'option "*fit to map*" de Chimera, qui réalise une optimisation locale.

Le positionnement de ces structures a été validé par le calcul des corrélations croisées locales entre la carte et les structures positionnées. Les corrélations croisées locales ont été calculées sur une distance de 1.1 nm (5 pixel) du point d'intérêt. Les modèles positionnés ont

été convertis en cartes de densité et filtré à des fréquences spatiales de 1/1.7 nm^{-1} (RNase P) et 1/1.5 nm^{-1} (RNase MRP).

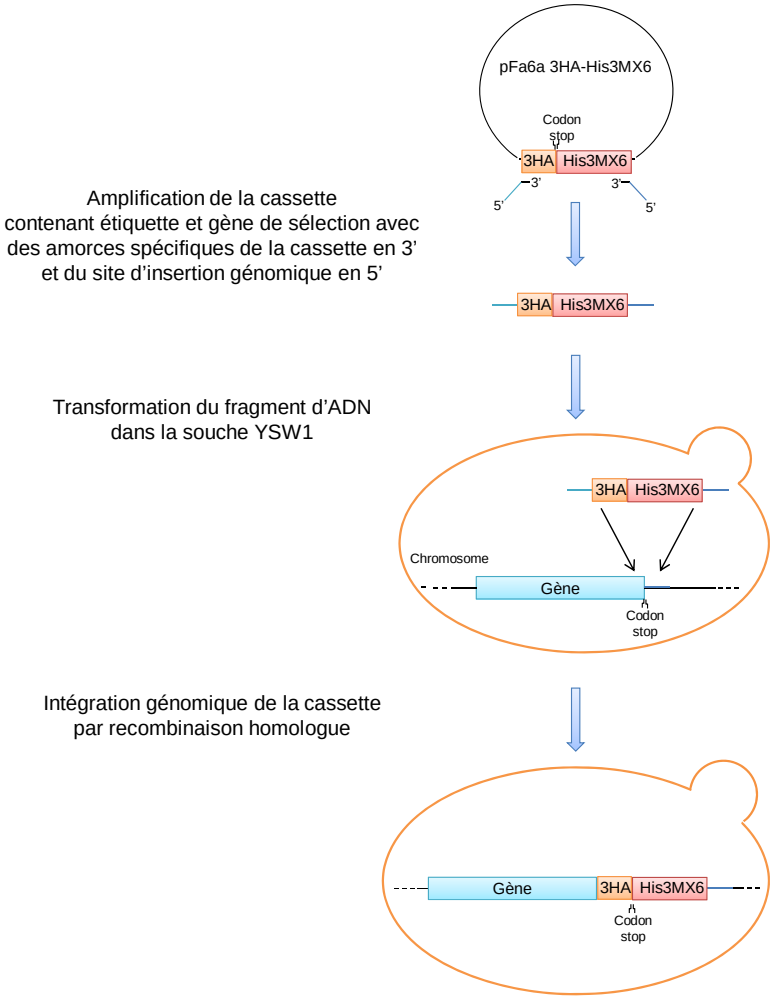


Figure 10. Marquage d'une protéine à son extrémité C-terminale avec une étiquette 3HA.

Une cassette contenant l'étiquette 3HA et un gène de sélection permettant la synthèse de l'histidine est amplifiée par PCR avec des amorces constituées en 3' d'une séquence spécifique de la cassette et en 5' d'une séquence homologue au point d'insertion. De plus, la séquence codant l'étiquette 3HA contient un codon stop à son extrémité 3'. Le fragment d'ADN obtenu est alors transformé dans la souche YSW1, auxotrophe à l'histidine. Il est alors inséré à l'extrémité 3' du gène d'intérêt avant le codon stop par recombinaison homologue à l'aide de la machinerie cellulaire propre aux levures. Le codon stop du gène est excisé tandis que la séquence codante l'étiquette 3HA est juxtaposée au gène. Un codon stop à la fin de cette séquence indique à la machinerie cellulaire l'extrémité de la protéine recombinante.

RESULTATS

1. Expression et caractérisation biochimique des complexes RNase P et MRP

Chez les eucaryotes, les RNases P et MRP sont des complexes ribonucléiques apparentés qui partagent une composition protéique très similaires, puisque huit protéines leur sont communes. La purification spécifique de chacun de ces complexes se révèle donc être un véritable challenge expérimental à cause de la sensibilité des noyaux ARN aux nucléases, de leur haute similarité au niveau protéique et de leur faible abondance.

Nous avons choisi de purifier différentiellement les complexes RNases P et MRP par une stratégie basée sur une double purification d'affinité séquentielle en utilisant des marqueurs spécifiques présents exclusivement dans chacun de ces deux complexes. J'ai utilisée pour cette étude la souche YSW1 de *S. cerevisiae* dans laquelle une étiquette TAP est fusionnée à l'extrémité C-terminale de Pop4, protéine commune aux deux complexes. Il s'agit d'une souche dans laquelle les gènes des deux principales nucléases cellulaires, Nuc1 et Xrn1 (Vincent *et al.*, 1988; Larimer *et al.*, 1990), ont été mutés pour réduire les problèmes liés à la dégradation des noyaux ARN. Dans cette souche, j'ai marqué avec une étiquette HA soit la protéine Rpr2 (spécifique de la RNase P) obtenant la souche YSW1-Rpr2-HA soit la protéine Snm1 (spécifique de la RNase MRP) obtenant la souche YSW1-HA-Snm1.

1.1. Marquage et Expression des protéines d'intérêt

1.1.1. Principe du marquage de protéine de levure par recombinaison homologue

Une cassette contenant une étiquette et un gène de sélection est générée par PCR à l'aide d'amorces dont la partie 5' est spécifique du point d'insertion du gène d'intérêt. Puis, la cassette est insérée à l'extrémité choisie de ce gène par recombinaison homologue. Les recombinants sont alors sélectionnés en fonction de l'acquisition du gène de sélection, un gène permettant la synthèse de l'histidine (**Figure 10**).

1.1.2. Résultats

Pour déterminer quels clones obtenus après transformation sont les plus adéquats pour la purification des RNases P et MRP, j'ai testé pour chaque clone la localisation génomique de l'étiquette puis l'expression des protéines recombinantes.

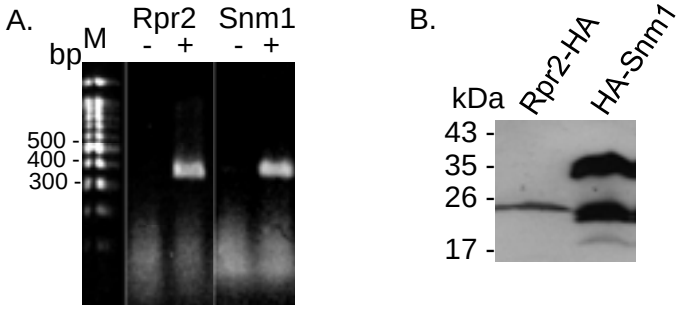


Figure 11. Marquage des protéines Rpr2 et Snm1 avec l'étiquette HA dans la souche YSW1.

A) Mise en évidence de l'intégration de l'étiquette 3HA à l'extrémité des gènes codant respectivement Rpr2 et Snm1 dans la souche YSW1 par amplification d'un fragment de 403 bp.
"-" : clone négatif sans intégration. "+" : clone positif avec intégration. M : marqueur de poids moléculaire.
B) Mise en évidence de l'expression des protéines recombinantes Rpr2-HA et HA-Snm1 dans la souche YSW1 par Western-blot avec anticorps anti-HA.

1.1.2.1. Test de l'intégration de l'étiquette.

Pour tester la localisation de la cassette contenant l'étiquette HA, j'ai réalisé une PCR à l'aide d'un couple d'amorces dont une s'hybride dans la séquence codant pour l'étiquette HA et l'autre dans la séquence codant pour la protéine à marquer. L'amplification d'un fragment d'ADN de 403 pb signifie que la cassette est adjacente au gène cible tandis que la non-amplification signifie que soit la cassette s'est intégrée dans un autre locus ou qu'elle ne s'est pas intégrée, le clone pouvant être un révertant, ayant acquis à nouveau la possibilité de pousser sur milieu sans histidine.

Par cette méthode, j'ai mis en évidence l'existence de plusieurs clones issus de la souche YSW1 ayant intégré la cassette avec 3HA dans le locus voulu à l'extrémité 5' du gène *YIR015w* codant Rpr2 ou à l'extrémité 3' du gène *YDR478w* codant Snm1 (**Figure 11.A**)

1.1.2.2. Test d'expression

Pour tester l'expression des protéines recombinantes Rpr2-HA et HA-Snm1, j'ai extrait le contenu cellulaire total à partir d'une culture des clones ayant intégrés génomiquement l'étiquette dans la position souhaitée. Puis, j'ai recherché la présence de l'étiquette 3HA par Western-blot avec un anticorps anti-HA (**Figure 11.B**).

J'ai ainsi mis en évidence une protéine recombinante Rpr2-HA de 25 kDa (dont 16,3 kDa pour Rpr2) dans l'extrait issu du clone ayant intégré l'étiquette à l'extrémité C-terminale de Rpr2. Dans l'extrait issu du clone ayant intégré l'étiquette à l'extrémité N-terminale de Snm1, deux signaux ont été mis évidence par Western-blot à 24 kDa et à 36 kDa. Comme la protéine Snm1 est prédite avec un poids moléculaire de 22.5 kDa et qu'aucune intégration d'étiquette n'a pu être réalisée avec succès à l'extrémité C-terminale de son gène, il semble que cette protéine existe sous au moins 2 formes dont la partie N-terminale serait constante.

En conclusion, j'ai obtenu les deux souches YSW1-Rpr2-HA et YSW1-HA-Snm1 exprimant les protéines recombinantes nécessaires à la réalisation des purifications différentielles des RNases P et MRP.

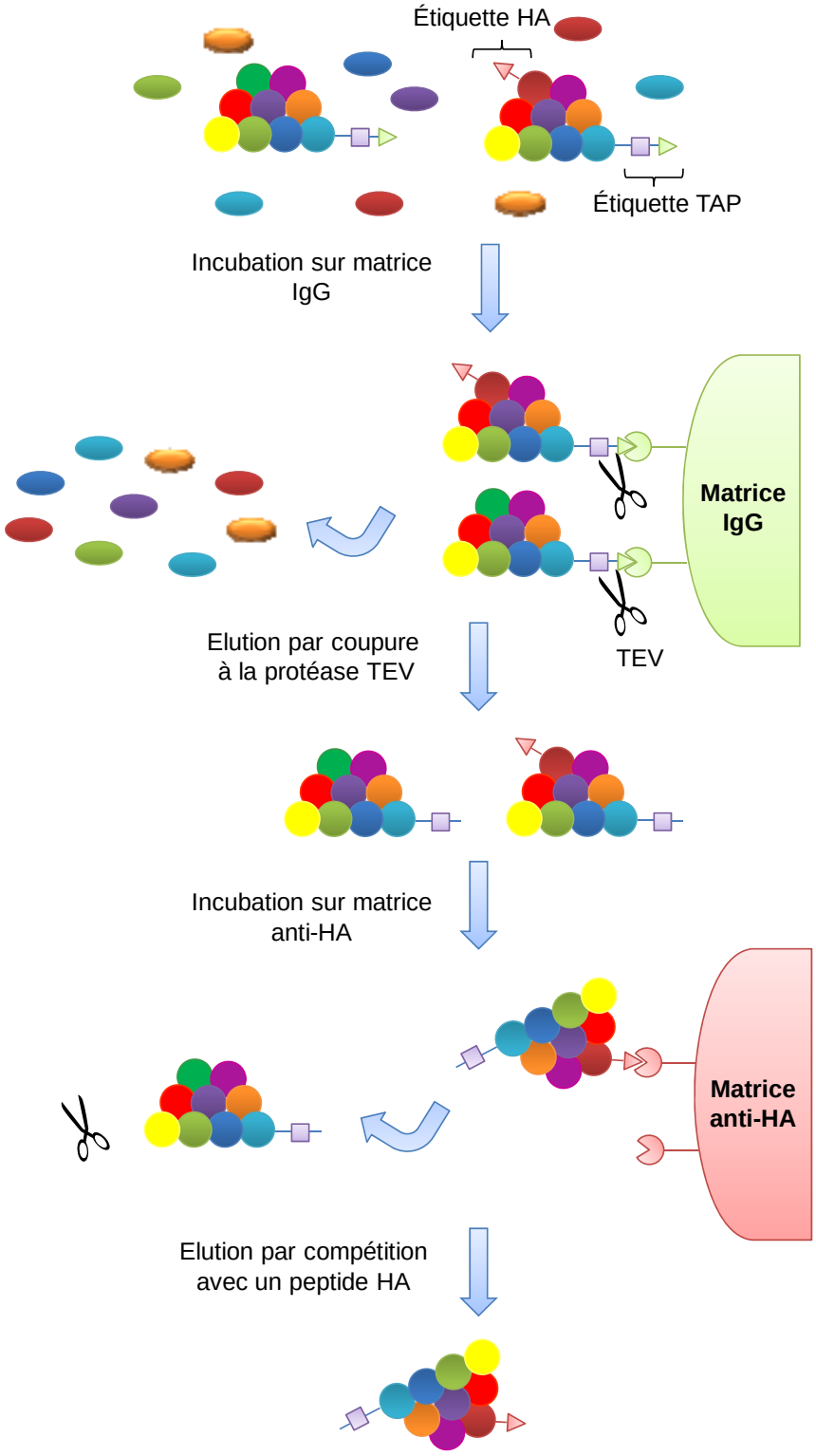


Figure 12. Stratégie de purification différentielle des complexes RNases P et MRP.
Un extrait cellulaire est réalisé à partir de la souche de levure YSW1 pour laquelle la protéine Pop4 porte l'étiquette TAP et une protéine spécifique (Rpr2 pour RNase P ou Snm1 pour RNase MRP) porte une étiquette HA (triangle rouge). L'étiquette TAP est composée d'un domaine CBP (*calmodulin-binding peptide*, carré violet), d'un site de coupure par la protéase TEV (trait bleu) et deux unités liant les IgG issues de la protéine A de *Staphylococcus aureus* (triangle vert). Les ciseaux représentent la protéase TEV.

1.2. Purification séquentielle des complexes RNase P et MRP

1.2.1. Principe de la purification

Une culture est réalisée avec une souche de levure soit YSW1-Rpr2-HA soit YSW1-HA-Snm1. Après lyse des cellules, l'extrait cellulaire est mis en présence d'une matrice avec des IgG. Les protéines Pop4-TAP se lient aux IgG par l'intermédiaire du domaine protéine A de l'étiquette TAP, les complexes RNases P et RNases MRP sont alors co-purifiés. Puis, la coupure par la protéase TEV permet d'éluer les complexes. L'éluat est alors mis en présence d'une matrice avec des anticorps anti-HA. Seul le complexe dont une protéine porte l'étiquette HA (Rpr2 ou Snm1 selon la souche) est retenu sur la matrice. Il est élué par compétition avec un peptide HA (**Figure 12**). L'efficacité de la purification est vérifiée par mise en évidence de l'étiquette HA par Western-blot (**Figure 13.A**). La spécificité de la purification est vérifiée par la détermination des composants protéiques (**Figure 13.B**) et par quantification relative des noyaux ARN de l'éluat (**Figure 14**). La qualité des complexes purifiés est vérifiée par microscopie électronique (**Figure 13.C**).

1.2.2. Vérification de l'efficacité de la purification

L'efficacité de la purification est testée par la mise en évidence de l'étiquette HA dans les différentes fractions issues de la purification. L'étiquette HA sur les protéines est révélée après SDS-PAGE par Western-blot avec un anticorps anti-HA.

Quelque soit la purification, l'étiquette HA est présente dans l'éluat issu de la coupure de l'étiquette TAP par la protéase TEV (**Figure 13.A**, pistes "Eluat TEV"). Après incubation de cet éluat sur la matrice anti-HA, la fraction non-retenue ne contient plus de protéines avec l'étiquette HA (**Figure 13.A**, pistes "FT"). Enfin, après compétition avec un excès de peptides HA, les protéines avec l'étiquette sont éluées comme attendue (**Figure 13.A**, piste "Eluat HA").

La purification des protéines portant l'étiquette HA issues des souches YSW1-Rpr2-HA et YSW1-HA-Snm1 que j'ai développée s'avère être donc un moyen efficace de purifier les protéines Rpr2 ou Snm1.

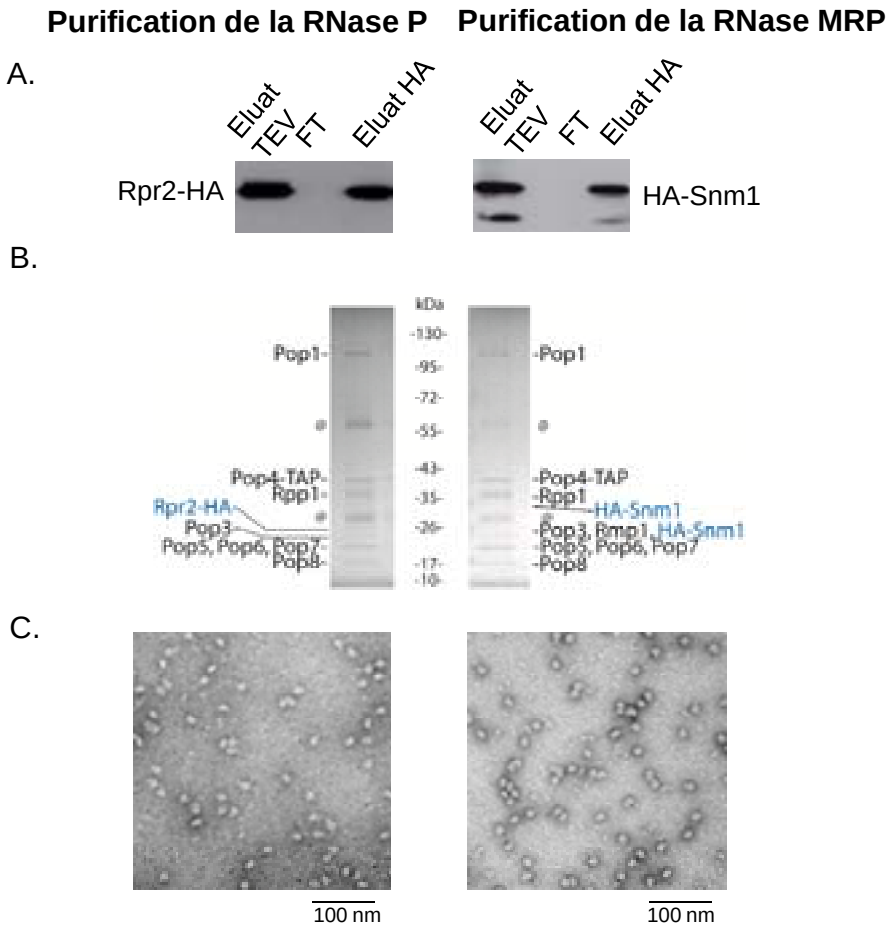


Figure 13. Purification des RNases P (à droite) et MRP (à gauche).
A) Identification par Western-blot des protéines marquées avec une étiquette HA : Rpr2 pour la RNase P, Snm1 pour la RNase MRP lors des étapes de la purification. Eluat TEV : Eluat après coupure à la protéase TEV. FT (Flow Through): Fraction non retenue sur la matrice anti-HA. Eluat HA : Eluat issu de la matrice anti-HA.
B) SDS-PAGE coloré au bleu de Coomassie des RNases P et MRP purifiées. A gauche: éluat HA issu de la purification à partir de la souche YSWI Rpr2-HA (RNase P) ; à droite: éluat HA issu de la purification à partir de la souche YSWI HA-Snm1 (RNase MRP). @, contaminants connus
C) Micrographes des éluats HA colorés à l'acétate d'uranyle 2% préparés selon la technique "sandwich" et visualisés par un microscope électronique Morgani.

1.2.3. Détermination de la composition des complexes purifiés

L'analyse par spectrométrie de masse de l'éluat issu de la purification à partir de la souche YSW1-Rpr2-HA a mis en évidence la présence des protéines communes aux deux complexes (Pop1, Pop3-8, Rpp1) ainsi que la protéine Rpr2 étiquetée. Comme attendu, la protéine Snm1 n'a pas été identifiée. (**Figure 13.B**, partie gauche)

Dans le cas de la purification à partir de la souche YSW1-HA-Snm1, nous avons également mis en évidence la présence des protéines communes aux deux complexes ainsi que les protéines Rmp1 et Snm1 étiquetée, retrouvées aux mêmes poids moléculaires que les signaux de Western-blot. La protéine Rpr2 n'est pas, quant à elle, retrouvée dans cette purification. (**Figure 13.B**, partie droite)

Ces résultats montrent donc que l'éluat issu de la purification à partir de YSW1-Rpr2-HA contient la protéine Rpr2 et ses partenaires, c'est-à-dire majoritairement le complexe RNase P. L'éluat issu de la purification à partir de YSW1-HA-Snm1 contient pour sa part Snm1 et ses partenaires, et donc majoritairement le complexe RNase MRP. Ces constatations valident la spécificité de notre purification séquentielle.

1.2.4. Quantification des noyaux ARN RPR1 et NME1 purifiés

Pour confirmer quantitativement l'efficacité des purifications, j'ai déterminé par qPCR la quantité relative du complexe considéré comme contaminant par rapport au complexe majoritaire en analysant les composantes ARN spécifiques de chaque complexe : RPR1 pour RNase P et NME1 pour la RNase MRP.

Dans chaque éluat, j'ai purifié les ARN en digérant les protéines par un traitement à la protéinase K puis l'ADN éventuellement présents par un traitement à la DNase I. Les réactions de transcription inverse se réalisant usuellement entre 37°C et 55°C, j'ai choisi les amorces de façon à ce qu'elles ne s'hybrident pas dans les hélices formées par l'ARN, non dénaturées à ces températures. J'ai également optimisé les conditions de transcription inverse par le choix d'une transcriptase adaptée, la transcriptase Transcriptor Reverse (Roche) thermostable jusqu'à 65°C, pour réaliser la réaction à plus haute température. Ainsi, pour obtenir une matrice ADN, une réaction de transcription inverse a été réalisée avec une amorce spécifique de chaque ARN à 60°C, température permettant de déployer ces ARN hautement repliés.

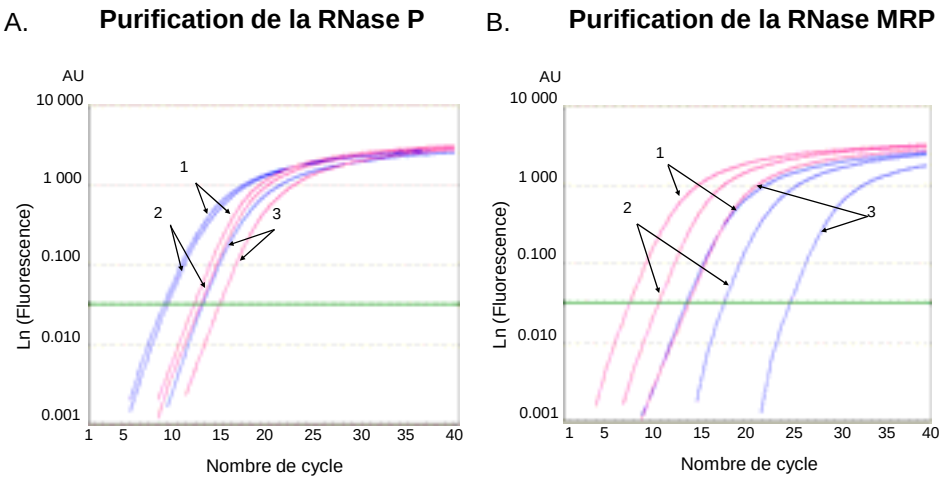


Figure 14. Quantification relative des ARN spécifiques de chaque complexe purifié.
Logarithme de l'intensité de la fluorescence en fonction du nombre de cycle de la PCR quantitative d'échantillons issus de trois purifications différentes de RNase P et RNase MRP.
A) Eluats HA issus de la souche YSW1- Rpr2-HA (RNase P). B) Eluats HA issus de la souche YSW1-HA-Snm1 (RNase MRP). Courbes bleues: ARN RPR1 (RNase P); Courbes rouges: ARN NME1 (RNase MRP). Ligne verte: seuil de détermination des valeurs Ct. Les quantités relatives d'un complexe par rapport à l'autre ont été déterminées pour trois expériences indépendantes indiquées par les flèches numérotées 1, 2 et 3.

Complexe	Purification	Ct moyenne		Δ Ct	Rapport	Rapport moyen	Ecart-type
		RPR1	NME1				
RNase P	1	12,3	15,49	3,19	9,12	12,17	3,36
	2	7,56	11,54	3,98	15,77		
	3	8,36	11,9	3,54	11,63		
RNase MRP	1	24	16,85	7,15	142,23	76,15	58.9
	2	21,03	15,19	5,83	57,07		
	3	16,59	11,73	4,87	29,16		

Tableau 5. Tableau des quantifications des noyaux ARN des trois purifications différentes de RNase P et RNase MRP.
La quantité relative d'un complexe par rapport à l'autre a été déterminée en calculant le rapport entre la quantité relative (Ct) d'ARN du complexe contaminant et celle du complexe attendu dans l'échantillon lors de la purification.

Enfin, une réaction d'amplification d'ADN avec des amorces spécifiques pour chaque matrice obtenue à l'étape précédente est réalisée en condition quantitative. Les deux réactions d'amplification des ADN complémentaires de RPR1 et de NME1 ont une efficacité très similaire d'environ 92,5% ce qui permet la comparaison des valeurs des cycles seuils des deux réactions entre elles.

Suite aux différentes purifications, les échantillons ne contenaient que peu d'espèces différentes d'ARN, idéalement une seule, il n'était donc pas possible d'amplifier un ARN référent interne. Dans ces conditions, à l'aide de la littérature et des conseils de l'ingénieur de la plateforme de génomique de l'institut, j'ai choisi d'utiliser un outil statistique adapté à cette problématique : la quantité relative d'un complexe par rapport à l'autre a été déterminée en calculant le rapport entre la quantité d'ARN du complexe contaminant et celle du complexe majoritaire dans l'échantillon lors de la purification selon la formule :

$$\textbf{[complexe contaminant]/[complexe majoritaire]} = 2^{(\textbf{Ct complexe contaminant} - \textbf{Ct complexe majoritaire})} = 2^{\Delta Ct}$$

où Ct correspond à la valeur du cycle seuil (*threshold cycle*) obtenu par qPCR.

Dans le cas d'une co-purification équimolaire des deux complexes, ce rapport serait égal à 1.

Après purification différentielle à partir de la souche YSW1-Rpr2-HA, l'éluat contenait en moyenne 1 molécule d'ARN NME1 (RNase MRP) pour 12 molécules RPR1 (RNase P) (**Figure 14.A**). Inversement, à partir de la souche marquée sur Snm1, l'éluat contenait en moyenne 1 molécule RPR1 pour 76 molécules de NME1 (**Figure 14.B**). Ces résultats détaillés dans le **Tableau 5** confirment ainsi l'efficacité de la purification.

1.2.5. Contrôle de la qualité des échantillons purifiés par microscopie électronique

Les éluats concentrés sont préparés sur grilles par la méthode de coloration négative en sandwich.

1.2.5.1. La méthode de coloration négative en sandwich

La méthode de coloration négative permet d'augmenter le contraste des particules observées par rapport à leur milieu. En microscopie électronique, le contraste entre deux points étant la différence de densité atomique entre ces deux points, l'utilisation en tant que colorant d'acétate d'uranyle, un sel d'uranium avec une densité atomique élevée (Z=92) permet de visualiser les complexes protéiques, composé d'atomes de carbone de plus faible numéro atomique (Z=6) (Brenner *et al.*, 1959). De plus, la prise en sandwich des particules entre deux fins films de carbone permet d'obtenir des grilles où les particules sont colorées de façon homogène (Valentine, 1961).

1.2.5.2. Résultats

L'analyse sous microscope électronique des deux types d'éluat préparés selon cette méthode a mis en évidence une répartition homogène de particules de taille et de forme similaire et l'absence d'agrégat (**Figure 13.C**). Ces constations témoignent de la qualité du matériel purifié.

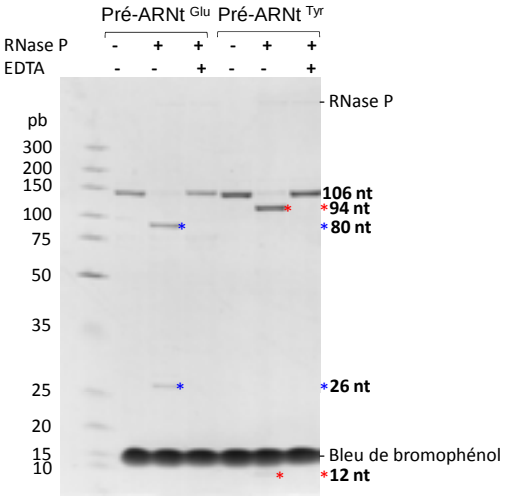


Figure 15. Test *in vitro* de l'activité de coupure en 5' de la RNase P purifiée sur 2 pré-ARNt différents.
Après incubation à 37°C pour 30 min, le mélange de réaction est séparé sur gel de polyacrylamide 12%- Urée 7M coloré avec une solution de bleu de toluidine 0.1 %. pré-ARNt ^{Glu} (*) pré-ARNt ^{Tyr} (*).

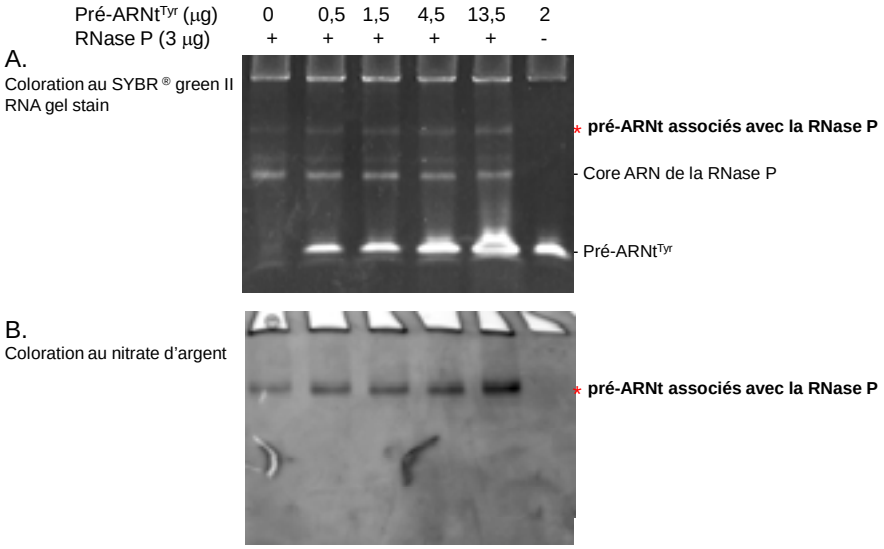


Figure 16. Test de liaison des pré-ARNt sur la RNase P purifiée par une expérience de retard de migration sur gel.
Après incubation pendant 30 min à température ambiante, le mélange réactionnel est séparé sur un gel de polyacrylamide 4.5%, coloré au *SYBR ® green II RNA gel stain* pour visualiser les ARN (A), puis au nitrate d'argent pour visualiser les protéines (B).

1.3. Test de l'activité de coupure endonucléolytique *in vitro* de la RNase P purifiée

Après purification, il était important de vérifier l'activité des complexes purifiés. Pour ma part, j'ai conçu un test d'activité *in vitro* de la RNase P. Bien que la radioactivité soit très couramment employée pour ce type de test, j'ai choisi de développer un test sans l'utiliser pour faciliter sa réalisation. A l'aide de la littérature et de la base de données SGD (*Saccharomyces Genome Database*), j'ai choisi de synthétiser deux pré-ARNt *in vitro* : un pré-ARNt ^{Tyr} précédemment décrit (Ziehler *et al.*, 2000) de 106 nucléotides (nt), générant 2 fragments de 12 et 94 nt après coupure et un pré-ARNt ^{Glu} dont j'ai optimisé la taille pour obtenir 2 fragments de 80 et 26 nt, plus différentiable après coupure. Après électrophorèse en conditions dénaturantes (PAGE), les fragments d'ARN sont alors révélés avec une solution de bleu de toluidine.

Le test de l'activité de coupure de la séquence leader 5' des ARNt prématures par la RNase P purifiée est illustré dans la **Figure 16**. Les séquences 5' leader (26 nt pour le pré-ARNt ^{Glu} et 12 nt pour le pré-ARNt ^{Tyr}) sont effectivement coupées par la RNase P purifiée pour générer un ARNt ^{Glu} (80 nt) et un ARNt ^{Tyr} (94 nt) avec une extrémité 5' mature. Aucune coupure n'a lieu en l'absence de RNase P ou en présence de RNase P et d'EDTA montrant la nécessité des ions divalents (ions magnésium,…) pour la réalisation de la réaction endonucléolytique mise en évidence par Pfeiffer *et al*, 2000.

J'ai ainsi mis en évidence que la RNase P purifiée est capable de couper les ARNt prématures. Notre méthode de purification conduit donc à l'obtention d'une enzyme fonctionnelle.

1.4. Test de la capacité de liaison des pré-ARNt à la RNase P purifiée

La RNase P purifiée étant active, il était complémentaire de vérifier la conservation de la capacité de liaison des ARNt prématures sur la RNase P purifiée. J'ai donc développé un test de retard sur gel où la mise en présence d'ARNt prématures et de RNase P purifiée conduit à une variation de la migration des composés lors de l'électrophorèse s'ils sont capables de se lier entre eux. La réaction a lieu en présence d'ions calcium, ions dits "structuraux" favorisant la capacité de liaison des ARNt prématures sur la RNase P (Warnecke *et al.*, 1999) en opposition aux ions magnésium, ions dits "catalytiques" essentiels à l'activité endonucléolytique.

La mise en présence de quantités croissantes de pré-ARNt^{Tyr} avec une quantité constante de RNase P purifiée conduit à l'augmentation de l'intensité d'un signal correspondant à la RNase P liée à un ARNt prémature (**Figure 16.A et B**). Ces données permettent de conclure que la RNase P purifiée a donc conservé sa capacité à lier les ARNt prématures.

1.5. Conclusions

J'ai ainsi développé une stratégie de purification séquentielle des complexes RNases P et MRP dont j'ai prouvé l'efficacité par spectrométrie de masse, par quantification des noyaux ARN et par la mise en évidence de l'activité de la RNase P purifiée et de sa conservation de ses propriétés de liaison aux ARNt prématures. Enfin, l'analyse préliminaire par microscopie électronique permet de conclure à la nature homogène des échantillons qui peuvent être utilisés pour acquérir un jeu de données par des techniques de microscopie électronique en vue d'une modélisation structurale des RNases P et MRP par traitement d'image.

2. Caractérisation structurale des complexes RNases P et MRP

Cette étude a été réalisée dans le cadre d'un travail d'équipe. J'ai réalisé la préparation des échantillons, la collecte initiale des particules, le développement de la méthode de marquage des sous-unités protéiques par les anticorps et la création des souches nécessaires à ces purifications (Annexe I). Le Dr. Böttcher, responsable de l'équipe, a réalisé le traitement des images pour la modélisation des complexes ainsi que le positionnement des modèles atomiques dans la carte de microscopie électronique. Le Dr. Hipp a réalisé le traitement des données pour la localisation des sous-unités protéiques par anticorps.

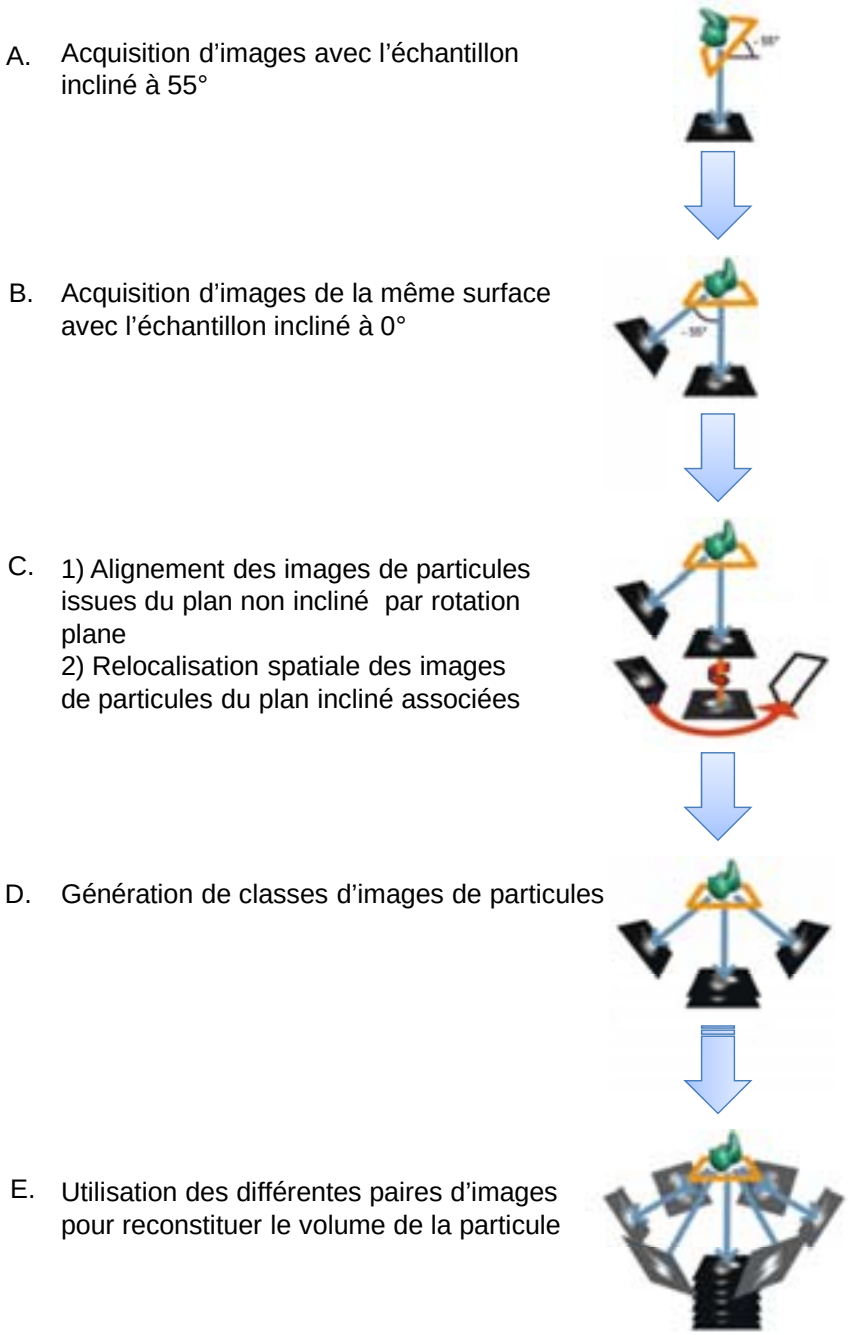
2.1. Modélisation des complexes RNases P et MRP

Les complexes RNases P et MRP étant peu abondants, nous avons choisi de réaliser leur modélisation par analyse d'images de particules préparées selon la méthode de coloration négative en sandwich collectées sur un microscope électronique. Dans un premier temps, il a été réalisé une reconstruction conique aléatoire à partir d'images de particules inclinées et non inclinées pour générer un modèle initial de chaque complexe. Puis, ces cartes ont été utilisées comme référence pour affiner la modélisation à l'aide d'un large ensemble d'images de particules non inclinées issues d'échantillons préparés selon la méthode de cryo-coloration négative en sandwich.

2.1.1. Modélisation initiale par reconstruction conique aléatoire

2.1.1.1. Principe de la reconstruction conique aléatoire (Radermacher *et al.* , 1987)

Les complexes macromoléculaires tels que les RNases P et MRP présentent souvent des orientations préférentielles sur une grille préparée selon la méthode de coloration négative (voir partie 1.2.5). La méthode de reconstruction conique aléatoire est une méthode qui tire avantage de ce biais en considérant des sous-ensembles de particules de même orientation. En effet, un sous-ensemble de particules présentant des orientations identiques sur un plan non incliné correspond en fait à des particules orientées aléatoirement vues sur un autre plan. Quand l'échantillon est incliné selon un angle choisi θ, le sous-ensemble ci-dessus apparait comme une série de projections coniques des particules avec des orientations aléatoires et un angle de cône θ .



Adaptée de Leschziner, 2010

Figure 17. Principe de la reconstruction conique aléatoire.
L'objet vert représente une particule, le quadrilatère orange son support, les fines flèches bleues l'orientation des images projetées et les flèches rouges les rotations planes appliquées lors de l'alignement des particules.
Les données relatives à une particule sont acquises en deux temps : pour chaque champ, (A) une image est d'abord acquise selon un angle de -55° (B) puis selon un angle de 0°. (C) Les images de particules issues du plan non incliné sont ensuite alignées si elles présentent la même vue par application d'une rotation plane. Cette rotation détermine une nouvelle localisation spatiale pour les images appariées issues du plan incliné. (D) Les images alignées sont alors membres d'une même classe, représentant la même vue et étant alignées les unes par rapport aux autres. Une image moyenne représentative de la classe est générée. (E) Pour chaque classe, les différentes vues issues des images de plan incliné sont alors utilisées pour calculer une reconstruction du volume de la particule initiale.

Pour mettre en œuvre la méthode de reconstruction conique aléatoire, une image des particules est acquise selon un plan incliné d'un angle θ, dans notre cas θ = -55° (**Figure 17.A**), puis selon un plan non incliné (**Figure 17.B**). Lors de la procédure d'alignement des images de particules issues du plan non incliné, les vues identiques sont alignées par rotation plane tandis que les données relatives aux orientations des images issues du plan incliné permettent d'assigner à ces images une localisation spatiale (**Figure 17.C**). Le sous-ensemble de particules du plan non incliné alignées entre elles forment une classe pour laquelle est générée une image moyenne de la classe de la particule (**Figure 17.D**). Une reconstruction du volume des particules est alors calculée pour toutes les classes bien définies par une faible variance entre les membres des classes en tenant compte de l'angle du cône et des orientations issues des images du plan incliné (**Figure 17.E**).

Cette méthode permettant de reconstituer le volume d'une macromolécule est cependant limitée par l'absence de données pour certains angles ce qui est due à la technique d'acquisition des images. Il n'est en effet pas possible techniquement d'acquérir des images à un angle de 90°. Le volume correspondant au manque d'information est appelé cône manquant, en raison de sa forme dans l'espace de Fourier, un espace mathématique contenant toutes les informations nécessaire à la formation du volume. Cela se traduit par une élongation du volume selon un axe vertical et à une perte de détails à sa surface. Pour pallier en partie à cette limitation, les volumes générés ont alors été affinés par appariement des projections des images avec et sans inclinaison.

2.1.1.2. Obtention d'une carte initiale de la surface des complexes RNases P et MRP

De nombreuses classes ont donné des volumes cohérents lors de la comparaison entre la rétroprojection des volumes obtenus et les images moyennes de classe de particules mais ils montrent un aplatissement dû aux procédés de coloration et de séchage de l'échantillon. De nouvelles images de particules non inclinées ont été ajoutées à l'ensemble de données donnant un total (inclinées + non inclinées) de 9 226 images de particules pour la RNase P et de 7 553 images pour la RNase MRP.

Les volumes finaux ne montrent aucune indication d'aplatissement (**Figure 18**, panel à gauche) et représentent correctement les images moyennes de classe initiale des particules non inclinées (**Figure 18**, panel au centre). Ainsi, les reconstructions coniques aléatoires des RNases P et MRP se présentent comme des modèles de départ fiables pour un affinement ultérieur de cartes tridimensionnelles. La résolution de ces cartes de densité ont été estimées à

2.7 nm pour la RNase P et de 3.1 nm pour la RNase MRP pour une corrélation des enveloppes de Fourrier (FSC) de 0.5 (**Figure 18**, panel à gauche).

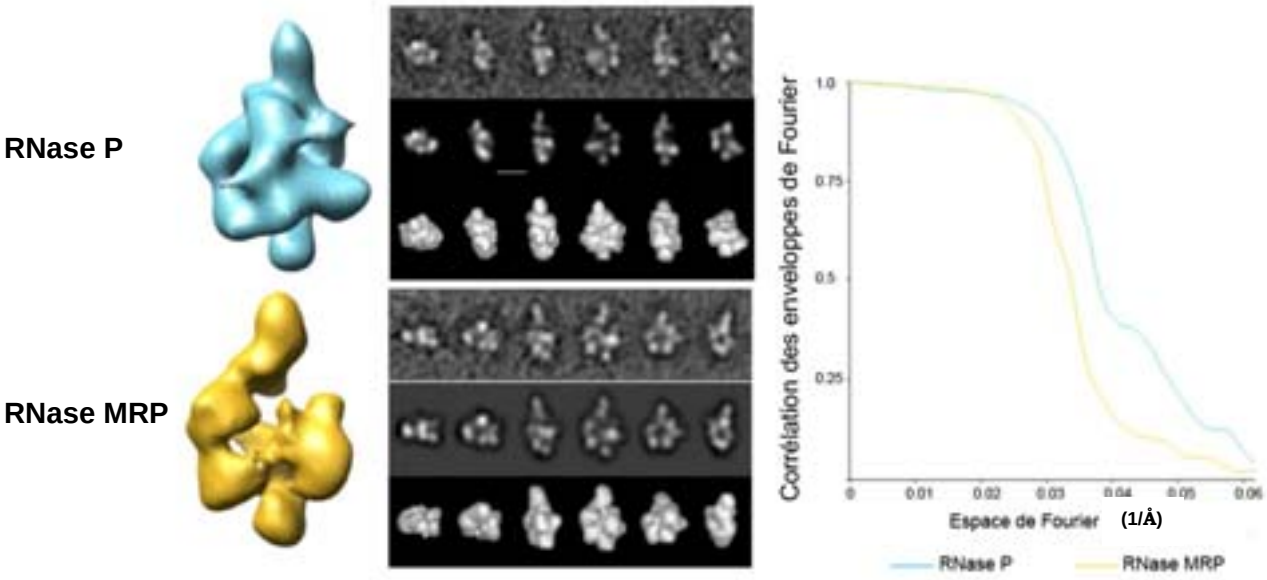


Figure 18. Reconstruction aléatoire conique initiale des complexes RNases P (en haut) et MRP (en bas).

Panel de gauche : Cartes initiales des RNases P et MRP issues de la reconstruction aléatoire conique initiale, cartes montrées dans la même orientation.

Panel de droite : Images moyennes des classes à partir d'appariement de projections (ligne du haut). Projection du volume dans la même direction que celle des images moyennes de classe (ligne au centre). Représentation de la surface vue dans la même direction que les projections (ligne du bas)

Panel de droite : Corrélation des enveloppes de Fourier (FSC) pour les cartes des RNases P et MRP. Pour une FSC de 0.5, la fréquence spatiale était de 1/2.7 nm^{-1} pour la RNase P (en bleu) et de 1/3.1 nm^{-1} pour la RNase MRP (en jaune).

2.1.2. Etablissement de cartes tridimensionnelles des RNases P et MRP

Pour la détermination d'une structure à plus haute résolution par microscopie électronique, les RNases P et MRP ont été préparées selon la méthode de cryo-coloration négative en sandwich qui permet de préserver l'hydratation des particules et de réduire leur aplatissement par comparaison avec la coloration négative standard (Golas *et al.*, 2003). Au total, 42 104 images de particules RNase P et 59 104 images de particules RNase MRP ont été sélectionnées pour analyse. La reconstruction finale est calculée à partir de 24 373 images de particules pour la RNase P et de 32 232 images de particules pour la RNase MRP. Ces particules se présentent sous la forme d'une goutte dont l'axe le plus long mesure approximativement 20 nm. L'inspection visuelle montre que les rétroprojections de la reconstruction correspondent avec les images moyennes de classes calculées par analyse statistique multivariée (**Figure 19.A**) et ne montrent pas d'hétérogénéité majeure dans les sous-classes soulignant la validité de la carte tridimensionnelle (3D).

Pour tester si les différences observées entre les cartes des RNase P et RNase MRP étaient liées à la variabilité du traitement des images ou à des réelles différences entre elles, les deux cartes ont été combinées en une référence commune initiale pour un affinement itératif. Les images des RNase P et RNase MRP ont été traitées séparément. L'affinement a été poursuivi jusqu'à ce que la plupart des orientations des particules ne change plus et qu'aucune amélioration de la résolution ne soit observée. L'affinement a convergé vers deux cartes distinctes (**Figure 19.C**), ce qui a mis en évidence les différences-clés (**Figure 20**), similaires à celles observées initialement avec les cartes individuelles des RNase P et MRP issues de la reconstruction conique aléatoire.

Les cartes 3D ont une résolution de 1.7 nm pour la RNase P et de 1.5 nm pour la RNase MRP correspondant à une FSC de 0.5 (**Figure 19.B**).

2.1.3. Une structure modulaire commune aux deux complexes

La reconstruction tridimensionnelle des RNases P et MRP révèlent que les deux complexes partagent une architecture similaire (**Figure 19.C**). Les RNases P et MRP se composent de 3 modules qui forment une cavité en leur centre. Les limites des modules ont été définies par les zones où leurs contacts sont les moins nombreux avec les modules voisins. Je me référerai à ces modules comme étant les modules de base, arrière et de tête.

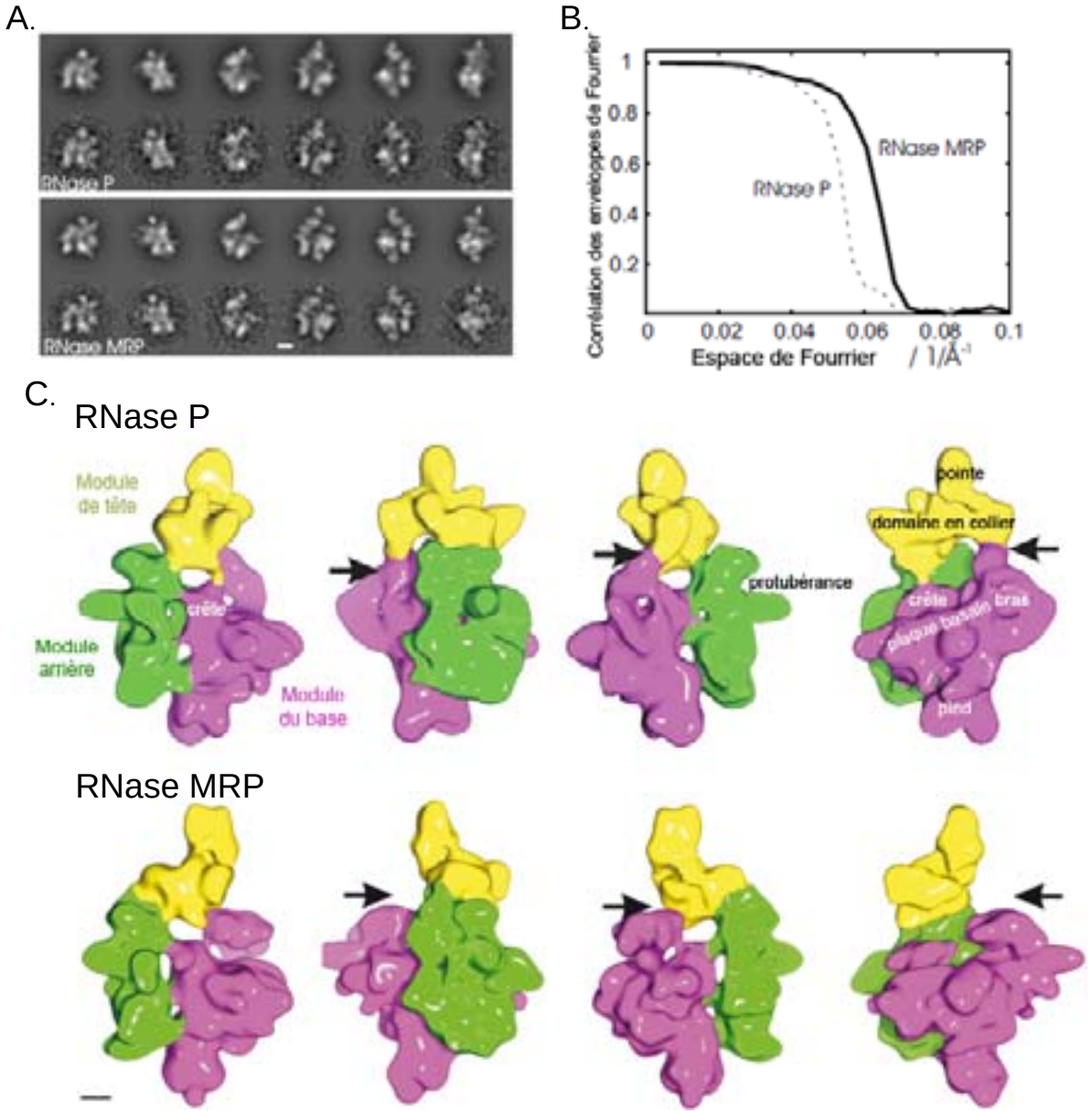


Figure 19. **Structures de la RNase P et de la RNase MRP obtenues à partir des images collectées en microscopie électronique.**

A) Sélection de projections de la carte 3D de la RNase P (panel du haut) et de la RNase MRP (panel du bas) ainsi que les images moyennes de classes correspondantes calculées par analyse statistique multivariée (ligne du bas de chaque panel). Echelle : 5 nm.

B) Corrélation des enveloppes de Fourier pour les modélisations de la RNase P (pointillé) et de la RNase MRP (trait continu). Pour une FSC de 0.5 (Bottcher *et al.*, 1997), la fréquence spatiale était de 1/1.7 nm^{-1} pour la RNase P et de 1/1.5 nm^{-1} pour la RNase MRP.

C) Structure tridimensionnelle de la surface de la RNase P (panel du haut) et de la RNase MRP (panel du bas). Les structures sont représentées selon des rotations successives de 90° autour de l'axe vertical (sens antihoraire). Distinction de 3 modules: module de base (magenta); module arrière (vert) et module de tête (jaune). Les flèches pointent l'interface entre le "bras" du module de base et le module de tête qui sont en contact dans la RNase P et séparés dans la RNase MRP.

Echelle : 2 nm.

2.1.3.1. Le module de base

Le module de base se compose de la plaque basale, une entité trapézoïdale plate avec une alternance de crêtes et de rainures, un domaine globulaire au pied du module et un domaine en forme de bras qui est en contact avec le module de tête dans la RNase P mais pas dans la RNase MRP (**Figure 19.C** et **Figure 20**, flèche noire).

La plaque basale forme la plus large surface de la cavité central et s'élève jusqu'à une crête près du module arrière. Dans la RNase P, la crête occupe deux côtés de l'entrée d'un conduit qui connecte les cavités entre le module de base et le module arrière. Au contraire, dans la RNase MRP, la crête est localisée sur un seul côté de l'entrée du conduit, permettant la formation d'un accès plus allongé à ces cavités (**Figure 20**, panel à gauche, flèche ouverte).

Dans la RNase P, le conduit traverse le complexe de part en part tandis qu'il est bloqué dans la RNase MRP (**Figure 20**, panel au centre : ●). Cela coïncide avec l'accessibilité du conduit à partir de l'extérieur de la RNase MRP et sa clôture dans la RNase P (**Figure 20**, panel à droite : o).

Dans les deux complexes, le bras du module émerge de l'autre côté de la plaque basale. Dans la RNase P, ce bras, plus fin et plus allongé permet le contact avec le module de tête (**Figure 19.C** et **Figure 20**, flèche noire) tandis que dans la RNase MRP, le bras, plus globulaire, n'est pas connecté au module de tête. Par ailleurs, la RNase MRP possède une densité supplémentaire près de l'endroit où le bras du module émerge de la plaque basale (**Figure 20**, panel à droite : *).

2.1.3.2. Le module de tête

Le module de tête est le plus variable entre les différentes reconstructions, ce qui pourrait être causé par une certaine flexibilité ou une haute susceptibilité à la dégradation. Ce module se compose de deux domaines auxquels je me réfère comme les domaines en collier et en pointe (**Figure 19.C**). Chez la RNase P, le domaine en collier entre en contact avec le bras du module de base ainsi qu'avec le module arrière. Par contre, chez la RNase MRP, le module de tête est seulement connecté au module arrière.

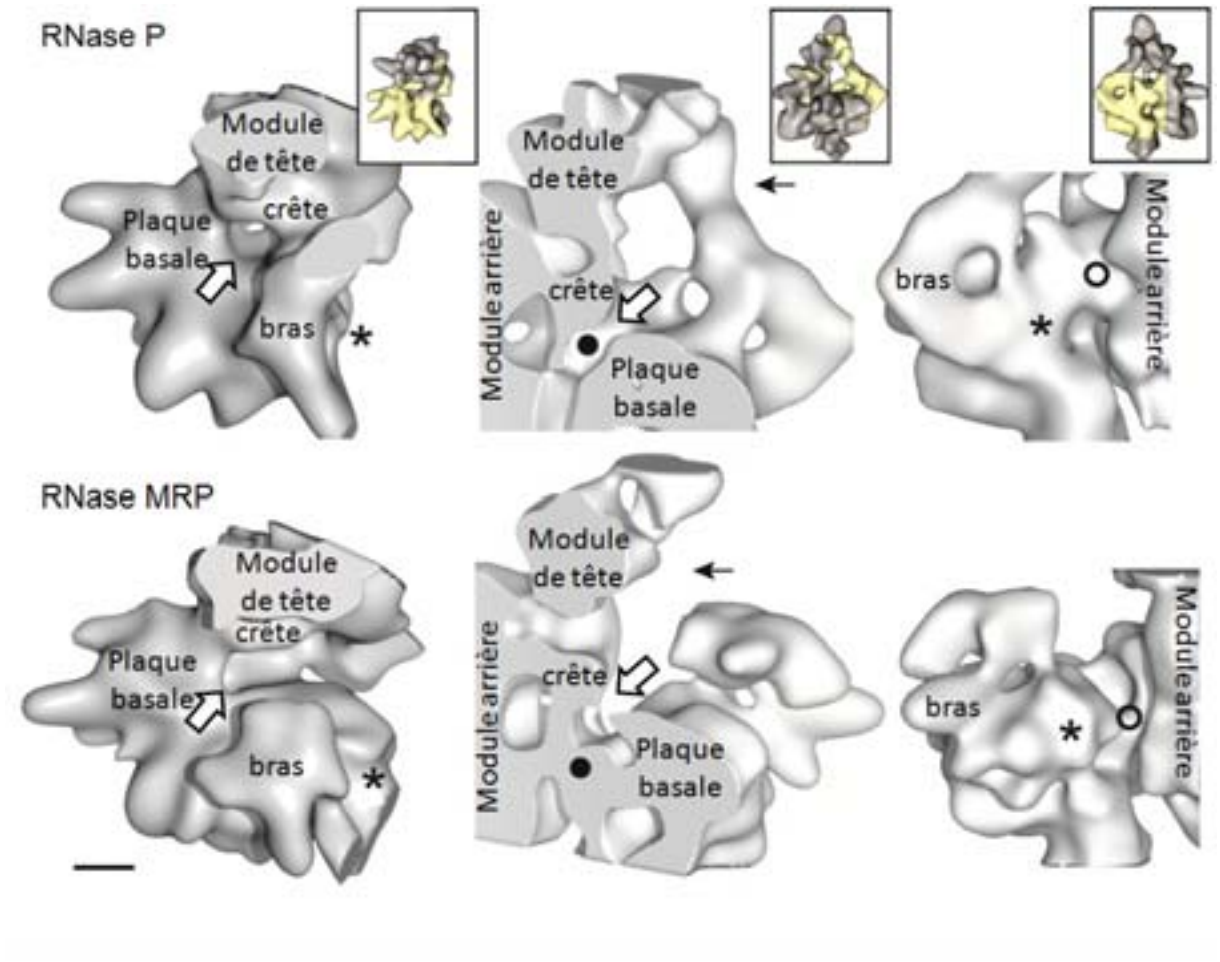


Figure 20. Vues agrandies des régions de RNase P (en haut) et de RNase MRP (en bas) dont les différences entre elles sont les plus prononcées.

Dans les images insérées ci-dessus, le complexe se trouve dans la même orientation que la vue agrandie, représentée en jaune.

Dans le panel à gauche, l'entrée du conduit (flèche ouverte) est close dans la RNase P et ouverte sur un côté dans la RNase MRP. La RNase MRP a une densité additionnelle près du module en forme de bras (*) qui est absente dans la RNase P.

Dans le panel au centre, les RNases P et MRP sont représentées selon une coupe dans la même position au travers du centre du conduit. Le conduit est ouvert dans la RNase P et fermé dans la RNase MRP (●). Le domaine en forme de bras (*) est connecté au module du tête dans la RNase P et en est séparé dans la RNase MRP (flèche noire).

Dans le panel de droite, le conduit est protégé de l'extérieur dans la RNase P et est accessible dans la RNase MRP (o). Enfin, la densité additionnelle de la RNase MRP (*) près du domaine en forme de bras donne lieu à un sillon qui est manquant dans la RNase P.

Echelle : 2 nm

2.1.3.3. Le module arrière

Le module arrière se compose de deux domaines similaires avec une protubérance centrale qui sont liés par une symétrie d'ordre 2, ce qui suggère un bloc dimérique. La partie inférieure du noyau dimérique contribue au conduit central, tandis que la moitié supérieure ancre le module de tête.

Le module arrière est connecté au module de base en plusieurs points de contact définis. La partie supérieure touche la crête tandis que la partie inférieure entre en contact principalement avec la partie inférieure de la plaque basale.

2.2. Localisation structurale des sous-unités protéiques

Pour construire un modèle pseudo-atomique des RNases P et MRP, il fallait positionner les structures à haute résolution des sous-unités homologues de différentes espèces ainsi que le modèle structural de la sous-unité ARN de *Schizosaccharomyces pombe* (Marquez *et al.*, 2006) dans nos cartes 3D. Cependant, nous nous sommes heurtés à des ambigüités quant à la position de certaines des structures atomiques liées à la résolution des cartes obtenues. Nous avons alors localisé plusieurs sous-unités des RNases P et MRP par microscopie électronique et traitement d'image de complexes avec des sous-unités marquées avec des anticorps. Ceci nous a alors permis de positionner les structures à haute résolution existantes au sein de nos cartes 3D.

2.2.1. Localisation de différentes sous-unités de la RNase P par marquage aux anticorps.

Pour réaliser ce marquage, j'ai tout d'abord produit des protéines recombinantes avec une étiquette GFP à l'extrémité N-terminale sur Pop1 et Rpr2-HA ou à l'extrémité C-terminale sur Pop5, Pop6, Pop8 et Rpp1. Après essai avec nos logiciels de traitement d'images, l'étiquette GFP s'est révélée trop petite pour être directement détectable par différence entre 2 images de particules issues du plan non incliné. Pour accroître la taille de l'étiquette, j'ai alors optimisé la méthode de purification des complexes afin qu'un anticorps monoclonal puisse se lier à l'étiquette GFP lors de l'étape de liaison du complexe à la matrice anti-HA. Les taux de marquage par les anticorps est typiquement en dessous des 10%, ce qui se trouve dans la gamme d'efficacité de marquage que nous avons obtenu pour d'autres complexes (Fernandez-Tornero *et al.*, 2007; Diepholz *et al.*, 2008; Ulbrich *et al.*, 2009).

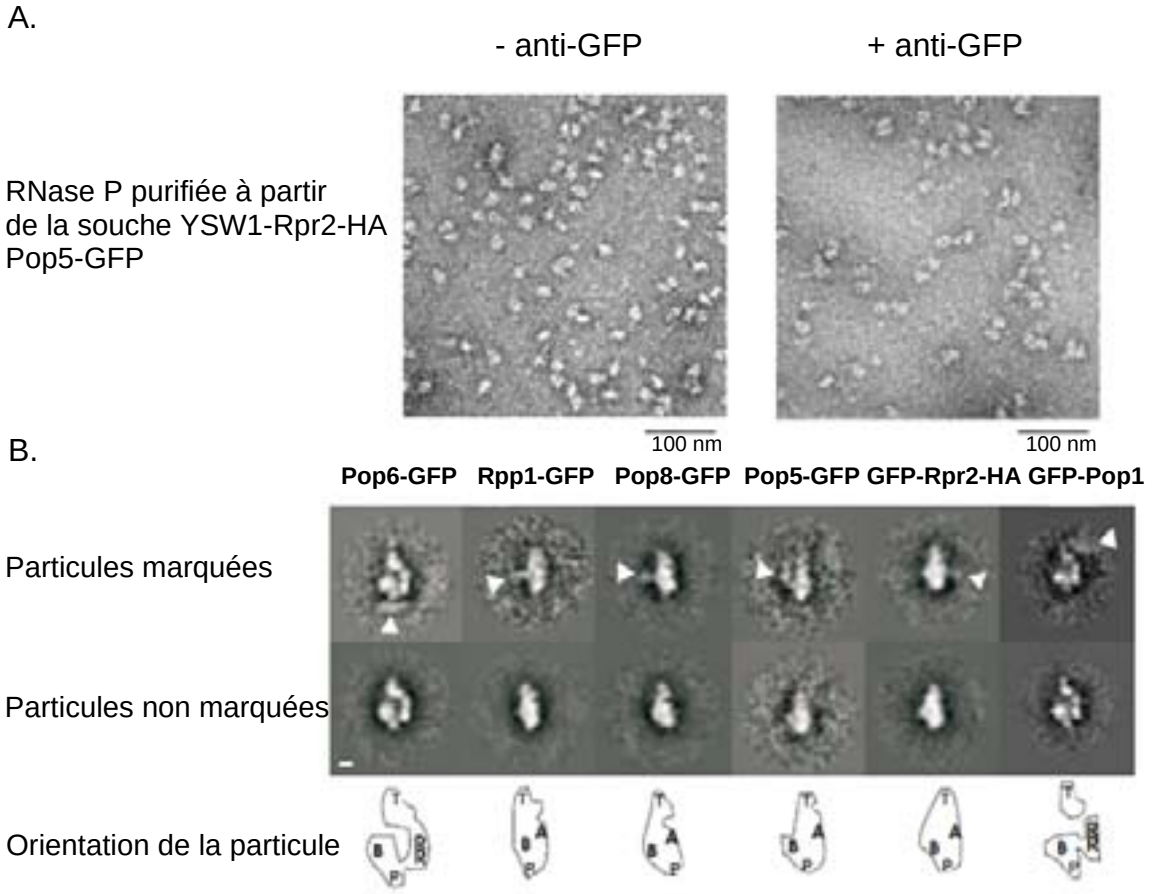


Figure 21. **Localisation de différentes sous-unités de la RNase P par marquage aux anticorps.**

A) Micrographes des éluats HA issus de la purification à partir de la souche YSW1-Rpr2-HA Pop5-GFP (RNase P) colorés à l'acétate d'uranyle 2% préparés selon la technique sandwich, visualisés par un microscope électronique Morgani. A gauche : éluat sans anticorps. A droite : éluat avec anticorps anti-GFP.
Echelle : 100 nm.

B) Galerie d'images moyennes de sous-classes avec une extra densité (ligne du haut) et sans extra densité (ligne au centre). Les flèches blanches pointent la position de l'extra densité due à l'anticorps. Echelle : 2 nm.
La ligne du bas montre l'orientation des complexes non marqués de la ligne au centre. (Grossissement : x2) Les positions approximatives des domaines sont indiquées par des lettres : A: domaine en forme de bras du module de base, T: module de tête; P: domaine au pied du complexe ; B: plaque basale; RR: module arrière.

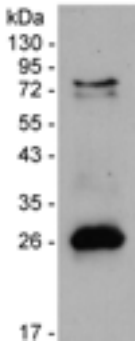


Figure 22. Expression de la protéine recombinante GFP-Pop1 (~130 kDa) dans la souche YSW1-Rpr2-HA révélée par Western-blot avec anticorps anti-GFP mettant en évidence un profil de dégradation de la protéine Pop1.

Les complexes marqués sont observés dans un nombre réduit d'orientation, ce qui est probablement du à la présence de l'étiquette GFP sur une des sous-unités du complexe (**Figure 21.A**). Cette orientation préférentielle rend impossible la cartographie tridimensionnelle des épitopes. Cependant, les complexes marqués peuvent être identifiés par des techniques de classification. Il est ainsi possible de déterminer la position approximative du complexe GFP-anticorps sur les particules (**Figure 21.B**).

La distribution bidimensionnelle (2D) des étiquettes liées aux anticorps a permis de localiser les protéines Rpp1, Pop5, Pop8 et Rpr2 au centre des particules. Les positions de Pop8 et Rpp1 se superposent virtuellement. Toutefois, à cause du manque d'information tridimensionnelle, il n'est pas possible de distinguer si ces sous-unités sont basées dans le module arrière ou à proximité de la plaque basale. La protéine Pop6, quant à elle, est localisée au pied du module de base tandis que la protéine Pop1 est localisée dans le module de tête. Pop1, d'un poids moléculaire de 101 kDa, pourrait par sa taille correspondre au module de tête dans son entier. En effet, la carte du module de tête est 100,000 $Å^3$ dans la RNase P, ce qui est un peu plus petit que ce qui est attendu pour une protéine de 101 kDa. Par ailleurs, l'analyse par spectrométrie de masse de purifications types montre que Pop1 est souvent partiellement dégradée. Cela est aussi fréquemment observé dans les Western-Blot mettant en évidence Pop1-GFP marquée à l'extrémité N-terminale en utilisant un anticorps anti-GFP (**Figure 22**). L'instabilité de Pop1 est cohérente avec la variabilité structurale que nous observons au niveau du module de tête et la sous-estimation du volume dans nos cartes 3D.

2.2.2. Positionnement des modèles atomiques

Grâce à la localisation des 6 sous-unités protéiques par marquage avec les anticorps, il est alors possible de positionner plusieurs structures à haute résolution de sous-unités protéiques et ribonucléiques mais aussi de sous-complexes dans nos cartes 3D. Ainsi, nous avons choisi d'utiliser :

- les structures atomiques archées homologues des sous-complexes Pop5-Rpp1 2:2 (Kawano *et al.*, 2006) et Rpr2-Pop4 (Honda *et al.*, 2008) issues de *P. horikoshii* (**Figure 5**),

- la composante ARN de *S. pombe* modélisée par homologie avec la structure à haute résolution de la composante ARN bactérienne de *B. stearothermophilus* (Marquez *et al.*, 2006) (**Figure 8**),

- la structure atomique de la protéine L7Ae issue de *M. jannashii*, homologue archée de Pop3 (Hamma *et al.*, 2004) (**Figure 5**),

- la structure atomique du sous-complexe Pop6-Pop7-P3 issu de *S. cerevisiae* (Perederina *et al.*, 2010a) (**Figure 7**).

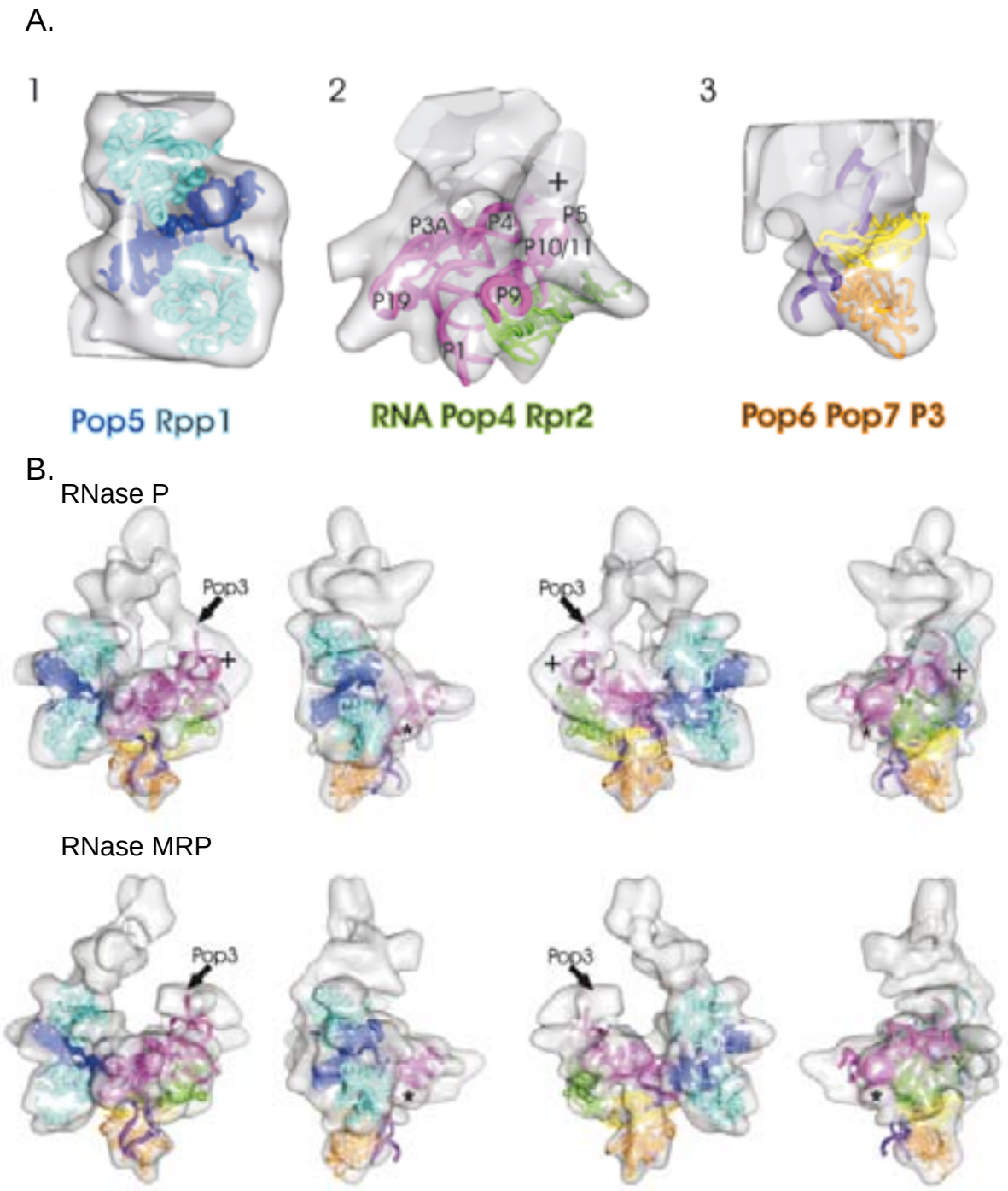


Figure 23. Alignement des structures atomiques des sous-unités avec les cartes des RNases P et MRP.

A) Vues agrandies de la carte de la RNase P alignée avec les modèles atomiques connus des sous-unités. Panel 1 : L'hétérotétramère [2CZV (Kawano *et al.*, 2006)] des homologues d'archée de Pop5 (en bleu foncé) et de Rpp1 (en bleu clair) est alignés au module arrière. Panel 2 : modèle du noyau ARN de *S. pombe* (Marquez *et al.*, 2006) (en magenta) et de l'homologue d'archée du dimère Pop4-Rpr2 [2ZAE, (Honda *et al.*, 2008)] (en vert) sont alignée à la plaque basale du domaine de base. Panel 3: Pop6 (en orange), Pop7 (en jaune) et la boucle ARN P3 (en violet) [3IAB (Perederina *et al.*, 2010a)] sont alignés au pied du module de base.

B) Sous-unités alignées avec la carte de la RNase P (panel du haut) et de la RNase MRP (panel du bas). Les vues sont mises en rotation autour de l'axe des ordonnées par saut de 90°. Pour le code couleur, se reporter au (A).

Le signe "+" souligne une densité non prise en compte à la fin de l'hélice P10/P11 où les 50 bases de l'hélice P12 sont absentes du modèle ARN. Le signe "*" souligne une densité en forme d'hélice au début de la boucle P3 où 30 bases sont absentes du modèle ARN. Les flèches soulignent la position approximative de Pop3.

2.2.2.1. Pop5 et Rpp1 forment le module arrière

Lors de la cristallisation du complexe d'archée aRpp1-aPop5, les protéines se sont organisées sous la forme d'un hétérotétramère asymétrique (Kawano *et al.*, 2006). A faible résolution, la structure de ce sous-complexe ressemble fortement au noyau dimérique du module arrière (**Figure 23.A**, panel 1). De plus, le module arrière correspond au meilleur appariement possible lors du positionnement automatique de la structure cristallisée du complexe d'archée Rpp1-Pop5 sur la carte segmentée avec une corrélation croisée de 0.89 (**Figure 24**). Cette position est aussi cohérente avec nos expériences de localisation qui indiquent que Rpp1 et Pop5 sont localisées dans le centre de la particule. Par conséquent, nous en concluons que le complexe Rpp1-Pop5 se trouve dans le module arrière. Dans cette position, le côté chargé positivement de Pop5, supposée liée l'ARN, fait face à la cavité entre le module de base et le module arrière.

D'autre part, si le complexe aRpp1-aPop5 correspond à la plupart de la densité du module arrière, une protubérance au centre est vacante. Cette caractéristique, proche de l'axe de symétrie local, suggère une certaine déviation à partir de l'organisation 2:2. Il semble que Pop8, localisée avec les anticorps sur la même position que Rpp1, ait remplacé une des deux protéines Pop5 dans la RNase P de *S. cerevisiae*.

2.2.2.2. La sous-unité ARN forme la plaque basale du domaine de base

Il a été montré par des expériences de protection que la liaison du complexe aPop5-aRpp1 à la composante ARN la protège contre l'action de nucléases autour de la région catalytique conservée formée par les hélices P3, P4 et P5 (Xu *et al.*, 2009). Cela situe la composante ARN avec son centre catalytique à proximité immédiate du module arrière, ce qui est cohérent avec son placement dans la plaque basale. Le modèle de la composante de la RNase P de *S. pombe* (Marquez *et al.*, 2006), considérablement plus petit que la sous-unité ARN de *S. cerevisiae*, a une forme trapézoïdale avec un motif alternant des crêtes et des rainures, caractéristiques concordant avec celles de la plaque basale. Le placement du modèle ARN dans la densité de la plaque basale avec les hélices P3A, P4 et P5 faisant face au module arrière met le noyau catalytique de la composante ARN à proximité immédiate de l'entrée du conduit et de la crête (**Figure 23.A**, panel 2). De plus, sur le côté opposé à la plaque basale, les hélices P10/11, P9, P1 et P19 dans le modèle ARN (Marquez *et al.*, 2006) correspondent approximativement à la crête.

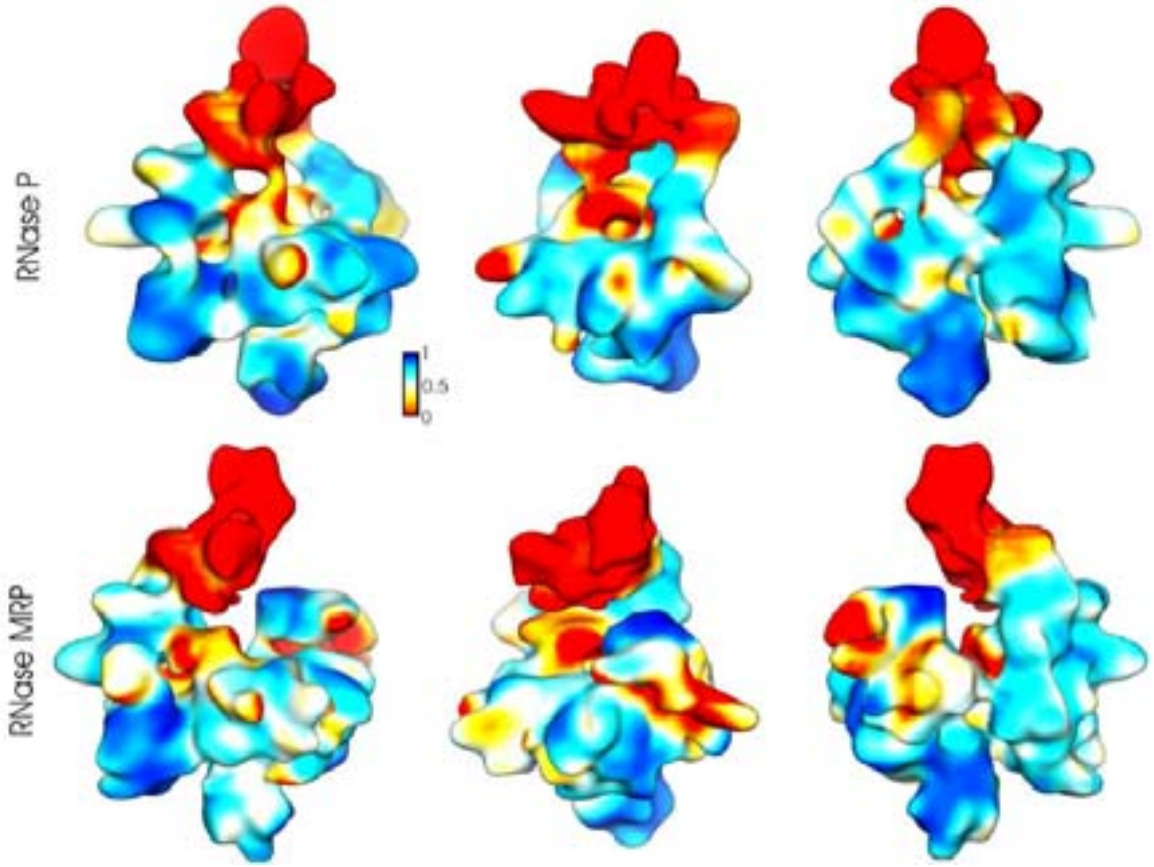


Figure 24. Représentation de la surface de la carte de la RNase P (panel du haut) et de la RNase MRP (panel du bas) colorée selon la corrélation locale entre les modèles atomiques positionnés et les cartes de microscopie électronique de la Figure 19.
(La comparaison inclut Pop3 dans la position indiquée dans la **Figure 23** par une flèche)
La légende des couleurs pour la corrélation est indiquée dans le panel supérieur. Une couleur blanche-bleue représente des corrélations croisées supérieures à 0.5, et les couleurs blanches-jaune-rouge des corrélations croisées inférieures à 0.5.
Les modèles ne présentent pas de densité pour Pop1, pour laquelle il n'existe pas de structure atomique. C'est pourquoi la corrélation est de zéro dans la partie supérieure des particules.

Dans le modèle ARN, les 30 bases du renflement de la boucle-hélice P3 n'ont pas pu être modélisées. Cependant, selon le positionnement du modèle ARN dans la carte 3D, la partie de l'hélice P3 (P3a) conduisant vers le renflement manquant pointe vers le module arrière des RNases P et MRP où une densité vacante se présente sous la forme d'une boucle vers le pied du domaine de base (**Figure 23.B**, *). Il est probable qu'au moins une partie de cette densité corresponde à la boucle-hélice P3.

Par ailleurs, dans ce modèle de la composante ARN, 50 bases de l'hélice P12 sont absentes à la fin de l'hélice P10/P11. Néanmoins, selon notre positionnement des hélices P10/P11, l'hélice P12 pointe directement vers le bras du module de base (**Figure 23.A**, panel 2, +), suggérant que l'ARN forme au moins une partie du bras du module de base. En accord avec cette hypothèse, le positionnement du domaine de spécificité de la composante ARN du complexe bactérien (Reiter *et al.*, 2010) dans notre carte 3D a montré l'étroite ressemblance des formes malgré les différences dans la structure secondaire (Hipp *et al.*, 2012).

2.2.2.3. Pop4 et Rpr2/Snm1 stabilisent le domaine S de la sous-unité ARN

Il a été montré par des expériences de double hybride et de copurification (GST-pulldown) que Pop4 interagit fortement avec Rpr2 aussi bien chez les eucaryotes (Houser-Scott *et al.*, 2002; Aspinall *et al.*, 2007) que chez les archées (Hall *et al.*, 2004; Kifusa *et al.*, 2005). De plus, lors de la résolution de la structure en solution de ce sous-complexe chez l'archée *P. furiosus* par spectroscopie de résonance magnétique nucléaire, il s'est avéré que la liaison du sous-complexe aRpr2-aPop4 à la composante ARN protège principalement les régions des hélices du domaine S (P9, P10, P11 et P12) (Xu *et al.*, 2009). Notre positionnement du modèle ARN de *S. pombe* place le sous-complexe aRpr2-aPop4 de *P. horikoshii* (Honda *et al.*, 2008) près du bras du module de base où une densité est laissée vacante par le modèle ARN sous la plaque basale (**Figure 23.A**, panel 2). Cela est cohérent avec nos expériences de localisation de Rpr2 mais incompatible avec les études qui rapportent une interaction de Pop5 avec Pop4 (Houser-Scott *et al.*, 2002; Hall *et al.*, 2004; Welting *et al.*, 2004). Une interaction entre Pop5 et Pop4 placerait le sous-complexe Pop4-Rpr2 près du module arrière, mais nous n'avons pas observé de densité suffisante dans cette région pour recevoir la structure de aPop4-aRpr2. Le sous-complexe Pop4-Rpr2 est donc plus susceptible d'être positionné près du bras du module de base. Cependant, nos expériences de marquage par anticorps ne nous autorisent pas à distinguer sans équivoque des positions proches sur le module arrière ou adjacentes au bras du domaine de base.

2.2.2.4. Pop3 contribue au bras du module de base par liaison à l'hélice P12

Dlakic a montré par analyse itérative avec PSI-BLAST que la structure secondaire de Pop3 est homologue à la protéine archée L7Ae (Dlakic, 2005) membre de la RNase P archée (Cho *et al.*, 2010). Il se présente sous la forme d'un domaine globulaire constitué de feuillets β pris en sandwich par deux séries d'hélices α. Les protéines contenant ce domaine sont connues pour lier une structure ARN secondaire conservé appelée motif kink-turn (K-turn) (Klein *et al.*, 2001) qui est prédit dans l'hélice P12 de la composante ARN de RNases P et MRP de nombreuses espèces (Rosenblad *et al.*, 2006). De plus, la liaison de Pop3 à l'hélice P12 a été mise en évidence par des expériences de liaison par UV (Pluk *et al.*, 1999) et de copurification (GST-pull-down) (Welting *et al.*, 2004). Selon notre positionnement de la composante ARN, l'interaction avec l'hélice P12 placerait Pop3 dans la région en forme de bras du module de base, qui serait assez large pour contenir Pop3 en tenant compte de la présence de l'hélice P12 issue de la plaque basale (**Figure 23.B**, flèche). Nous avons donc essayé de positionner la structure atomique de la protéine L7Ae (Hamma et al 2004), homologue archée de Pop3. Cependant, la résolution de nos cartes ne permet pas de déterminer sans ambiguïtés l'orientation exacte de Pop3, c'est pourquoi nous ne l'avons pas inclus dans la **Figure 23.B** mais seulement située avec une flèche.

2.2.2.5. Pop6 et Pop7 forment le domaine au pied du module de base

Une interaction spécifique entre l'hélice P3 du noyau ARN et l'hétérodimère Pop6-Pop7 a été mis en évidence par plusieurs équipes (Pluk *et al.*, 1999; Perederina *et al.*, 2007; Welting *et al.*, 2007). Notre étude a mis en évidence que Pop6 est localisée dans le pied du domaine de base ce qui est cohérent avec la position de l'hélice P3 issue de notre placement de la composante ARN. Récemment, la structure du sous-complexe composé de Pop6, Pop7 et de l'hélice P3 a été résolue (Perederina *et al.*, 2010b). Le positionnement automatique de ce sous-complexe sur la carte 3D segmentée donne le meilleur appariement au pied du domaine de base (corrélation croisée de 0.84 sur le segment correspondant, **Figure 24**) soulignant la validité de cette assignation. Ce positionnement met en évidence que le pied du domaine de base contient dans son entier Pop6, Pop7 et l'hélice P3 (**Figure 23.A**, panel 3). De plus, nous avons obtenu le meilleur positionnement avec l'hélice P3 pointant vers le module arrière à proximité des 30 nucléotides correspondant au renflement de l'hélice P3 non modélisée, ce qui confirme notre placement de la composante ARN de *S. pombe* dans la plaque basale.

2.3. Conclusions

Ce travail a permis de déterminer les structures des RNases P et MRP eucaryotes à une résolution intermédiaire et de localiser 6 des 10 sous-unités de ces complexes.

La superposition de structures atomiques des protéines et d'une modélisation d'un noyau ARN sur la carte de microscopie électronique nous a permis de construire un modèle pseudo-atomique des complexes RNases P et MRP. Ce modèle est cohérent avec les données des expériences de protection des sous-unités protéiques sur la composante ARN (Xu *et al.*, 2009; Honda *et al.*, 2010), nos expériences de marquage aux anticorps et le volume des structures à haute résolution. Pour la RNase P, ce modèle prend en compte environ 75% du volume observé (**Figure 23.B**), seuls manquent Pop1 dans le module de tête, des domaines non modélisés de l'ARN et certaines protéines (Pop8, Snm1, Rmp1). Il permet pour la première fois de proposer un modèle de l'architecture de RNases P et MRP eucaryotes.

Le modèle avec nos cartes des RNases P et MRP propose une meilleure compréhension du rôle des composants protéiques dans les particules. De plus, les différences structurales entre RNase P et MRP suggèrent des mécanismes d'identification du substrat différents pour ces deux complexes.

PUBLICATION

relative au travail de recherche

1. Préambule

La RNase P est une endonucléase qui coupe la séquence 5' leader des ARNt prématures. Holoenzyme très conservée, elle est constituée d'une sous-unité ARN qui forme le noyau catalytique et d'une composante protéique dont le nombre de sous-unité est variable : une protéine chez les bactéries, quatre à cinq protéines chez les archées et d'au moins neuf protéines chez les eucaryotes.

Les eucaryotes possèdent également une autre endonucléase, la RNase MRP dont la composition est proche de la RNase P tant au niveau ribonucléique que protéique. Ainsi, ces complexes partagent chez *S. cerevisiae* huit protéines (Pop1, Pop3-Pop8, Rpp1). Seule la protéine Rpr2 est spécifique de la RNase P tandis que Snm1 et Rmp1 sont propres à la RNase MRP. Toutes deux localisées dans le nucléole, la RNase MRP a plus particulièrement pour substrat un précurseur de l'ARNr 5.8S, des ARN impliqués dans la réplication de l'ADN mitochondrial et un ARNm, CLB2, dont la dégradation participe à la régulation du cycle cellulaire.

De nombreuses études se sont focalisées sur la résolution de la structure de la composante ARN, sous-unité catalytique pouvant lier directement les ARNt prématures. Les analyses des séquences des composantes ARN des RNases P et MRP suggèrent que ces ARN forment des structures secondaires similaires et s'assemblent en structures tridimensionnelles apparentées. La structure de plusieurs sous-unités protéiques de la RNase P a été résolue : les homologues archées de Pop4, Rpr2, Rpp1, Pop4 et Pop3 et le complexe hélice ARN P3-Pop6-Pop7 chez la levure. De plus, la structure de la RNase P bactérienne avec son substrat a été résolue récemment. Cependant, les structures des RNases P archées et eucaryotes avec une composante protéique plus complexe restent inconnues. De plus, le rôle exact des protéines additionnelles dans les holoenzymes eucaryotes reste peu compris, bien que leurs structures soient résolues. De nombreuses interactions protéine-protéine ont été rapportées mais varient selon les études, seules les interactions Pop4-Pop5, Pop4-Rpr2, Pop5-Rpp1, Pop6-Pop7, Pop1-Rpp1 et Pop1-Pop4 sont rapportées de façon répétée.

Malgré toutes ces informations, il n'existe pas de modèle de l'architecture des RNases P et MRP eucaryotes, principalement à cause d'un manque d'information structurale sur ces holoenzymes.

2. Publication

Hipp* K., Galani* K., **Batisse* C**., Prinz S. and Böttcher B. Modular architecture of eukaryotic RNase P and RNase MRP revealed by electron microscopy. *Nucleic Acids Research*, 2012, vol.40 (7), p.3275-3288.
* authors contributed equally

Dans cette étude, nous proposons une méthode originale et spécifique pour purifier la RNase P et la RNase MRP de *S. cerevisiae*. Grâce à la microscopie électronique et au traitement d'images, nous avons déterminé la première structure de ces deux complexes ribonucléoprotéiques à une résolution d'environ 1.5 - 1.7 nm. Ces structures révèlent une architecture modulaire des RNases P et MRP eucaryotes où les protéines stabilisent la composante ARN et contribuent à l'édification de cavités et de conduits entre les modules. Des spécificités structurales sont localisées en des positions stratégiques pour l'identification du substrat. De plus, ces sites sont aussi ceux où les différences structurales entre les RNases P et MRP sont les plus prononcées, soulignant l'importance de l'adaptation de ces régions aux différents substrats coupés par ces deux holoenzymes.

Cette étude a constitué le fils conducteur de mon travail dans l'équipe du Dr. Böttcher durant quatre ans. J'ai ainsi créé toutes les souches de *S. cerevisiae* nécessaires à la purification des complexes mais aussi à la localisation des sous-unités protéiques. J'ai aussi développé et réalisé la purification des complexes avec et sans anticorps, choisi et réalisé les tests d'activité enzymatique, préparé les échantillons et collecté les données pour l'analyse en microscopie électronique. En outre, j'ai activement participé à la publication des résultats en rédigeant la partie "Matériel et Méthode", en participant à la mise en forme de la figure 1, en réalisant la figure S4 et en contribuant à la rédaction de la partie "Discussion".

Published online 13 December 2011 Nucleic Acids Research, 2012, Vol. 40, No. 7 3275–3288
doi:10.1093/nar/gkr1217

Modular architecture of eukaryotic RNase P and RNase MRP revealed by electron microscopy

Katharina Hipp[1,2], Kyriaki Galani[2], Claire Batisse[2], Simone Prinz[2] and Bettina Böttcher[1,2,*]

[1]School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JR, Scotland, UK and [2]Structural and Computational Biology Unit, EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Received July 8, 2011; Revised November 18, 2011; Accepted November 21, 2011

ABSTRACT

Ribonuclease P (RNase P) and RNase MRP are closely related ribonucleoprotein enzymes, which process RNA substrates including tRNA precursors for RNase P and 5.8 S rRNA precursors, as well as some mRNAs, for RNase MRP. The structures of RNase P and RNase MRP have not yet been solved, so it is unclear how the proteins contribute to the structure of the complexes and how substrate specificity is determined. Using electron microscopy and image processing we show that eukaryotic RNase P and RNase MRP have a modular architecture, where proteins stabilize the RNA fold and contribute to cavities, channels and chambers between the modules. Such features are located at strategic positions for substrate recognition by shape and coordination of the cleaved-off sequence. These are also the sites of greatest difference between RNase P and RNase MRP, highlighting the importance of the adaptation of this region to the different substrates.

INTRODUCTION

Ribonuclease P (RNase P) is an endonuclease that cleaves the 5′-leader sequence of pre-tRNAs [for review see (1)]. RNase P is conserved between all taxonomic kingdoms, demonstrating its early appearance in evolution. The complex consists of an RNA subunit [Rpr1 for *Saccharomyces cerevisiae*, (2)] that forms part of the catalytic core and protein components of variable size. Bacterial RNase P has a single protein component, whereas there are four to five protein subunits in archaea and at least nine in the eukaryotic system [Pop1, Pop3–Pop8, Rpp1, Rpr2 in *S. cerevisiae* for review see (3)]. In the eukaryotic complex, these proteins comprise more than half of the mass of the ribonucleoprotein (RNP) particle, which has an approximate total mass of 410 kDa.

In addition to RNase P, eukaryotes possess the mitochondrial RNA processing ribonuclease [RNase MRP, (4–5)]. RNase MRP has a related RNA core [Nme1 for *S. cerevisiae*, (6)] and shares eight protein subunits with RNase P [Pop1, Pop3–Pop8, Rpp1 in *S. cerevisiae* (7)]. Only two subunits are specific for RNase MRP, Rmp1 (8) (24 kDa) and Snm1 (9) (23 kDa), the latter of which is homologous to Rpr2 in RNase P. RNase MRP has a wider range of known substrates than does RNase P. As its name implies, it was first described as a mitochondrial complex that is involved in processing of an RNA primer required for initiation of DNA replication (4). Subsequent fluorescence studies revealed that RNase MRP, like RNase P, is mainly localized in the nucleolus (10–12) and is involved in processing of pre-mature 5.8 S rRNA (13–15). But it is also involved in regulation of cell cycle by cleavage of specific mRNAs (16). Similar to RNase P, RNase MRP is essential for the survival of the cell.

Many investigations have focused on elucidating the structure of the RNA core, which consists of a catalytic domain (C-domain) and a specificity domain (S-domain), which can bind pre-tRNA directly. Sequence analysis suggests that the catalytic RNA cores of RNase P and RNase MRP form similar secondary structures (17–20), and assemble into related 3D core structures (21–25). Structures have been determined for several protein subunits of RNase P; the archaeal homologs of Pop3, Pop4, Pop5, Rpp1 and Rpr2 (26–34), two eukaryotic proteins Pop6 and Pop7 (35), and the bacterial protein (36). This is complemented by a recent structure of the whole bacterial complex (24) with bound tRNA. However, the structures of the archaeal and eukaryotic particles, with more complex protein components,

*To whom correspondence should be addressed. Tel: +44 131 650 5699; Fax + 44 131 650 8650; Email: bettina.boettcher@ed.ac.uk
Present address:
Kyriaki Galani, Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

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

© The Author(s) 2011. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

are still unknown. Thus, the exact role of the additional proteins in these complexes is still poorly understood, despite most of the individual protein structures being known.

Analysis of scans of gels of RNase P/MRP stained with SYPRO ruby provide additional information on the relative subunit composition suggesting that most protein subunits are present in multiple copies (8). These data are complemented by protein–protein interaction studies in different species (37–41), which among other interactions repeatedly report interactions between homologs of Pop4–Pop5, Pop4–Rpr2, Pop5–Rpp1 and Pop6–Pop7. Apart from the Pop4–Pop5 subcomplex, all of the predicted binary complexes have been crystallized either from archaea or from yeast (27,29,35).

Even with this wealth of information, there is still no consensus model for the architecture of eukaryotic RNase P and RNase MRP, due mainly to a lack of structural information on the holoenzymes. Therefore, we have used electron microscopy (EM) and single particle image processing to determine structures of both complexes at a resolution of ~1.5–1.7 nm. These structures reveal the modular architecture of RNase P and MRP and show differences in the substrate-binding cavity.

MATERIALS AND METHODS

Yeast strains

Genomic integration of the HA tag (HIS3MX6-marker) was performed to create fusion proteins of Rpr2 (HA tag C-terminally) and Snm1 (HA tag N-terminally) as described (42) into the *S. cerevisiae* strain YSW1 (MATa POP4::TAPTAG::TRP1ks pep4:LEU2 nuc1::LEU2 sep1::URA3 trp-1his3–11,15 can-100 ura3–1 leu2–3,112) described previously (9). Snm1 was tagged N-terminally due to frequent C-terminal degradation of Snm1. Presence of the HA tag in the fusion proteins was confirmed by western blot. Genomic integrations of the green fluorescent protein (GFP) tag were performed on above described Rpr2-HA strain by tagging the C-terminus of Pop5, Pop6, Pop8 or Rpp1 with yeGFP tag (*hphNT1* marker) and N-terminus of Rpr2-HA with promoter Gal1 : yeGFP tag (*natNT2* marker) as described (43). Presence of the GFP tag in the fusion proteins was confirmed by western blot.

Affinity purification

Cells, grown in 12 l of YPD, were harvested at OD 3.0–3.5 and disrupted with glass beads in lysis buffer [50 mM Tris–HCl pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, protease inhibitor FY (Serva), 0.0375% NP40] using a Pulverisette 5 (Fritsch, three cycles milling with power set to 350 for 4 min, 1 min break between cycles). TAP-tagged Pop4 protein was purified from cell lysate as described previously (44) until the step of the TEV eluate, which was then loaded on an HA affinity matrix (Roche) and incubated for 5 h at 4°C. The bound fraction was eluted via competition with the HA peptide (Roche) for 30 min at 37°C. The HA eluates were concentrated, analyzed using NuPAGE 4–12% gradient or 12% SDS–PAGE (Invitrogen) stained with colloidal Coomassie (Sigma) and verified by mass spectrometry. Furthermore, the separation of RNase P and RNase MRP was tested by quantitative RT–PCR.

For antibody labeling, the TEV eluate was split into two halves and one half was processed as described before. To the other half, anti-GFP monoclonal antibodies (Roche, 40 μg/ml final concentration) were added after 1 h incubation on the HA matrix and further incubated for 4 h.

Mass spectrometry

Mass spectrometry using tryptic digest from Coomassie-stained SDS–PAGE was performed as described previously (45). Proteins were identified using Mascot (Matrix Science) and the MSDB protein database.

Quantitative real-time reverse-transcription PCR

The HA eluates of Pop4–TAP Rpr2–HA and Pop4–TAP HA–Snm1 were treated with Proteinase K (10 μg/ml) to digest the proteins. The RNAs were then extracted by standard phenol/chloroform extraction (46). Contaminant DNA was digested by DNase I treatment.

Nme1 RNA (RNase MRP) and Rpr1 RNA (RNase P) were reverse-transcribed using Transcriptor First Strand cDNA Synthesis Kit (Roche) at 60°C to improve the efficiency of the reaction by unfolding the RNA. The primers used were 5′-TCCATTGGGTTACTCGAT CC-3′ for Nme1 RNA and 5′-TGGAACAGCAGCAGT AATCG-3′ for Rpr1 RNA.

qPCR was performed using the SYBR®GREEN PCR Master Mix (Applied Biosystem) following the manufacturer's instructions. Reactions were performed in triplicate from three different purifications of each complex using an ABI 7500 according to these conditions: 50°C for 2 min and 95°C for 10 min followed by 40 amplification cycles (95°C for 15 s and 60°C for 1 min). A dissociation stage was performed to check the specificity of the reactions. Pairs of primers (5′-CGGTACTGGATTCCGTTTGT-3′ and 5′-AGCACGGGAAAGAGCAATC-3′ for Nme1 cDNA and 5′-CCGGGTTAATGTCGCTTTT-3′ and 5′-CGGTCGGTAAAGACTGGTTC-3′ for RPR1 cDNA) were designed by Primer3. The three pairs of primers were chosen with care taken that they did not anneal the regions of secondary structure of target molecules predicted by Mfold (47). Primer efficiencies were tested as described previously (48).

The ratio of the RNA subunits of RNase P (Rpr1) and RNase MRP (Nme1) was calculated as follows:

$$[\text{Rpr1}]/[\text{Nme1}] = 2^{(\text{Ct Rpr1} - \text{Ct Nme1})} = 2^{\Delta ct}$$

Cleavage assay

Six micrograms of premature tRNA were incubated with 3 μg affinity purified RNase P (approximate molar ratio 20:1) in buffer (50 mM Tris–HCl pH 8.0, 100 mM NH_4Cl, 10 mM $MgCl_2$ or 2 mM EDTA) for 30 min at 37°C (49). Reactions were stopped by adding 9 M urea, 1 mM EDTA, 1 mg/ml bromophenol blue in 1× TBE (89 mM Tris, 89 mM boric acid; 2 mM EDTA). The mixtures were denatured for 1 min at 95°C, and separated on a

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

12% polyacrylamide / 7 M urea gel. The gel was stained with 0.1% toluidine blue.

Pre-tRNA synthesis

Saccharomyces cerevisiae genomic DNA was used as the initial PCR template to amplify the tRNA Tyr gene (GenBank accession number J01380) with primers described in (50) and the tRNA Glu gene [tE(CUC)D GenBank accession number X06132.1] with primers ptRNA-Glu_F (5′-GTAATACGACTCACTATAGAGG TCAAACTATCACCAAGGTACA-3′) and ptRNA-Glu_R (5′-AAAAAGAAACTCCGAAGCGG-3′). The PCR products were transcribed using T7 Transcription kit (Fermentas, #K0412) to produce pre-tRNA^{Tyr} and pre-tRNA^{Glu}, respectively, following the manufacturer's instruction.

Electron microscopy

Grids were prepared by sandwich negative staining (51). For random conical tilt reconstruction, pairs of tilted (−55°) and untilted fields were recorded on a Philips CM200 FEG on a 2kx2k CCD (TVIPS GmbH) at a nominal magnification of 27 500 (5.2 Å/pixel) at room temperature. Micrographs were recorded under low dose conditions. Tracking of the area of interest was done manually.

For high-resolution image analysis, grids were prepared as outlined above and frozen in liquid nitrogen immediately after drying. Thereafter, grids were kept at liquid nitrogen temperature. Grids were transferred to the CM200 FEG with a Gatan 626 cryo-transfer holder. Micrographs were recorded on CCD at a nominal magnification of 66 000 (2.2 Å at specimen level) at 200 kV under low dose conditions. 2006 micrographs were collected for RNase MRP and 1861 micrographs for RNase P, respectively.

For antibody labeling, grids were prepared by sandwich negative staining (51), and microscopy and imaging were performed as described for random conical tilt reconstruction.

Image processing

Initial maps of RNase P and RNase MRP were calculated by random conical tilt reconstruction using Spider (52) for particle selection and windowing and Imagic 5 (53) for alignment, classification and reconstruction. The relative orientations of some of the random conical tilt maps were determined and used to calculate a new map from the tilted particles contributing to the successfully aligned maps. The maps were further refined by projection matching of the tilted and untilted images (Supplementary Figure S1 and Supplementary Methods).

These refined maps of RNase P and RNase MRP were used to generate starting references for iterative refinement by projection matching (Spider) of the larger datasets of negatively stained, frozen RNase P (42 009 particle images) and RNase MRP (59 400 particle images). For further processing, the effects of the contrast transfer function were reduced by appropriate phase reversal (Supplementary Methods). In addition, for each complex, the data was processed in 20 defocus groups; for each defocus group, a map was calculated by weighted back projection using Spider. The maps were added to give the final map (Supplementary Data).

Maps of RNase P and RNase MRP were similar but also had distinct differences. In order to test whether these differences were linked to variability in the image processing or to true differences between RNase P and RNase MRP, both maps were combined to a common starting reference for iterative refinement of both complexes. Images of RNase P and RNase MRP were processed separately. In the subsequent refinement, the current best map of each complex was used as a new reference. Refinement was continued until most orientations of particle images did not change and no further improvement in resolution was observed (see Supplementary Data for more details on iterative refinement). Refinement converged to two distinct maps, which showed similar key-differences as previously observed by starting with individual random conical tilt maps for RNase P and RNase MRP.

Furthermore, we assessed the heterogeneity of the particle populations for RNase P and RNase MRP separately following two approaches. (i) We used Xmipp Maximum-likelihood 3D-sorting (54) to sort particle images of either RNase P or RNase MRP into five classes. (ii) We used IMAGIC 4D, to generate 10 maps from particles randomly assigned to the 10 groups and then iteratively refined in projection matching and supervised classification. Both approaches showed small differences between the maps, but no indication for major intrinsic flexibility or heterogeneous composition of the particles in the population. This justified the refinement of particle images against one representative map in retrospect.

To check whether the orientations were assigned consistently, we sub-classified particle images, which were assigned with similar Euler angles using multivariate statistical analysis (see Supplementary Methods and Supplementary Figure S2). Visual inspection showed that the averages of the sub-classes had similar features as the reference projections (Figure 1D and Supplementary Figure S2), further highlighting that the majority of particle images was assigned with a correct overall orientation.

Resolution of the maps was estimated by Fourier shell correlation [FSC, (55)]. The FSC was 0.5 (56) at a spatial frequency of $1/1.7\,nm^{-1}$ for RNase P and $1/1.5\,nm^{-1}$ for RNase MRP [FSC = 0.14 (57) at $1/1.6\,nm^{-1}$ for RNase P and $1/1.4\,nm^{-1}$ for RNase MRP; Figure 1], respectively.

Image processing of antibody labeled particles

For image analysis in antibody labeling experiments, particle images were selected from micrographs using Boxer, which is part of EMAN (58). Subsequent alignment and classification were done with IMAGIC 5 (53) and were repeated until classes remained stable. To localize additional densities in the periphery of particles arising from bound antibodies, class averages were sub-classified using a mask focusing on the periphery of the particles.

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

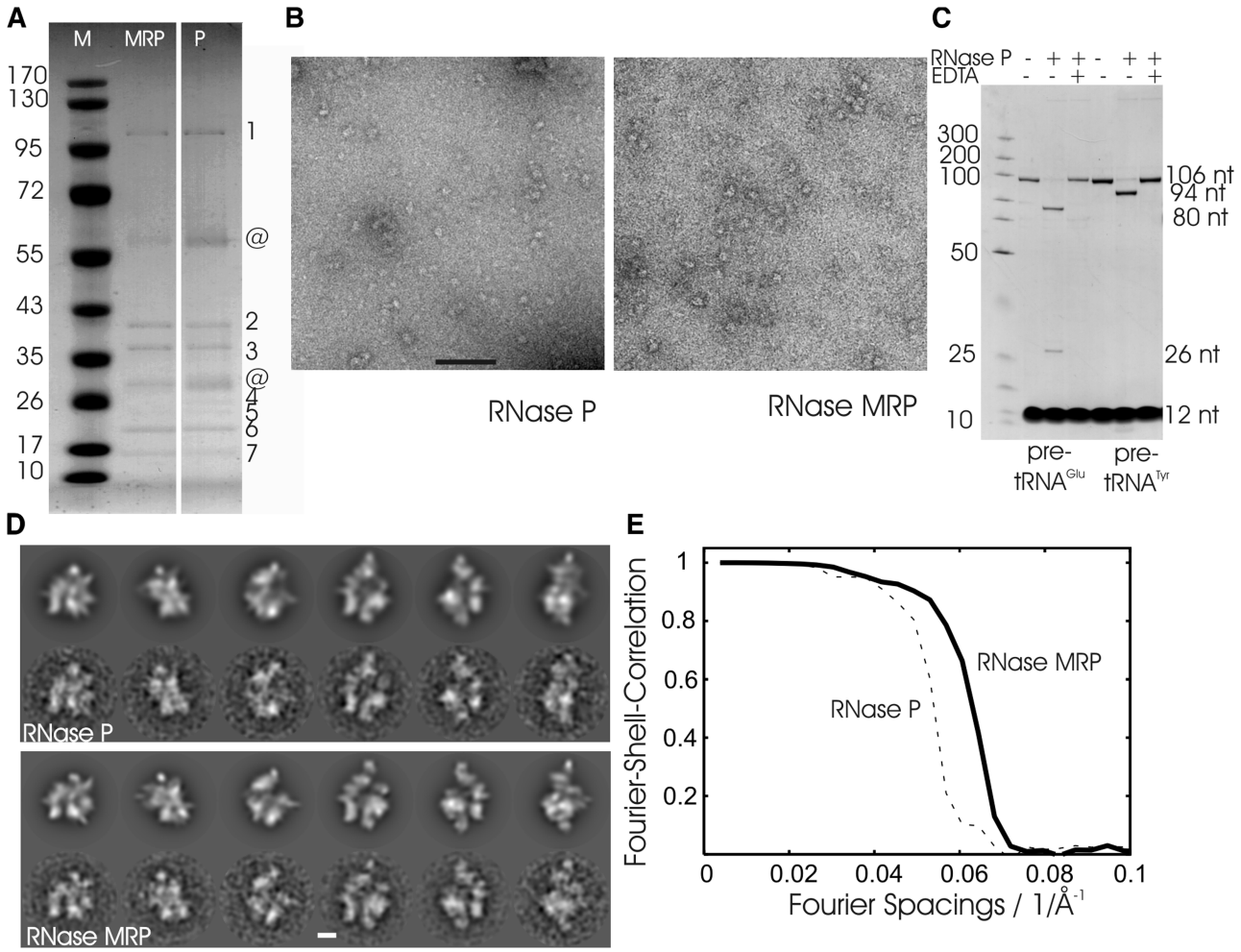


Figure 1. Characterization of the RNase P/MRP preparation. (**A**) Coomassie-stained SDS–PAGE of RNase P and RNase MRP. Left: marker with given molecular masses (Page Ruler™ Prestained Protein Ladder; Fermentas) in kDa; centre: HA eluate of Pop4–TAP HA-Snm1 (RNase MRP); right: HA eluate of Pop4–TAP Rpr2–HA (RNase P). Bands identified by mass spectrometry are indicated by numbers: (1) Pop1; (2) Pop4; (3) Snm1, Rpp1 for RNase MRP and Rpp1 for RNase P; (4) Nothing for RNase MRP and Rpr2 for RNase P; (5) Pop3, Rmp1, Snm1 for RNase MRP and Pop3, Pop1 for RNase P; (6): Pop5, Pop6, Pop7; (7): Pop8; @ contaminations. (**B**) Micrographs of negatively stained RNase P (left panel) and RNase MRP (right panel) from representative preparations. Scale bar: 100 nm. (**C**) *In vitro* activity assay of RNase P using two different pre-tRNAs. Leader sequences (26 nt and 12 nt for pre-tRNA^{Glu} and pre-tRNA^{Tyr}, respectively) are cleaved off by RNase P to give rise to tRNA^{Glu} (80 nt) and tRNA^{Tyr} (94 nt) with mature 5′-ends. No cleavage of pre-tRNA molecules occurs in the absence of RNase P or in the presence of RNase P with EDTA added to the reaction showing the dependency of the catalysis on magnesium ions. The reaction mixtures were separated on 12% polyacrylamide / 7 M urea gels and stained with 0.1% toluidine blue. The strong bands at the bottom are from bromphenol blue in the sample buffer. The position of marker fragments (GeneRuler™, UltraLow Range DNA Ladder, Fermentas) in bp is shown on the left. (**D**) Selected projections of the 3D map of RNase P (upper panel, upper row) and RNase MRP (lower panel, upper row) and matching class-averages calculated by multivariate statistical analysis (lower rows). Scale bar: 5 nm. (**E**) Fourier-shell-correlation for maps of RNase P (dashed) and of RNase MRP (solid).

Fitting of atomic models

Chimera (59) was used to represent results graphically and to fit atomic models. Some components [Pop6–Pop7–P3 subcomplex (35) and the archaeal homolog of the 2:2 Pop5–Rpp1 subcomplex (29)] were fitted automatically to the segmented map using the Segger option of Chimera (fit to groups of regions). Other components [model of the *Schizosaccharomyces pombe* RNA (60), the archaeal homolog of the Rpr2–Pop4 subcomplex (27) and the L7Ae protein, which is homologous to Pop3 (34)] were placed manually, taking constrains from published interaction studies and our antibody labeling into account. The manual fits were further optimized with the fit to map option of Chimera, which performs a local optimization. The fit of these components was further validated by calculating local cross-correlations between the map and the fitted components (Supplementary Figure S3).

RESULTS

RNase P and RNase MRP share a highly similar protein subunit composition. Therefore, a two-step affinity purification strategy was employed to separate both complexes. A strain with genomically integrated TAP tagged Pop4 subunit was used for tagging either Rpr2 (unique to RNase P) or Snm1 (unique to RNase MRP) with an

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

HA tag. Affinity tags on different subunits led to efficient separation of RNase P and RNase MRP as determined by SDS–PAGE followed by mass spectrometry (Figure 1A) and further confirmed by quantitative RT–PCR (see also Supplementary Figure S4). Preparations of negatively stained RNase P and RNase MRP show a homogeneous spread of elongated particles of similar size (Figure 1B). The purified RNase P was active in cleaving pre-tRNA in the presence of magnesium ions indicating the recovery of a functional enzyme (Figure 1C).

For structure determination by EM, RNase P and RNase MRP were prepared by cryo-negative staining, which keeps residual structural water and reduces flattening compared to conventional negative staining. A total of 42 104 individual particle images for RNase P and of 59 104 for RNase MRP were selected for further processing. The final reconstructions were calculated from 24 373 (RNase P) and 32 232 (RNase MRP) particle images, respectively. The longest axes of the complexes were ~20 nm. Visual inspection showed that projections of the reconstructions agreed well with class averages calculated by multivariate statistical analysis (Figure 1D) and showed no major heterogeneity in sub-classes (Supplementary Figure S2) highlighting the validity of the 3D maps. The maps had a resolution of 1.7 nm for RNase P and 1.5 nm for RNase MRP (Figure 1E). However, considering that the particles were stained, these maps only reproduce the stain-accessible surfaces of the particles but not the internal density distribution of protein and RNA.

RNase P and RNase MRP consist of three modules

Image reconstruction of RNase P and RNase MRP revealed that both enzymes share a similar architecture (Figure 2). RNase P/MRP consists of three modules, which form a cavity in their centre. The boundaries of the modules were defined by their limited contacts with neighboring modules. We refer to these modules as the base, rear and lid modules (see Figure 2A for labeling).

Base module

The base module consists of the base plate, a flat trapezoidal entity with alternating ridges and grooves, a globular foot domain underneath it and an arm domain that contacts the lid module in RNase P but not in RNase MRP (Figure 2A). The base plate forms the largest surface of the central cavity and rises to a crest close to the rear module. In RNase P, the crest occupies both sides of the entry to a channel that is connected to chambers between the base module and the rear module (Figure 2B, channel entry: open arrows). In contrast, in RNase MRP the crest is located to one side of the channel entrance, resulting in more elongated access to the chambers between the base and rear module (Figure 2B, open arrow). In RNase P, the channel traverses the complex in a straight line, whereas it is blocked in RNase MRP (Figure 2B, block: filled circle). This coincides with the channel being accessible from the outside (Figure 2B, unfilled circle) in RNase MRP and being enclosed in RNase P.

The arm domain emerges at the other side of the base plate. In RNase P the arm domain is thinner and more elongated and contacts the lid module (Figure 2A and B, solid arrow), whereas in RNase MRP the arm is more globular and is not connected to the lid module. RNase MRP has additional density close to where the arm module emerges from the base plate, which is absent from RNase P (Figure 2B, asterisk).

Lid module

The lid module is most variable between different reconstructions, which might be caused by flexibility or high susceptibility to degradation. The module consists of two domains, which we refer to as the collar and the spike domains (Figure 2A for labeling). In RNase P, the collar domain connects the lid module to the arm domain of the base module and to the rear module. In contrast, the lid module in RNase MRP is only connected to the rear module.

Rear module

The rear module consists of two similar domains, which are related by ~2-fold symmetry, suggesting a dimeric building block. The lower half of the dimeric core contributes to the channel, whereas the upper half anchors the lid module. The rear module is connected to the base module at several defined contact points. The upper half contacts the crest whereas the lower half contacts mainly the lower part of the base plate.

Localization of the subunits in RNase P/MRP

To build a pseudo-atomic model of RNase P/MRP, we wanted to fit the various high-resolution structures of homologous subunits and subcomplexes from different species as well as a structural model for the RNA subunit of *S. pombe* (60) into our EM maps. However, at the current resolution there were ambiguities where to place some of the high-resolution structures. We therefore localized several subunits of RNase P/MRP by EM and image analysis of complexes with antibody labeled subunits. For labeling we either fused GFP N-terminally (Pop1, Rpr2–HA) or C-terminally (Pop6, Pop8, Pop5, Rpp1) to the subunit of interest. In our hands the GFP tag was too small to be directly detectable in difference imaging. To increase the size we bound a commercially available monoclonal antibody to GFP. Labeling rates for the antibodies were typically <10%, which is in the range of labeling efficiency that we obtain for other complexes (61–63). Labeled complexes were observed in preferred orientations, which is possibly due to the additional GFP moiety on one of the protein subunits of the complex. The preferred orientation makes mapping of the epitopes in 3D impossible. However, labeled complexes could be identified by classification techniques and the approximate position of the GFP–antibody complex in the particles could be determined (Figure 3). The 2D distribution of labels localized Rpp1, Pop5, Pop8 and Rpr2 to the centre of the particle. The positions of Pop8 and Rpp1 were virtually superimposed. However, due to the lack of 3D information it was impossible to distinguish

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

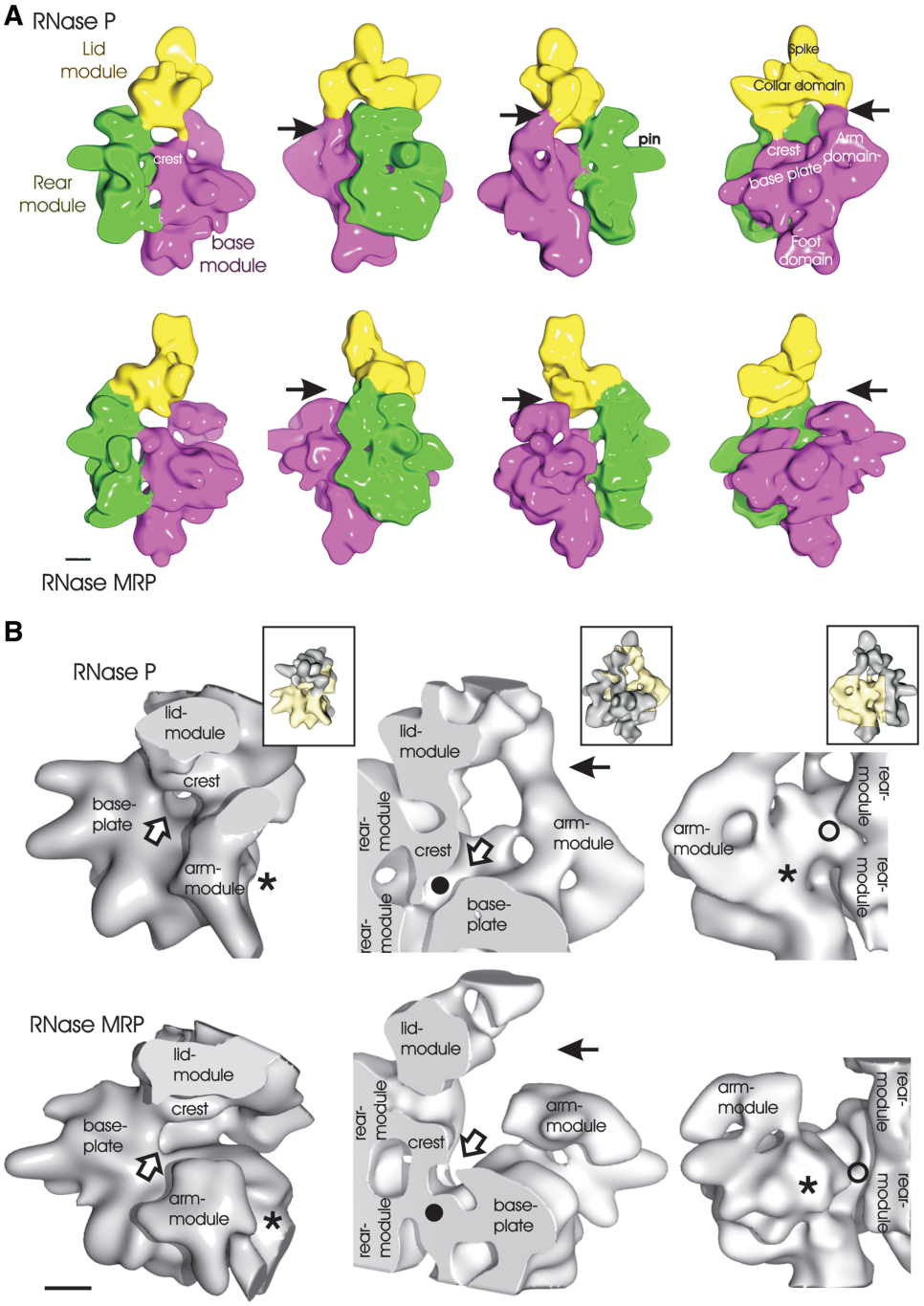


Figure 2. Surface representation of RNase P and RNase MRP. (**A**) RNase P is shown in the top panel and RNase MRP in the bottom panel. The maps are rotated anti-clockwise in 90° steps around the long molecular axis. The different modules are colored, and the names of the modules are shown in the same color: base module (magenta); rear module (green); lid module (yellow). Different landmarks in the map are indicated. The arrow points at the interface between arm domain and lid module, which are in contact in RNase P and are separated in RNase MRP. (**B**) Close-up views of selected regions of RNase P (top) and RNase MRP (bottom), which differ most in the two complexes. In the inserts the whole complex is shown in the same orientation as the respective close-up and the part of the close-up is highlighted in yellow. Left panel: the entry to the channel (open arrow) is enclosed in RNase P and open to one side in RNase MRP (open arrow). RNase MRP has additional density close to the arm module (asterisk), which is missing in RNase P (asterisk). Central panel: RNase P and RNase MRP are cut at the same position through the centre of the channel. The channel is open in RNase P (filled circle) and blocked in RNase MRP (filled circle). The arm module is connected to the lid module in RNase P and separated in RNase MRP (arrow). Right panel: the channel is protected to the outside in RNase P (unfilled circle) and accessible in RNase MRP (unfilled circle). The additional density of RNase MRP (asterisk) is close to the arm module and gives rise to a protrusion that is missing in RNase P (asterisk). Scale bar: 2 nm.

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

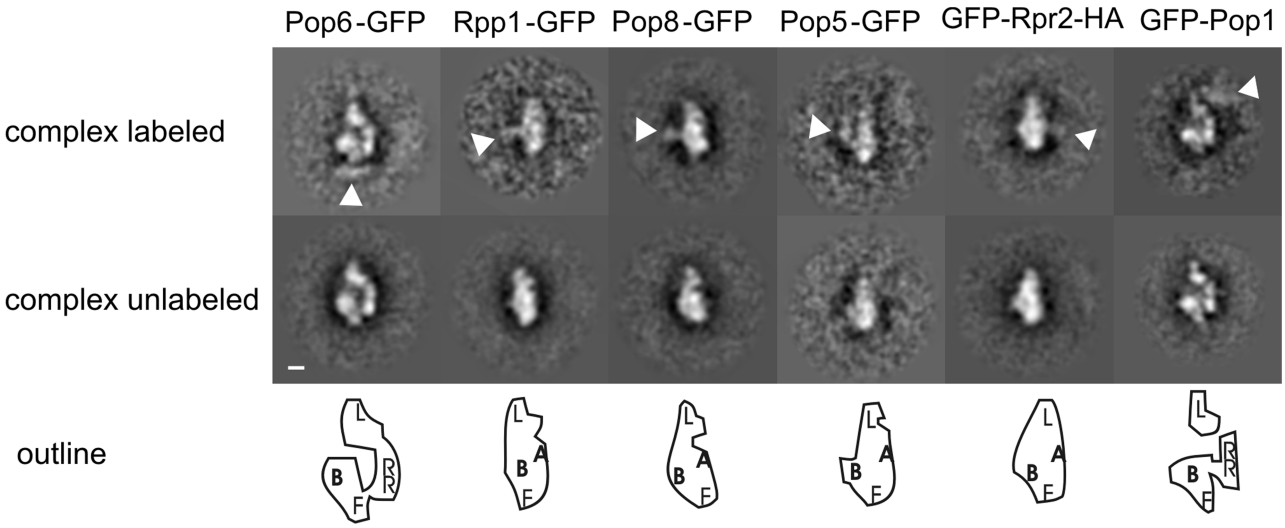


Figure 3. Localization of RNase P subunits by antibody labeling. Sub-class averages with an extra density are shown in the upper row, in contrast to sub-class averages with no extra densities shown in the middle row. White arrowheads point to the position of the extra density due to the bound antibody. Scale bar: 2 nm. The bottom row shows the outline of the class averages of the unlabeled complex above. For clarity, the outline is enlarged by a factor of 2. The approximate positions of the domains are indicated by letters: A: arm domain, L: lid module; F: foot domain; B: base plate; RR: rear module [only shown in side views (panel 1 and 6) where it does not overlap with the arm module and base plate].

whether these subunits were based in the rear-module or close to the base plate. Pop6 was located in the foot domain and Pop1 mapped to the lid module.

Pop1 has a molecular mass of 101 kDa and large enough to account for the complete lid module. Indeed, the map of the lid module was 100 000 $Å^3$ in RNase P, which is somewhat smaller than expected for a 100-kDa protein. Mass spectrometry on typical purifications showed that Pop1 was partially degraded. This was also frequently observed in western blots of C-terminally labeled Pop1–GFP using anti-GFP (data not shown). This instability of Pop1 is consistent with the structural variability that we observe for the lid module and the underestimation of the volume.

Pop5 and Rpp1 form the rear module

The archaeal homologs of Rpp1 and Pop5 form a 2:2 complex in crystals (29). At low resolution, the structure of this subcomplex closely resembles the dimeric core of the rear module (Figure 4A, panel 1). The rear module also gives the best fit in automatic fitting of the crystal structure of the Rpp1–Pop5 complex to segments of the segmented map (cross-correlation = 0.89 to the respective segment). This position is also consistent with our localization experiments that indicated that Rpp1 and Pop5 were located in the centre of the particle. Therefore, the Rpp1–Pop5 complex was fitted into the rear module. In this fit, the positively charged side of Pop5 faces the chamber between the base module and the rear module. The Rpp1–Pop5 complex accounts for most of the density in the rear module apart from the pin (see Figure 2A for definition). This feature, close to the local symmetry axis, suggests some deviation from the 2:2 organization. Possibly Pop8, which is mapped with antibody to the same position as Rpp1 (Figure 3) has replaced one of the related Pop5 proteins.

The RNA subunit forms the base plate

Binding of Pop5 and Rpp1 to the RNA subunit leads to a protection of the RNA against nucleases around the conserved catalytic region formed by stems P3, P4 and P5 (64). This puts the RNA subunit with its catalytic centre into close proximity to the rear module (see also Supplementary Figure S5 for overview), which is consistent with placing it into the base plate. The modeled RNA subunit of RNase P from *S. pombe* (60), which is considerably smaller than the RNA subunit of *S. cerevisiae*, has a similar trapezoidal shape as the base plate. The modeled RNA subunit shows a pattern of alternating ridges and grooves that closely resembled the features of the base plate. Placing the RNA model into the density of the base plate with stems P3A, P4 and P5 facing the rear module brings the catalytic core of the RNA subunit into close proximity to the channel entry and the crest (Figure 4A, panel 2). At the opposite side of the base plate, the P10/11, P9, P1 and P19 stems in the RNA model (60) approximately match the ridges.

The RNA model lacks a further 44 bases including the P3 bulge loop (see also Supplementary Figure S5). According to the fitting of the RNA model into the density map, the stem (P3A) leading towards the missing P3 bulge loop points towards the rear module of RNase P and RNase MRP, where an unaccounted density loops back towards the foot domain (Figure 4B, denoted by asterisk). It is likely that at least part of this density is accounted for by the stem leading to the P3 bulge loop.

The RNA model also lacks 50 bases of the P12 stem at the end of the P10/P11 stem (see also Supplementary Figure S5). According to the fit, the P12 stem points directly towards the arm domain (Figure 4A, denoted by plus), suggesting that RNA forms at least part of the arm domain. In agreement with this assumption, fitting the

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

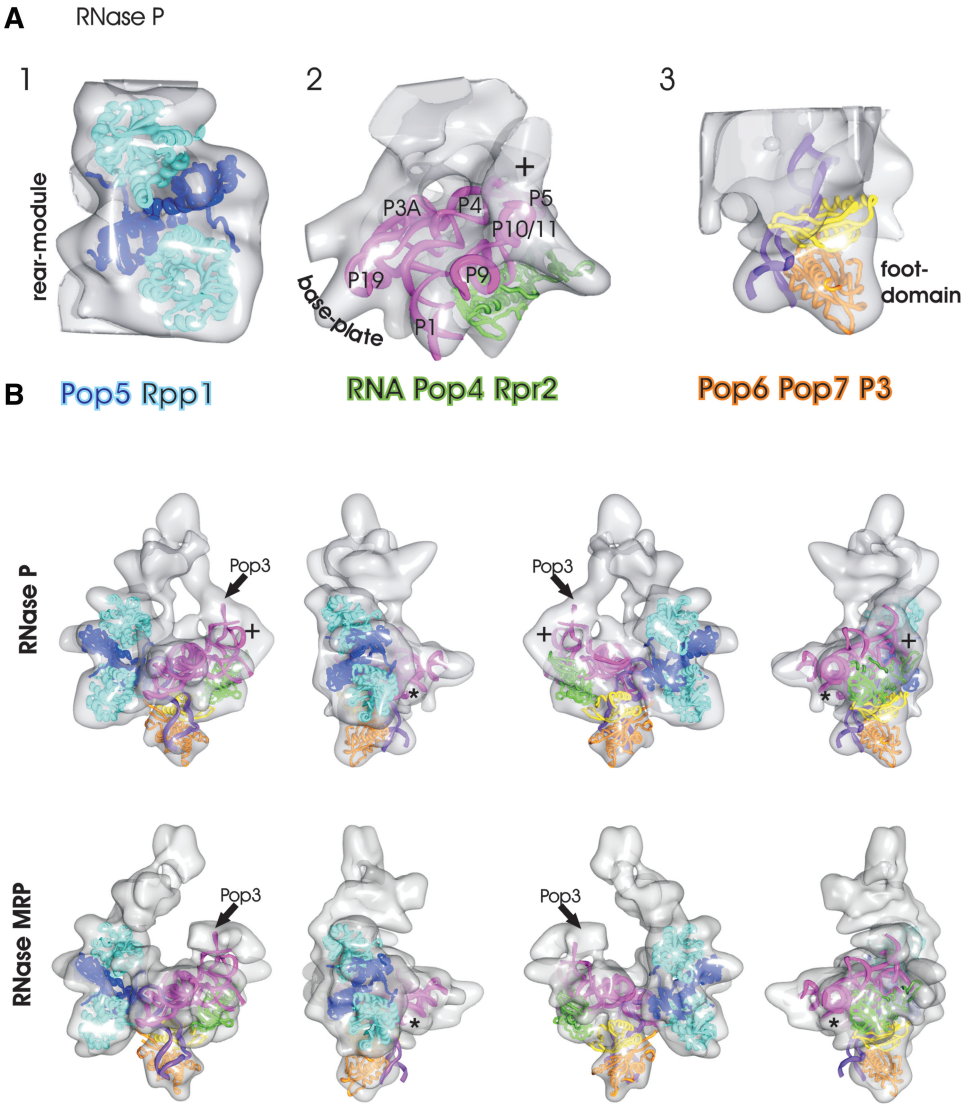


Figure 4. Surface representations of the RNase P with fitted subunits: (**A**) Panel 1: heterotetramer [2CZV (29)] of the archaeal homologs of Pop5 (dark blue) and Rpp1 (light blue) are fitted to the rear module. Panel 2: model of *S. pombe* RNA (60) (magenta) and archaeal homologs of the Pop4–Rpr2 dimer [2ZAE, green (27)] fitted to the base plate. Panel 3: Pop6 (orange), Pop7 (yellow) and P3 loop (purple) [3IAB, (35)] fitted to the foot domain. (**B**) Fitted subunits (see A for color coding) superimposed to RNase P (top panel) and to RNase MRP (bottom panel). The views are rotated in 90°-steps around the y-axes. The 'plus' highlights unaccounted density at the end of the P10/P11 stem where 50 bases of the P12 stem are missing in the RNA model. The 'asterisk' highlights an unaccounted rod-shaped density at the start of the P3 bulge loop where 30 bases are missing in the RNA model. The arrow highlights the approximate position of Pop3. The local cross correlation between the maps and the fitted models are shown in Supplementary Figure S3.

specificity domain of the RNA subunit of the bacterial complex (24) into our EM-map showed the close resemblance of the shapes (Supplementary Figure S6) despite the differences in secondary structure.

Pop4 and Rpr2/Snm1 stabilizes the S-domain of the RNA subunit

Pop4 tightly interacts with Rpr2 (38,40), and a solution structure of the archaeal homolog of this complex has been determined (64). Binding of this complex to the RNA subunit mainly protects regions in the stems of the specificity domain (S-domain, P9, P10, P11 and P12) (64). Our fitting of the RNA model of *S. pombe* places this close to the arm module (Supplementary Figure S7 position 4), where there is unaccounted density underneath the base plate (Figure 4A, panel 2). This is consistent with our localization experiments of Rpr2 but incompatible with studies that report an interaction of Pop5 with Pop4 (37–38,41). A Pop5–Pop4 interaction would place the Pop4–Rpr2 subcomplex close to the rear module, but we did not observe sufficient density in this region to accommodate the structure of the Pop4–Rpr2 homologs. The Pop4–Rpr2 complex is therefore more likely to be positioned close to the arm domain. Our labeling experiments do not allow us to distinguish unequivocally between positions close to the rear module or adjacent to the arm domain.

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

Pop6 and Pop7 form the foot domain

The P3 stem of the RNA subunit binds the Pop6–Pop7 heterodimer (65). Our labeling studies localized Pop6 of this subcomplex in the foot domain, which would be consistent with the emerging P3 stem. Recently, the structure of the whole subcomplex consisting of Pop6, Pop7 and the P3 stem has been solved (66). Automatic fitting of the Pop6–Pop7–P3 complex to the segmented map (Chimera segger option) also gave the best match for the foot domain (correlation 0.84 to the respective segment) further highlighting the validity of the assignment. Fitting this subcomplex into the foot domain showed that the Pop6–Pop7–P3 subcomplex accounts for it completely (Figure 4A, panel 3). We obtained the best fit with the P3 stem pointing towards the rear module (Supplementary Figure S7, position 1) close to the missing 44 nt in the RNA model, supporting our fitting of the modeled RNA subunit of *S. pombe* to the base plate.

DISCUSSION

We have determined the structures of eukaryotic RNase P and RNase MRP at intermediate resolution and mapped a number of subunits in these complexes. Superimposition of known protein structures and a computational model of the eukaryotic RNA subunit onto the EM structure allowed us to build pseudo-atomic models of the RNase P and MRP complexes. The model is consistent with foot-printing data between subunits and the RNA subunit (64,67), our labeling experiments and the shape of the high-resolution structures. The supporting evidence is summarized in Supplementary Table S1 and reported interactions between protein and the RNA-subunit are highlighted in the pseudo-atomic model in Supplementary Figure S7. For RNase P this model accounts for ca. 75% of the observed volume (Figure 4B), lacking only Pop1 in the lid module and individual domains of the RNA and some proteins. The model, together with our maps of RNase P and RNase MRP, offers important insights into the role of the protein components in the particles. Furthermore, the structural differences between RNase P and RNase MRP suggest a mechanism for substrate recognition by RNase P and its differences from RNase MRP.

Proteins in RNase P/MRP support the structure of the RNA

During evolution from bacteria to archaea and eukaryotes, the RNA subunit of RNase P/MRP became smaller, while additional protein components were gained. This led to the assumption that a major function of the protein subunits is stabilizing the folding of the RNA. Our model suggests that this is true for the Pop4 and Rpr2 subunits. These were acquired in archaea and retained in eukaryotes and bind underneath the arm domain, which is formed by the S-domain of the RNA subunit. In RNase MRP, Rpr2 is replaced by the larger Snm1 subunit. The S-domain also differs between RNase P and RNase MRP, much more than the rest of the RNA subunit, suggesting that binding to this area involves a specifically adapted subunit. It has been reported that Snm1 interacts with Rmp1 (68), another protein that is unique to RNase MRP and has no counterpart in RNase P. Consistent with Snm1 being larger than its homolog Rpr2 and the additional presence of Rmp1, we observe extra density close to the arm region in RNase MRP (Figure 2B, asterisk) next to the predicted location of Rpr2.

Pop3 is another protein that potentially binds to the arm domain. It is homologous to the ribosomal L7Ae/L30e protein (69), which is incorporated into the archaeal enzyme (70) and elevates the optimal temperature for catalytic activity (71). Pop3 binds to kink-turn motifs, which are predicted for the P12 stem of the RNA subunits of RNase P and MRP from various species (72). Binding of Pop3 to the P12 stem is supported by cross-link data (73) and GST-pull downs (41). An interaction with P12 would place Pop3 in the arm regions of RNase P and MRP, which are large enough to accommodate Pop3 at similar positions with respect to the base plate (see Figure 4B, arrow). However, the resolution of the maps is insufficient to unambiguously determine the exact orientation of Pop3. At the assigned position, Pop3 would be strategically placed for mediating substrate recognition and substrate binding at a site distant from the catalytic core. Such a role in enhancing and stabilizing substrate binding would be compatible with the requirement for Pop3 binding in catalysis at elevated temperatures (71).

The protein subunits Pop6 and Pop7 form a heterodimer and are specific to eukaryotes. Pop6–Pop7 binds to the P3 stem of the RNA subunit in both RNase P and RNase MRP (65) and account for the foot domain. The Pop6–Pop7 heterodimer holds the P3 stem in an orientation where it kinks back and is positioned directly underneath the base plate. This location is similar to the P15.2 stem of the bacterial B-type RNA (23) and the P18 stem of the bacterial A-type RNA (25), which also loop underneath the core of the RNA subunit. This suggests that the P15.2/P18 stem of bacteria might have been replaced by the P3 stem, together with the associated Pop6–Pop7 heterodimer, in eukaryotes. A channel is formed between the foot domain and the rear module, providing a protective environment that restricts the conformational space for the RNA substrate that is cleaved off.

Proteins contribute to substrate binding

Other subunits of RNase P/MRP are less tightly intermingled with the RNA subunit and are, therefore, likely to contribute to functions other than direct stabilization of RNA folding. Among these subunits are Pop5 and Rpp1, which form a heterotetramer (29) that comprises the rear module. Pop8 belongs to the same family as Pop5 (72) suggesting that in yeast one of the two Pop5 proteins in the heterotetramer is replaced by the smaller Pop8. This would be consistent with the slight asymmetry observed in the rear module and with our localization of Pop5 and Pop8 at similar positions adjacent to Rpp1.

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

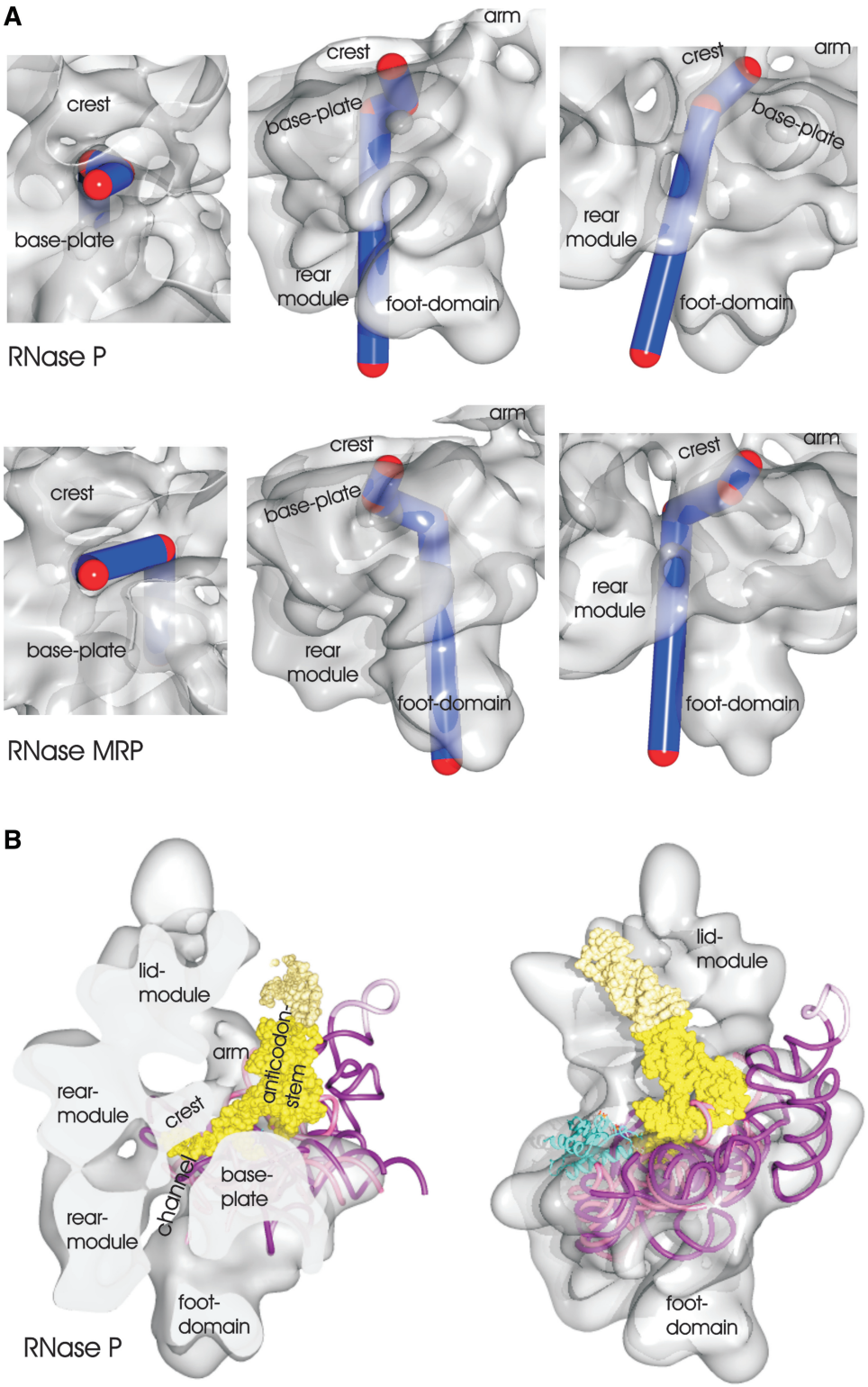


Figure 5. (**A**) Substrate channel in RNase P (upper row) and RNase MRP (lower row). Turning points in the channels are marked by red spheres (diameter 10 Å) and linked by blue lines. Different views highlight the different routes of the channels close to the catalytic domain. (**B**) Surface representation of the map of RNase P with the structure of the bacterial RNase P with bound tRNA (24) superimposed. The bacterial RNA subunit is shown in purple with the bound tRNA as yellow surface and the bacterial protein in cyan. The core-domain of the bacterial RNA subunit was used for fitting and is superimposed to the fitted model of the *S. pombe* RNA subunit (magenta). The fit also shows the likely position of the bound tRNA, which is part of the crystal structure but not included in the EM-map. The crystal structure of the bacterial complex contains engineered crystal contacts, which are shown in light yellow for the tRNA and in light-magenta for the RNA subunit. Left panel: the map and the fitted structure are clipped at the central channel of RNase P. According to the fit of the bacterial RNA subunit to the map, the 3′-end of the pre-tRNA enters the channel. Right panel: the view is rotated by 90° around the long axis in respect to the view on the right.

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

The rear module is comprised of Pop5 and Rpp1, which were acquired in archaea. The rear module binds to the RNA subunit, forming extensive cavities and channels between the rear and base modules (base plate and foot-domain). As for the foot domain, the cavities are strategically positioned to contribute to coordinating the cleaved RNA sequence. Pop5–Rpp1 in eukaryotes and the protein in bacteria fulfill a similar function in enhancing catalysis (67). This suggests that the rear module is a replacement for the bacterial protein, but provides more sophisticated coordination of the cleaved-off leader sequence.

Interestingly, the system of channels and cavities is different in RNase P and RNase MRP. Whereas RNase P has an enclosed channel, RNase MRP provides a slit that is open towards the side of the S-domain. Accordingly, the possible exit path for the cleaved-off sequence varies between RNase P and RNase MRP (Figure 5A).

The Pop1 subunit does not directly contact the RNA but forms the lid module and delineates the substrate-binding site. In RNase P, the lid module contacts the arm, forming an enclosed substrate-binding site with a specific shape that resembles a tRNA. In RNase MRP, the lid module does not contact the arm module. This could give rise to a greater flexibility, allowing the substrate-binding cavity to accept a wider range of substrates.

The tRNA is coordinated by the arm and lid modules

Superimposing the structure of a bacterial RNAse P complex with bound tRNA (24) onto our model of RNase P (Figure 5B) shows the likely contributions of the arm domain and the lid module to substrate binding for tRNA. The tRNA in the fitted ternary complex is oriented such that the 3′-end of the tRNA is pointing towards the rear module and enters the channel between the rear module and the base module (Figure 5B). The fit suggests that the channel entrance in the eukaryotic RNase P is some 2 nm away from the residues identified as being part of the active site in the bacterial enzyme. In the bacterial enzyme the 5′-leader sequence is coordinated by the bacterial protein, which has no counterpart in the EM-density map of the eukaryotic RNase P. This suggests that the 5′-leader sequence is coordinated in a different way. Possibly, it also enters the channel and interacts with the extensive system of channels and cavities between base module and rear module.

Further from the catalytic site the structural match is less good, but it is conceivable that the arm domain, where we placed Pop3, coordinates the anticodon stem of the pre-tRNA opposite to the catalytic site. In addition, the lid module restricts the maximal possible length of the anticodon stem of the tRNA (Figure 5B). This is consistent with a substrate binding pocket that not only recognizes specific sequences but also identifies the characteristic L-shape of the tRNA.

In the bacterial complex, the S-domain, which we think forms part of the arm domain in the eukaryotic complex, points away from the anticodon stem of the tRNA (24) and is not accommodated by the EM-density (Figure 5B). This suggests that the S-domain in the bacterial complex has a more open conformation than in the eukaryotic complex. However, it is difficult to understand, why such an extensive structure is acquired in the S-domain, if it does not directly contribute to substrate binding or catalysis. We think that in the bacterial complex the S-domain was possibly stabilized in the open conformation by the crystal contacts, which were engineered between the tip of the anticodon arm of the tRNA and the end of the P12-stem of the RNA-subunit. Thus the bacterial RNA-subunit could have a greater flexibility adopting an open conformation to facilitate the initial, loose substrate binding and a closed conformation, similar to what we observe in the eukaryotic complex, for further strengthening the substrate binding (induced fit).

According to our model of RNase P (Figure 6), the components that were specifically gained in eukaryotes (Pop1, Pop6 and Pop7) are located at the ends of the complex, where they contribute to substrate recognition and binding. This evolution might have been driven by the need to cope with the increased complexity of RNA species in the cell. We speculate that the gains provide a selection for specific shapes, which sample a larger proportion of the surface of potential substrates than in the bacterial complex.

Architecture of the complex and binary interactions

We present a map of the architecture of the complex (Figure 6), which accounts for most of the observed density. However, our model agrees only partly with the numerous interaction studies (37–38,41,68,74), since many

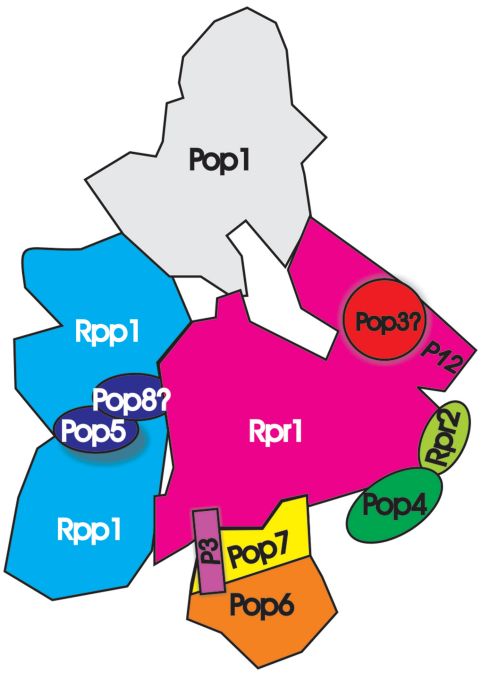


Figure 6. Cartoon of the subunit organization of RNase P. Subunits where structures could be fitted are shown in the same color as in Figure 4. Regions, which are not accounted by fitted subunits, are shown in grey.

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

predicted binary protein–protein interactions cannot be confirmed. However, the published data indicate that these small (20–40 kDa) proteins make five or more protein–protein contacts, and it is unclear if this is really the case. According to our model, no protein subunit in the complex is in direct contact with more than three other proteins. This highlights the difficulties in deducing the architecture of complexes from interaction studies alone. Furthermore, apart from Rpp1 and possibly Pop5 we have no evidence that any of the subunits of RNase P/MRP occurs in multiple copies, in contrast to estimates of the stoichiometry using SYPRO ruby (8). A potential explanation for these discrepancies might be the formation of dimers that occur in the bacterial RNase P under certain conditions (75–76) or high-order structures in the cell.

CONCLUSION

In conclusion, we have found that RNase P and RNase MRP show a modular architecture. Modules acquired in archaea and eukaryotes can be understood as evolutionary gains that contribute to substrate recognition and coordination. In eukaryotes, subtle changes in the interplay between the modules in RNase P and RNase MRP alter the shape of the substrate-binding cavities and the channel. This is in agreement with their different substrate specificities and highlights the importance of coordination of the cleaved leader sequence of the substrate.

ACCESSION NUMBERS

emd-1929, emd-1928.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online: Supplementary Table 1, Supplementary Figures 1–7, Supplementary methods and supplementary references [77–80].

ACKNOWLEDGEMENTS

The authors thank Stephanie Kronenberg for her contribution in the early steps of the purification, Stephan Kästner and Margot Scheffer for their contribution in data collection of random conical tilt data and in the initial steps of the image processing, and Tony Crowther and David Tollervey for comments on the manuscript. The authors thank the Proteomic Core Facility at EMBL-Heidelberg and the Lechner-group at University Heidelberg (BZH) for mass spectrometry and the Genomic Core Facility at EMBL-Heidelberg for support of the RT–qPCR.

FUNDING

The EU-grant '3D repertoire,' contract number LSHG-CT-2005-512028; the Deutsche Forschungsgemeinschaft (DFG, grant number BO1150); the Scottish Universities Life Sciences Alliance (SULSA) and a Wellcome Trust equipment grant (grant number WT087658MA) for the work carried out at the Electron Cryo Microscopy facility at the University of Edinburgh; the Darwin Trust of Edinburgh to B.B; Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081). Funding for open access charge: DFG.

Conflict of interest statement. None declared.

REFERENCES

1. Nakanishi,K. and Nureki,O. (2005) Recent progress of structural biology of tRNA processing and modification. *Mol. Cell.*, **19**, 157–166.
2. 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.
3. Walker,S.C. and Engelke,D.R. (2006) Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit. Rev. Biochem. Mol. Biol.*, **41**, 77–102.
4. Chang,D.D. and Clayton,D.A. (1987) A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. *EMBO J.*, **6**, 409–417.
5. Karwan,R., Bennett,J.L. and Clayton,D.A. (1991) Nuclear RNase MRP processes RNA at multiple discrete sites: interaction with an upstream G box is required for subsequent downstream cleavages. *Genes Dev.*, **5**, 1264–1276.
6. Schmitt,M.E. and Clayton,D.A. (1992) Yeast site-specific ribonucleoprotein endoribonuclease MRP contains an RNA component homologous to mammalian RNase MRP RNA and essential for cell viability. *Genes Dev.*, **6**, 1975–1985.
7. 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.
8. Salinas,K., Wierzbicki,S., Zhou,L. and Schmitt,M.E. (2005) Characterization and purification of Saccharomyces cerevisiae RNase MRP reveals a new unique protein component. *J. Biol. Chem.*, **280**, 11352–11360.
9. Schmitt,M.E. and Clayton,D.A. (1994) Characterization of a unique protein component of yeast RNase MRP: an RNA-binding protein with a zinc-cluster domain. *Genes Dev.*, **8**, 2617–2628.
10. Bertrand,E., Houser-Scott,F., Kendall,A., Singer,R.H. and Engelke,D.R. (1998) Nucleolar localization of early tRNA processing. *Genes Dev.*, **12**, 2463–2468.
11. Jacobson,M.R., Cao,L.G., Wang,Y.L. and Pederson,T. (1995) Dynamic localization of RNase MRP RNA in the nucleolus observed by fluorescent RNA cytochemistry in living cells. *J. Cell. Biol.*, **131**, 1649–1658.
12. Kiss,T. and Filipowicz,W. (1992) Evidence against a mitochondrial location of the 7-2/MRP RNA in mammalian cells. *Cell*, **70**, 11–16.
13. Lygerou,Z., Allmang,C., Tollervey,D. and Seraphin,B. (1996) Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP in vitro. *Science*, **272**, 268–270.
14. Schmitt,M.E. and Clayton,D.A. (1993) Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in Saccharomyces cerevisiae. *Mol. Cell. Biol.*, **13**, 7935–7941.
15. Lindahl,L., Bommankanti,A., Li,X., Hayden,L., Jones,A., Khan,M., Oni,T. and Zengel,J.M. (2009) RNase MRP is required for entry of 35S precursor rRNA into the canonical processing pathway. *RNA*, **15**, 1407–1416.
16. Gill,T., Cai,T., Aulds,J., Wierzbicki,S. and Schmitt,M.E. (2004) RNase MRP cleaves the CLB2 mRNA to promote cell cycle

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

progression: novel method of mRNA degradation. *Mol. Cell. Biol.*, **24**, 945–953.
17. Piccinelli,P., Rosenblad,M.A. and Samuelsson,T. (2005) Identification and analysis of ribonuclease P and MRP RNA in a broad range of eukaryotes. *Nucleic Acids Res.*, **33**, 4485–4495.
18. Chen,J.L. and Pace,N.R. (1997) Identification of the universally conserved core of ribonuclease P RNA. *RNA*, **3**, 557–560.
19. Lindahl,L., Fretz,S., Epps,N. and Zengel,J.M. (2000) Functional equivalence of hairpins in the RNA subunits of RNase MRP and RNase P in Saccharomyces cerevisiae. *RNA*, **6**, 653–658.
20. Li,X., Frank,D.N., Pace,N., Zengel,J.M. and Lindahl,L. (2002) Phylogenetic analysis of the structure of RNase MRP RNA in yeasts. *RNA*, **8**, 740–751.
21. Krasilnikov,A.S., Xiao,Y., Pan,T. and Mondragon,A. (2004) Basis for structural diversity in homologous RNAs. *Science*, **306**, 104–107.
22. Krasilnikov,A.S., Yang,X., Pan,T. and Mondragon,A. (2003) Crystal structure of the specificity domain of ribonuclease P. *Nature*, **421**, 760–764.
23. Kazantsev,A.V., Krivenko,A.A., Harrington,D.J., Holbrook,S.R., Adams,P.D. and Pace,N.R. (2005) Crystal structure of a bacterial ribonuclease P RNA. *Proc. Natl Acad. Sci. USA*, **102**, 13392–13397.
24. Reiter,N.J., Osterman,A., Torres-Larios,A., Swinger,K.K., Pan,T. and Mondragon,A. (2010) Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*, **468**, 784–789.
25. Torres-Larios,A., Swinger,K.K., Krasilnikov,A.S., Pan,T. and Mondragon,A. (2005) Crystal structure of the RNA component of bacterial ribonuclease P. *Nature*, **437**, 584–587.
26. Amero,C.D., Boomershine,W.P., Xu,Y. and Foster,M. (2008) Solution structure of Pyrococcus furiosus RPP21, a component of the archaeal RNase P holoenzyme, and interactions with its RPP29 protein partner. *Biochemistry*, **47**, 11704–11710.
27. Honda,T., Kakuta,Y., Kimura,K., Saho,J. and Kimura,M. (2008) Structure of an archaeal homolog of the human protein complex Rpp21-Rpp29 that is a key core component for the assembly of active ribonuclease P. *J. Mol. Biol.*, **384**, 652–662.
28. Kakuta,Y., Ishimatsu,I., Numata,T., Kimura,K., Yao,M., Tanaka,I. and Kimura,M. (2005) Crystal structure of a ribonuclease P protein Ph1601p from Pyrococcus horikoshii OT3: an archaeal homologue of human nuclear ribonuclease P protein Rpp21. *Biochemistry*, **44**, 12086–12093.
29. Kawano,S., Nakashima,T., Kakuta,Y., Tanaka,I. and Kimura,M. (2006) Crystal structure of protein Ph1481p in complex with protein Ph1877p of archaeal RNase P from Pyrococcus horikoshii OT3: implication of dimer formation of the holoenzyme. *J. Mol. Biol.*, **357**, 583–591.
30. Numata,T., Ishimatsu,I., Kakuta,Y., Tanaka,I. and Kimura,M. (2004) Crystal structure of archaeal ribonuclease P protein Ph1771p from Pyrococcus horikoshii OT3: an archaeal homolog of eukaryotic ribonuclease P protein Rpp29. *RNA*, **10**, 1423–1432.
31. Sidote,D.J. and Hoffman,D.W. (2003) NMR structure of an archaeal homologue of ribonuclease P protein Rpp29. *Biochemistry*, **42**, 13541–13550.
32. Takagi,H., Watanabe,M., Kakuta,Y., Kamachi,R., Numata,T., Tanaka,I. and Kimura,M. (2004) Crystal structure of the ribonuclease P protein Ph1877p from hyperthermophilic archaeon Pyrococcus horikoshii OT3. *Biochem. Biophys. Res. Commun.*, **319**, 787–794.
33. Wilson,R.C., Bohlen,C.J., Foster,M.P. and Bell,C.E. (2006) Structure of Pfu Pop5, an archaeal RNase P protein. *Proc. Natl Acad. Sci. USA*, **103**, 873–878.
34. Hamma,T. and Ferre-D'Amare,A.R. (2004) Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 A resolution. *Structure*, **12**, 893–903.
35. Perederina,A., Esakova,O., Quan,C., Khanova,E. and Krasilnikov,A.S. (2010) Eukaryotic ribonucleases P/MRP: the crystal structure of the P3 domain. *EMBO J.*, **29**, 761–769.
36. Stams,T., Niranjanakumari,S., Fierke,C.A. and Christianson,D.W. (1998) Ribonuclease P protein structure: evolutionary origins in the translational apparatus. *Science*, **280**, 752–755.
37. Hall,T.A. and Brown,J.W. (2004) Interactions between RNase P protein subunits in archaea. *Archaea*, **1**, 247–254.
38. Houser-Scott,F., Xiao,S., Millikin,C.E., Zengel,J.M., Lindahl,L. and Engelke,D.R. (2002) Interactions among the protein and RNA subunits of Saccharomyces cerevisiae nuclear RNase P. *Proc. Natl Acad. Sci. USA*, **99**, 2684–2689.
39. Jiang,T. and Altman,S. (2001) Protein-protein interactions with subunits of human nuclear RNase P. *Proc. Natl Acad. Sci. USA*, **98**, 920–925.
40. Kifusa,M., Fukuhara,H., Hayashi,T. and Kimura,M. (2005) Protein-protein interactions in the subunits of ribonuclease P in the hyperthermophilic archaeon Pyrococcus horikoshii OT3. *Biosci. Biotechnol. Biochem.*, **69**, 1209–1212.
41. 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.
42. Longtine,M.S., McKenzie,A. 3rd, Demarini,D.J., Shah,N.G., Wach,A., Brachat,A., Philippsen,P. and Pringle,J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast*, **14**, 953–961.
43. Janke,C., Magiera,M.M., Rathfelder,N., Taxis,C., Reber,S., Maekawa,H., Moreno-Borchart,A., Doenges,G., Schwob,E., Schiebel,E. *et al.* (2004) A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast*, **21**, 947–962.
44. Puig,O., Caspary,F., Rigaut,G., Rutz,B., Bouveret,E., Bragado-Nilsson,E., Wilm,M. and Seraphin,B. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods*, **24**, 218–229.
45. Bassler,J., Grandi,P., Gadal,O., Lessmann,T., Petfalski,E., Tollervey,D., Lechner,J. and Hurt,E. (2001) Identification of a 60S preribosomal particle that is closely linked to nuclear export. *Mol. Cell.*, **8**, 517–529.
46. Sambrook,J. and Russell,D.W. (2001) *Molecular cloning: a laboratory manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
47. Zuker,M., Mathews,D.H. and Turner,D.H. (1999) Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In: Barciszewski,J. and Clark,B.F.C. (eds), *RNA biochemistry and biotechnology*. Kluwer Academic Publishers, Dordrecht, pp. 11–43.
48. Nolan,T., Hands,R.E. and Bustin,S.A. (2006) Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.*, **1**, 1559–1582.
49. Yang,L. and Altman,S. (2007) A noncoding RNA in Saccharomyces cerevisiae is an RNase P substrate. *RNA*, **13**, 682–690.
50. Ziehler,W.A., Day,J.J., Fierke,C.A. and Engelke,D.R. (2000) Effects of 5′ leader and 3′ trailer structures on pre-tRNA processing by nuclear RNase P. *Biochemistry*, **39**, 9909–9916.
51. Golas,M.M., Sander,B., Will,C.L., Luhrmann,R. and Stark,H. (2003) Molecular architecture of the multiprotein splicing factor SF3b. *Science*, **300**, 980–984.
52. Frank,J., Radermacher,M., Penczek,P., Zhu,J., Li,Y., Ladjadj,M. and Leith,A. (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.*, **116**, 190–199.
53. van Heel,M., Harauz,G., Orlova,E.V., Schmidt,R. and Schatz,M. (1996) A new generation of the IMAGIC image processing system. *J. Struct. Biol.*, **116**, 17–24.
54. Scheres,S.H., Gao,H., Valle,M., Herman,G.T., Eggermont,P.P., Frank,J. and Carazo,J.M. (2007) Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat. Methods*, **4**, 27–29.
55. Harauz,G. and Van Heel,M. (1986) Exact filters for general geometry 3-dimensional reconstruction. *Optik*, **73**, 146–156.
56. Bottcher,B., Wynne,S.A. and Crowther,R.A. (1997) Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature*, **386**, 88–91.
57. Rosenthal,P.B. and Henderson,R. (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.*, **333**, 721–745.
58. Ludtke,S.J., Baldwin,P.R. and Chiu,W. (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.*, **128**, 82–97.

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

59. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
60. Marquez,S.M., Chen,J.L., Evans,D. and Pace,N.R. (2006) Structure and function of eukaryotic Ribonuclease P RNA. *Mol. Cell.*, **24**, 445–456.
61. Fernandez-Tornero,C., Bottcher,B., Riva,M., Carles,C., Steuerwald,U., Ruigrok,R.W., Sentenac,A., Muller,C.W. and Schoehn,G. (2007) Insights into transcription initiation and termination from the electron microscopy structure of yeast RNA polymerase III. *Mol. Cell.*, **25**, 813–823.
62. Diepholz,M., Venzke,D., Prinz,S., Batisse,C., Florchinger,B., Rossle,M., Svergun,D.I., Bottcher,B. and Fethiere,J. (2008) A different conformation for EGC stator subcomplex in solution and in the assembled yeast V-ATPase: possible implications for regulatory disassembly. *Structure*, **16**, 1789–1798.
63. Ulbrich,C., Diepholz,M., Bassler,J., Kressler,D., Pertschy,B., Galani,K., Bottcher,B. and Hurt,E. (2009) Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. *Cell*, **138**, 911–922.
64. Xu,Y., Amero,C.D., Pulukkunat,D.K., Gopalan,V. and Foster,M.P. (2009) Solution structure of an archaeal RNase P binary protein complex: formation of the 30-kDa complex between Pyrococcus furiosus RPP21 and RPP29 is accompanied by coupled protein folding and highlights critical features for protein-protein and protein-RNA interactions. *J. Mol. Biol.*, **393**, 1043–1055.
65. Perederina,A., Esakova,O., Koc,H., Schmitt,M.E. and Krasilnikov,A.S. (2007) Specific binding of a Pop6/Pop7 heterodimer to the P3 stem of the yeast RNase MRP and RNase P RNAs. *RNA*, **13**, 1648–1655.
66. Perederina,A. and Krasilnikov,A.S. (2010) The P3 domain of eukaryotic RNases P/MRP: making a protein-rich RNA-based enzyme. *RNA Biol.*, **7**, 534–539.
67. Honda,T., Hara,T., Nan,J., Zhang,X. and Kimura,M. (2010) Archaeal homologs of human RNase P protein pairs Pop5 with Rpp30 and Rpp21 with Rpp29 work on distinct functional domains of the RNA subunit. *Biosci. Biotechnol. Biochem.*, **74**, 266–273.
68. Aspinall,T.V., Gordon,J.M., Bennett,H.J., Karahalios,P., Bukowski,J.P., Walker,S.C., Engelke,D.R. and Avis,J.M. (2007) Interactions between subunits of Saccharomyces cerevisiae RNase MRP support a conserved eukaryotic RNase P/MRP architecture. *Nucleic Acids Res.*, **35**, 6439–6450.
69. Dlakic,M. (2005) 3D models of yeast RNase P/MRP proteins Rpp1p and Pop3p. *RNA*, **11**, 123–127.
70. Cho,I.M., Lai,L.B., Susanti,D., Mukhopadhyay,B. and Gopalan,V. (2010) Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc. Natl Acad. Sci. USA*, **107**, 14573–14578.
71. Fukuhara,H., Kifusa,M., Watanabe,M., Terada,A., Honda,T., Numata,T., Kakuta,Y. and Kimura,M. (2006) A fifth protein subunit Ph1496p elevates the optimum temperature for the ribonuclease P activity from Pyrococcus horikoshii OT3. *Biochem. Biophys. Res. Commun.*, **343**, 956–964.
72. Rosenblad,M.A., Lopez,M.D., Piccinelli,P. and Samuelsson,T. (2006) Inventory and analysis of the protein subunits of the ribonucleases P and MRP provides further evidence of homology between the yeast and human enzymes. *Nucleic Acids Res.*, **34**, 5145–5156.
73. Pluk,H., van Eenennaam,H., Rutjes,S.A., Pruijn,G.J. and van Venrooij,W.J. (1999) RNA-protein interactions in the human RNase MRP ribonucleoprotein complex. *RNA*, **5**, 512–524.
74. Welting,T.J., Kikkert,B.J., van Venrooij,W.J. and Pruijn,G.J. (2006) Differential association of protein subunits with the human RNase MRP and RNase P complexes. *RNA*, **12**, 1373–1382.
75. Barrera,A., Fang,X., Jacob,J., Casey,E., Thiyagarajan,P. and Pan,T. (2002) Dimeric and monomeric Bacillus subtilis RNase P holoenzyme in the absence and presence of pre-tRNA substrates. *Biochemistry*, **41**, 12986–12994.
76. Buck,A.H., Dalby,A.B., Poole,A.W., Kazantsev,A.V. and Pace,N.R. (2005) Protein activation of a ribozyme: the role of bacterial RNase P protein. *EMBO J.*, **24**, 3360–3368.
77. Radermacher,M., Wagenknecht,T., Verschoor,A. and Frank,J. (1987) 3-dimensional reconstruction from a single exposure, random conical tilt series applied to the 50S-risomal subunit of escherichia-coli. *J. Microscopy*, **146**, 113–136.
78. Mindell,J.A. and Grigorieff,N. (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.*, **142**, 334–347.
79. van Heel,M., Harauz,G. and Orlova,E.V. (1996) A new generation of the IMAGIC image processing system. *J. Struct. Biol.*, **116**, 17–24.
80. van Heel,M. and Frank,J. (1981) Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy*, **6**, 187–194.

Downloaded from http://nar.oxfordjournals.org/ by guest on November 1, 2012

Supplementary Information

Random Conical Tilt Reconstruction

To determine starting models for RNase P and RNase MRP, we collected pairs of tilted and untilted negatively stained particles (-55°, 0°, Philips CM200, 200 kV, 27,500 magnification, 2kx2k TVIPS CCD camera) for random conical tilt reconstruction (1). We selected 2,156 pairs of particle images for RNase P and 1,588 pairs for RNase MRP. The untilted particle images were aligned and classified (100 classes for RNase P and 80 classes for RNase MRP) and the tilted images were centered. For each class of the untilted images, a volume was calculated from the related tilted images using the tilt angle from the microscope and the in-plane rotation determined for the corresponding untilted particle image. Many classes gave consistent volumes, but showed flattening due to the staining and drying process. Matching volumes were combined and used as a reference in further projection matching of tilted and untilted particle images. More untilted particle images were added to the data giving a total (tilted + untilted) of 9,226 particle images for RNase P and 7,553 particle images for RNase MRP. The results are summarized in Figure S1. The final volumes do not show any indications of flattening and represent the initial class averages of the untilted particles well. The random conical tilt reconstructions established the absolute hand off the models and were reliable starting models for the refinement of the larger data set of cryo-negative stain images.

CTF-correction and normalization of particle images

For each micrograph, the ctf-parameters were determined with ctffind2 (2). Particle images were boxed with a box size of 160x160 pixels, 35.2x35.2 nm. For phase reversal the boxed particle images were padded to a box size of 600x600 pixels and the phases were reversed with Spider (3) using the ctf-parameters determined with ctffind2. After phase reversal, the grey values were normalized, the box size was reduced to 120x120 pixels and a smooth circular mask was applied.

Determination of orientations of cryo-negative stain images by projection matching

The orientations of particle images were determined by projection matching. During iterative refinement projection matching was carried out in 2 steps: First the global orientation of the particle images was determined by a coarse search. Therefore, the reference volume was low-pass filtered to exclude spatial frequencies > 1/ 3.5 nm^{-1} and projected in 10° steps using quasi-evenly spaced projection directions as determined by the VO EA option of Spider. Particle images were cross-correlated with each of the reference projections. Then the orientation of the reference projection with which the particle image had the highest cross-correlation was assigned as particle orientation. For the global search the AP SH option of Spider was used, which performs an exhaustive grid search for the in-plane rotation and the origin. Changes in the origin of the particles were limited to 9 pixels. As first starting

model the refined volumes of the random conical tilt reconstruction were used (see above).

The coarse search was followed by a finer search in 2° steps in a limited search range (+/- 20° around the previously determined orientation). In the finer search the reference volume was low-pass filtered to include all spatial frequencies up to the frequency, where the Fourier-shell correlation in the previous iteration dropped to 0.5. The filtered reference volume was projected in 2° steps. Particle images were correlated with each of the projections within the limited search range. For the limited search, the AP REF option of Spider was used. In the limited search the change in origin was restricted to 4 pixels (0.88 nm).

Calculation of image reconstructions

The data was split into 20 defocus groups. For each defocus group three three-dimensional maps were calculated using the BP 32F option of Spider, which does a back-projection interpolated in Fourier space. The maps were calculated from particle images showing correlations with the reference projections above a certain threshold value. Because the correlation strongly depends on the defocus, we used a dynamic cut-off, which was determined by a linear fit of the correlation values of all particle images in dependence of the defocus value. The threshold value was chosen to include approximately 80% of all particle images. Some orientations were more frequent than others giving rise to anisotropic resolution. To reduce this effect, we restricted the number of particles with similar orientations in the final reconstructions. A maximum of 180 (RNase P) or 220 (RNase MRP) particle images respectively, which correlated best to the same reference in the global search, was included.

The maps, which were calculated with BP 32F, represented either all particle images above the threshold criterion (full-map) or alternating halves of the set of particle images (A- and B-maps). The maps of the different defocus groups were weighted according to the relative particle numbers and added to the final maps (full-map, A-map and B-map). The Fourier-shell correlation (4) was determined between A- and B-maps.

For preparing the full-map as reference map for projection matching, the map was masked with a binary mask with smooth edges, which was slightly larger than the volume of the particle. This reduced the background noise in the reference projections.

Checking the consistency of the assignment of particle orientations in the global search

To test whether the determination of particle orientations during the global orientation search was consistent, we sub-classified particle images, which were assigned with the same Euler angles. Sub-classification of the aligned particles was done using the multivariate statistical analysis of IMAGIC (5). In general sub-classes were consistent with the reference projections (Figure S2). This suggested that the large majority of particle images was assigned with the correct Euler angles.

Supplementary Figures and Legends

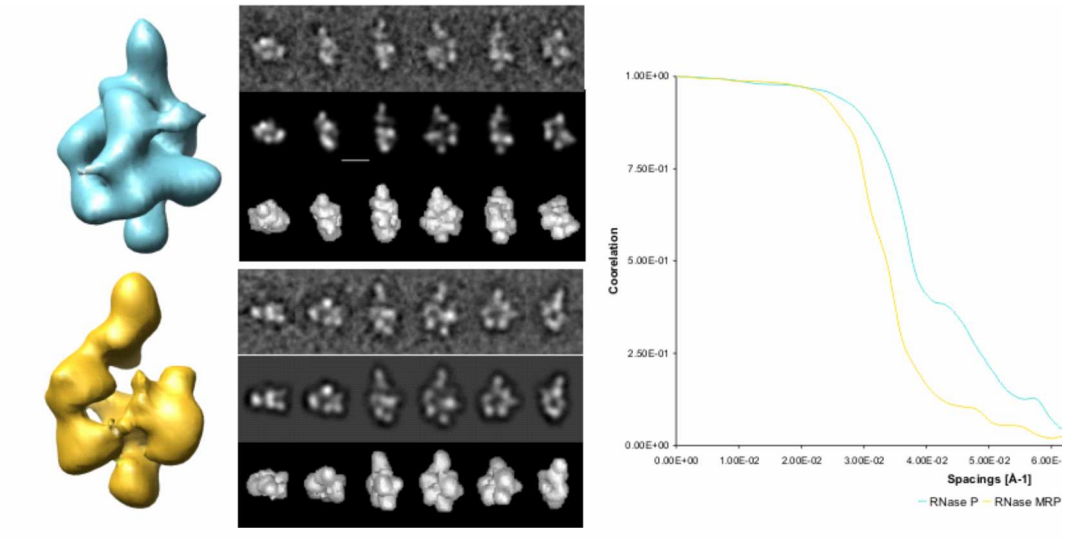
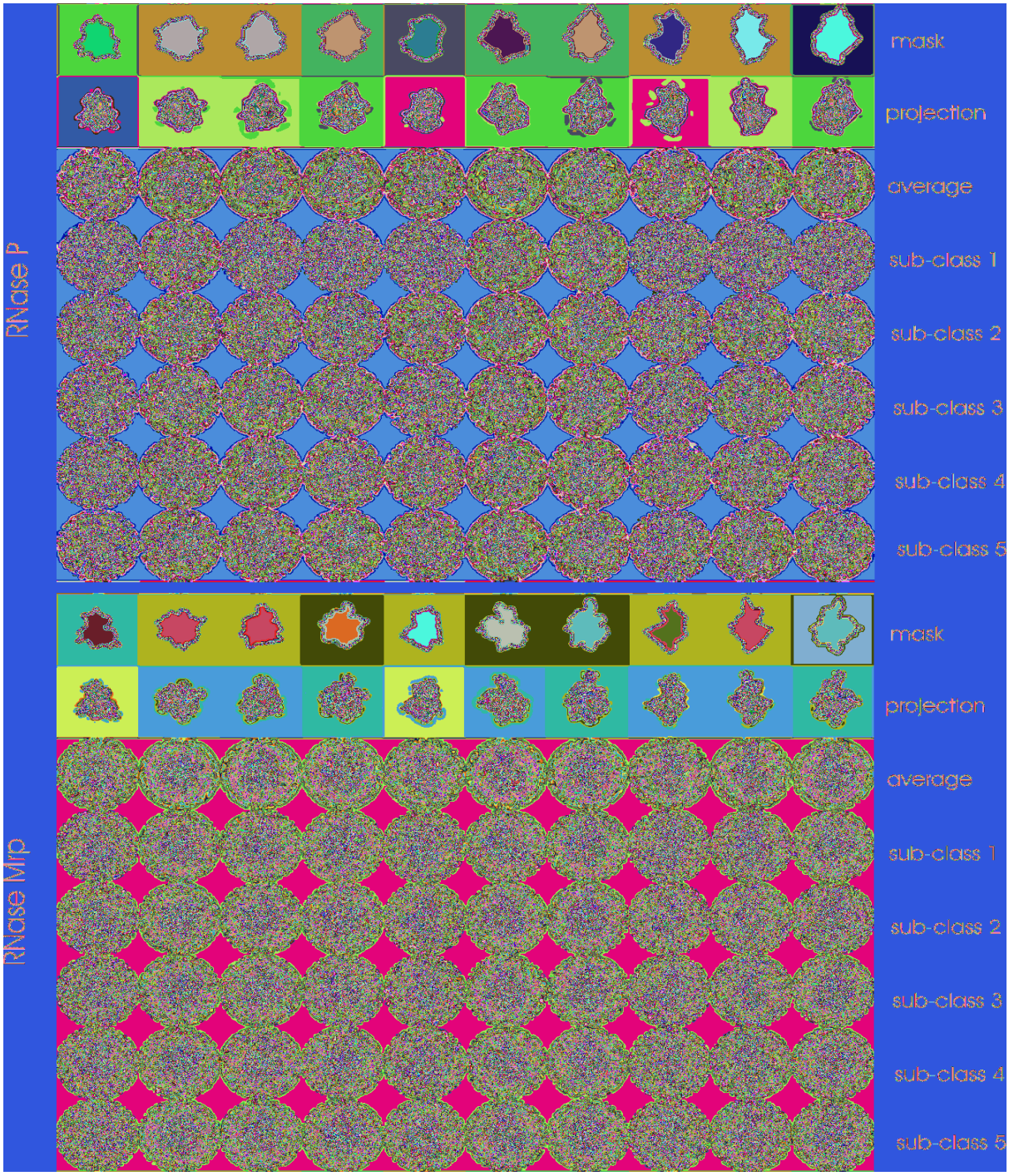


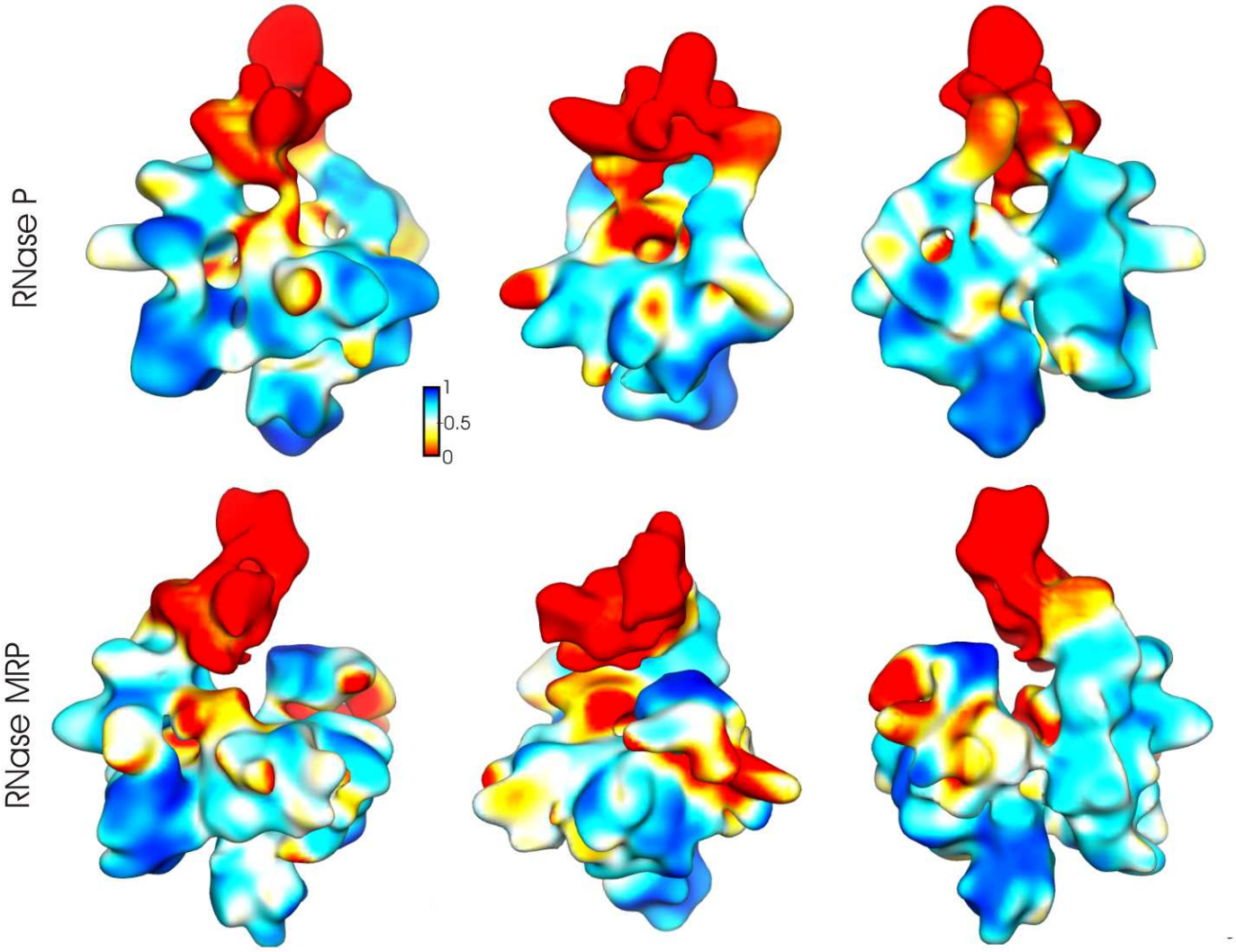
Figure S1. Initial reconstructions of RNase P (top) and RNase MRP (bottom). Left: surface representations are shown in a similar orientation. Centre: class averages from projection matching (top), projections of the volume in the same direction as determined for class averages (middle), surface representations of the volume viewed in the same direction as the projections (bottom). Right: Fourier shell correlation. The correlation drops to 0.5 at 1/27 $Å^{-1}$ for RNase P (blue) and 1/31 $Å^{-1}$ for RNase MRP (yellow).



S2 Sub-classification of particles, which aligned to the same reference in the global orientation search. The upper panel shows RNase P, the lower panel shows RNase MRP. The analysis is shown for every 10th reference projection in the global search (columns 1-10). Row 1 shows the binary mask that was used in multivariate statistical analysis (6) in the sub-classification process. Row 2 shows the projection of the map, which were used as reference projections in the global search. Row 3 shows the average of all particles, which aligned best to the same reference projection in the global search. Reference projections (row 2) and averages (row 3) were similar highlighting the overall validity of the map.

Rows 4-8 show class averages of sub-classes of particles that aligned to the same reference in the global search. The classes were derived by multivariate statistical analysis. The majority of class averages are similar to the reference projection. This indicates that there are no systematic errors in the assignments of orientations and that there is very little heterogeneity in the data. This is consistent with observation from maximum-likelihood sorting using XMIPP (7). However, some of the less densely populated orientations contain images that do not represent particles and were accidentally included in the data set.

Figure S3 Surface representation of the map of RNase P (top panel) and RNase MRP (bottom panel) colored with the local correlation between the fitted atomic models and the EM-maps as shown in Figure 4 (the comparison also included Pop3 at the position indicated in Figure 4 by an arrow). The color-key for the correlation is given in the upper panel. White-blue colors represent cross-correlations >0.5 and white-yellow-red-colors cross-correlations < 0.5. The local cross-correlations were calculated within a distance of 1.1 nm (5 pixels) from the point of interest. For cross-correlation, the fitted models were converted to density maps and low-pass filtered to spatial frequencies of 1/1.7 nm^{-1} (RNase P) and 1/1.5 nm^{-1} (RNase MRP). The models do not represent density for Pop1, which explains the correlation



of 0 at the top of the particle.

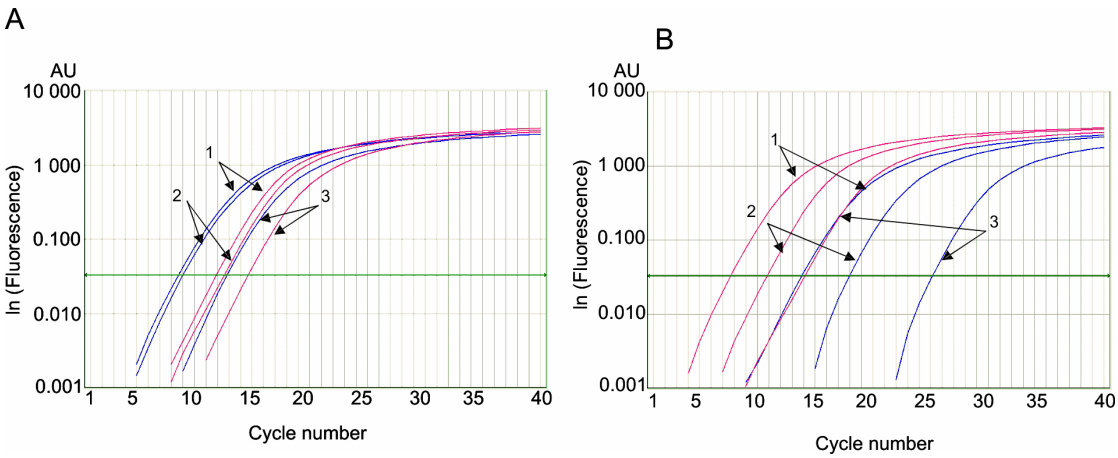


Figure S4. Logarithm of the fluorescence intensity over the number of cycles during qPCR of samples from three different purifications of RNase P and RNase MRP. (A) HA eluates of Pop4-TAP Rpr2-HA (RNase P) and (B) Pop4-TAP HA-Snm1 (RNase MRP). Red curves: Nme1 RNA (RNase MRP); blue curves: Rpr1 RNA (RNase P); green line: Ct threshold. The Δct ratios of Nme1 to Rpr1 were determined for three independent purifications (indicated by the arrows and numbers 1, 2 and 3). Amplification efficiencies were similar for both RNAs (ca. 92.5 %). For all purifications, there was at least 10 fold more Rpr1 than Nme1 for RNase P purifications and more than 50 fold Nme1 than Rpr1 for RNase MRP purifications.

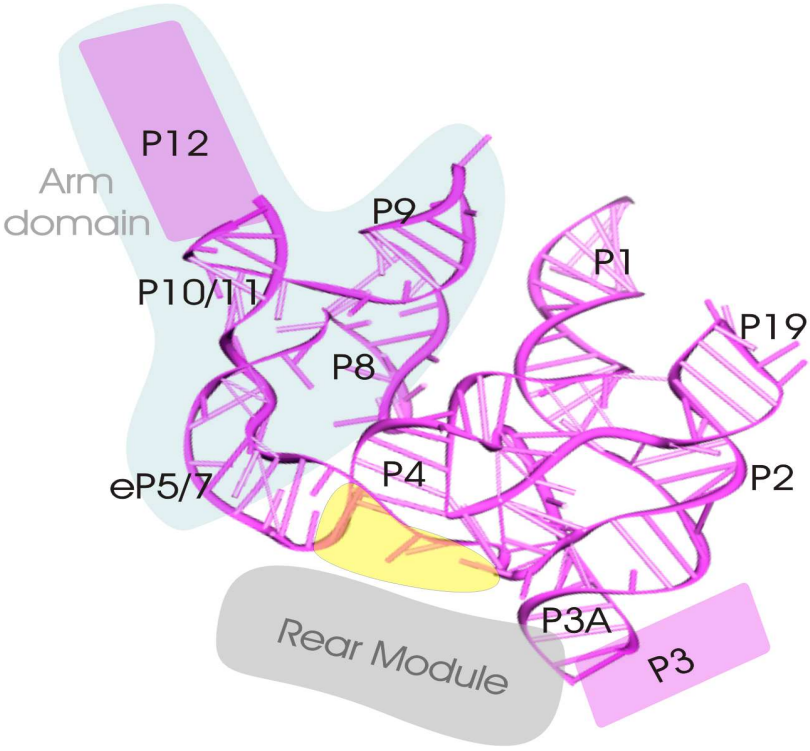


Figure S5 Cartoon of the model of the fitted *S. pombe* RNA (8) relative to the rear module. The RNA is viewed from the lid module towards the foot domain. The rear module and the approximate position of the arm domain are shown in grey. The approximate position of the catalytic site is shown in yellow. Larger stretches of RNA that are missing in the model are shown as boxes colored in magenta. The stems or the RNA model are labeled in black according to (8). The S-domain of the RNA-subunit is indicated by a light blue background.

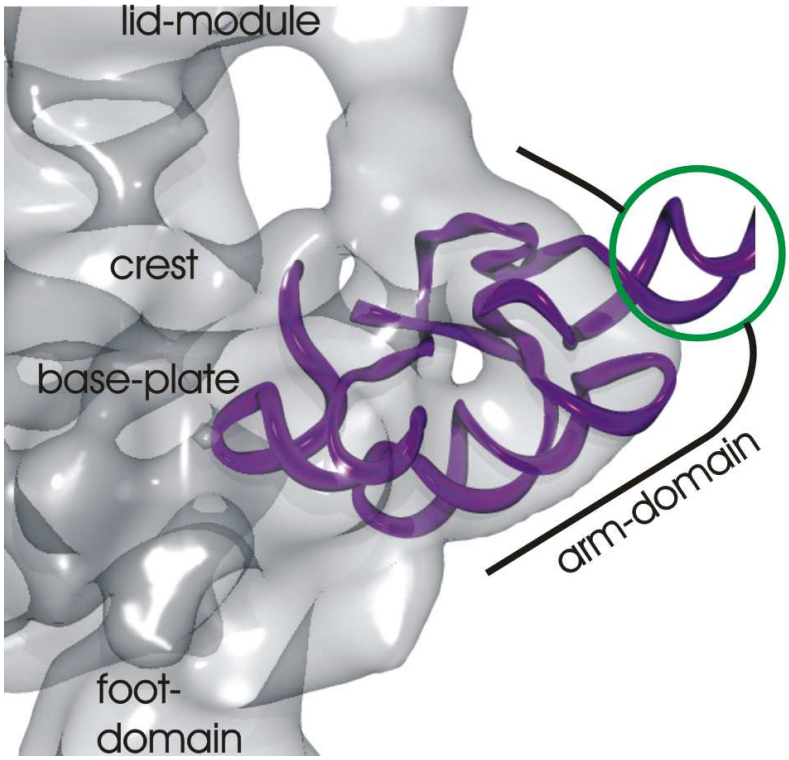


Figure S6 S-domain of the RNA-subunit of *Thermotoga maritima* ((9), purple) superimposed to the EM-map of the RNase P of *S. cerevisiae*. The subunit was placed manually. The green circle highlights an engineered region, which was introduced for optimizing crystal contacts and is not accommodated by the EM-density.

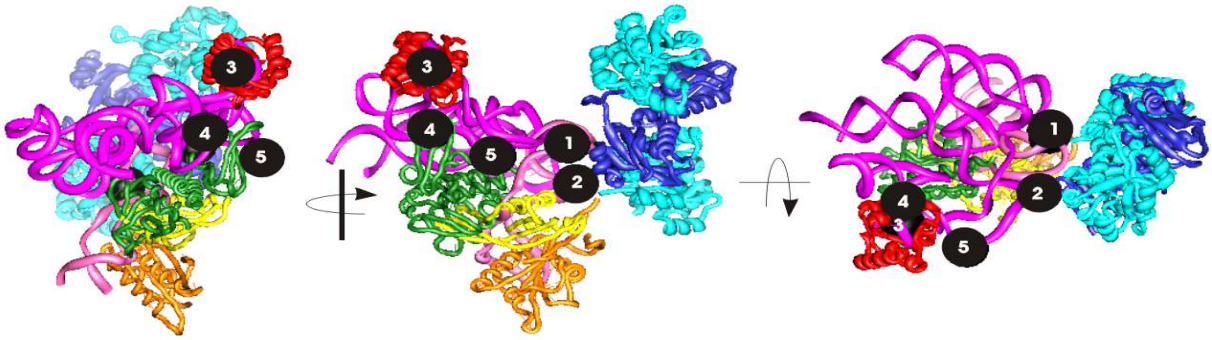


Figure S7 Model of RNase P as shown in Figure 4. The views are rotated by 90° around the axes as indicated. The black spheres have a diameter of 15Å and indicate positions of reported interactions that are consistent with our model: 1) interaction of P3a with the P3 stem bound to Pop6 and Pop7 (10) and interaction of the P3 stem with the Pop5-Rpp1 sub-complex (11). 2) Interaction of the P5 stem with Pop5-Rpp1 (11); 3) Interaction of the Pop3 homologue L7Ae with the P12-stem in Archaea (12). 4) Interaction of Pop4-Rpr2 with stems P9, P10 and P11 (11). 5) Snm1 in RNase MRP, which is homologous to Rpr2, is adjacent to the RNase MRP specific subunit Rmp1 (13). Position 5 highlights a region where the map of RNase MRP shows extra density, which is absent in the map of RNase P. For further details of the see also table S1.

Table S1 Summary of the evidence that supports the assignment and fitting of the subunits in the map and an estimation of the accuracy of the fits of the individual subunits and sub-complexes

subunit /sub-complex	**Evidence for the positioning**	**Estimated accuracy of the orientation of the fitted subunit/ sub-complex**	**comment**
Pop6-Pop7-P3	shape similarity of crystal structure (10) and map (this study) specific binding of Pop6 and Pop7 to the the P3 stem (14) (Figure S7, position 1) labeling (this study)	unique fit to the foot domain	
Pop5-Rpp1	shape similarity between the crystal structure of the homologous $(Ph1871\text{-}Ph1481)_2$ (15) with the rear module (the rear module is the only pseudo 2-fold symmetrical entity of the EM-density map) consistent with labeling (this study) archaeal Pop5-Rpp1 homologues bind to the C-domain of the respective RNA-subunit (16) (Figure S7, positions 1 and 2) foot-printing of archaeal homologues to the stems P3 and P5 (Figure S7) of the respective RNA-subunit (11) (Figure S7, positions 1 and 2)	unique fit to the rear module	one Pop5 might be exchanged for a Pop8 as suggested by the evolutionary relation between Pop5 and Pop8 (17) and as also consistent with labeling (this study)

RNA subunit	orientation relative to the P3 stem in the foot domain (Figure S7, interaction 1) foot-printing of archaeal homologues of Pop5-Rpp1 to stems P3 and P5 (11) support the relative orientation of the rear module to the RNA-subunit (Figure S7, positions 1 and 2)	some rotational freedom (<50°,mainly out of plane tilt of the RNA subunit) small positional freedom (<10Å)	
Pop3	In archaea Pop3 is replaced by L7Ae (12) which binds kink-turn motifs. These kink-turn motifs are preserved in the P12 stem of the RNA-subunit (12) (Figure S7, position 3).	full rotational freedom and small shifts (<10 Å) inside the arm module; therefore not included in the fit shown in Figure 4	
Rpr2-Pop4	consistent with labeling (this study) elongated shape of the crystal structure of the homologous Rpp21-Rpp29 (18) matches the unaccounted density underneath the base plate in the EM-map the archaeal Rpr2-Pop4 homologues bind to the S-domain of the respective RNA-subunit (16) (Figure S7, position 4) foot-printing of archaeal homologues to stems P9, P10, P11 and P12 of the respective RNA-subunit (11), (Figure S7, position 4) in RNase MRP the Rpr2 homologue Snm1	some rotational freedom mainly around the long axis of the Pop4-Rpr2-heterodimer (<40°) and small positional freedom (<10 A)	

	interacts with Rmp1 (13). This is consistent with the vicinity of the fitted Rpr2 to the unaccounted extra density in the map of RNase MRP (this study, Fig 2B *, Figure S7 position 5)		
Pop1	labeling (this study) the 100 $Å^3$ of the lid module agrees approximately with a molecular weight of a 100 kDa protein		

Supplementary References

1. Radermacher, M., Wagenknecht, T., Verschoor, A. and Frank, J. (1987) 3-dimensional reconstruction from a single exposure, random conical tilt series applied to the 50S-risomal subunit of escherichia-coli. *J of Microscopy-Oxford*, **146,** 113-136.
2. Mindell, J.A. and Grigorieff, N. (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol*, **142,** 334-347.
3. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. and Leith, A. (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol*, **116,** 190-199.
4. Harauz, G. and Van Heel, M. (1986) Exact Filters for General Geometry 3-Dimensional Reconstruction. *Optik*, **73,** 146-156.
5. van Heel, M., Harauz, G. and Orlova, E.V. (1996) A new generation of the IMAGIC image processing system. *J Struct Biol*, **116,** 17-24.
6. van Heel, M. and Frank, J. (1981) Use of multivariate statistics in analysing the images of biological macromolecules. *Ultramicroscopy*, **6,** 187-194.
7. Scheres, S.H., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P., Frank, J. and Carazo, J.M. (2007) Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Methods*, **4,** 27-29.
8. Marquez, S.M., Chen, J.L., Evans, D. and Pace, N.R. (2006) Structure and function of eukaryotic Ribonuclease P RNA. *Mol Cell*, **24,** 445-456.
9. Reiter, N.J., Osterman, A., Torres-Larios, A., Swinger, K.K., Pan, T. and Mondragon, A. (2010) Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*, **468,** 784-789.
10. Perederina, A., Esakova, O., Quan, C., Khanova, E. and Krasilnikov, A.S. (2010) Eukaryotic ribonucleases P/MRP: the crystal structure of the P3 domain. *EMBO J*, **29,** 761-769.
11. Xu, Y., Amero, C.D., Pulukkunat, D.K., Gopalan, V. and Foster, M.P. (2009) Solution structure of an archaeal RNase P binary protein complex: formation of the 30-kDa complex between Pyrococcus furiosus RPP21 and RPP29 is accompanied by coupled protein folding and highlights critical features for protein-protein and protein-RNA interactions. *J Mol Biol*, **393,** 1043-1055.
12. Cho, I.M., Lai, L.B., Susanti, D., Mukhopadhyay, B. and Gopalan, V. (2010) Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc Natl Acad Sci U S A*, **107,** 14573-14578.
13. Aspinall, T.V., Gordon, J.M., Bennett, H.J., Karahalios, P., Bukowski, J.P., Walker, S.C., Engelke, D.R. and Avis, J.M. (2007) Interactions between subunits of Saccharomyces cerevisiae RNase MRP support a conserved eukaryotic RNase P/MRP architecture. *Nucleic Acids Res*, **35,** 6439-6450.
14. Perederina, A., Esakova, O., Koc, H., Schmitt, M.E. and Krasilnikov, A.S. (2007) Specific binding of a Pop6/Pop7 heterodimer to the P3

stem of the yeast RNase MRP and RNase P RNAs. *RNA*, **13,** 1648-1655.
15. Kawano, S., Nakashima, T., Kakuta, Y., Tanaka, I. and Kimura, M. (2006) Crystal structure of protein Ph1481p in complex with protein Ph1877p of archaeal RNase P from Pyrococcus horikoshii OT3: implication of dimer formation of the holoenzyme. *J Mol Biol*, **357,** 583-591.
16. Honda, T., Hara, T., Nan, J., Zhang, X. and Kimura, M. (2010) Archaeal homologs of human RNase P protein pairs Pop5 with Rpp30 and Rpp21 with Rpp29 work on distinct functional domains of the RNA subunit. *Biosci Biotechnol Biochem*, **74,** 266-273.
17. Rosenblad, M.A., Lopez, M.D., Piccinelli, P. and Samuelsson, T. (2006) Inventory and analysis of the protein subunits of the ribonucleases P and MRP provides further evidence of homology between the yeast and human enzymes. *Nucleic Acids Res*, **34,** 5145-5156.
18. Honda, T., Kakuta, Y., Kimura, K., Saho, J. and Kimura, M. (2008) Structure of an archaeal homolog of the human protein complex Rpp21-Rpp29 that is a key core component for the assembly of active ribonuclease P. *J Mol Biol*, **384,** 652-662.

DISCUSSION

1. Purification de complexes RNases P et MRP intacts et fonctionnels

La réalisation de la modélisation des complexes RNases P et MRP a nécessité la purification de ces complexes en quantité suffisante et d'une qualité élevée. La mise au point de la stratégie de purification séquentielle a permis de purifier des complexes intacts contenant un ARN et 9 protéines pour la RNase P et 10 protéines pour la RNase MRP comme décrit précédemment (Lee *et al.*, 1991b; Schmitt *et al.*, 1992; Lygerou *et al.*, 1994; Schmitt *et al.*, 1994; Dichtl *et al.*, 1997; Stolc *et al.*, 1997; Chamberlain *et al.*, 1998; Salinas *et al.*, 2005). Par ailleurs, le complexe RNase P purifié est capable de fixer un ARNt prémature et de procéder à sa coupure. Pour compléter cette étude, il serait nécessaire de tester l'activité catalytique du complexe RNase MRP sur un de ces deux principaux substrats, des ARN avec un ITS1, espace interne transcrits 1 contenant un site A3, (Lygerou *et al.*, 1996a) ou les ARNm CLB2 (Gill *et al.*, 2004). Enfin, l'estimation de la quantité relative des deux composantes ARN dans les deux purifications nous conduit à conclure à une pureté élevée du complexe purifié dans les échantillons, de nature comparable à celle obtenue par Salinas *et al.*, 2005 lors de l'identification de la protéine Rmp1 dans la RNase MRP.

2. Rôle des protéines dans les complexes RNases P/MRP

Notre étude permet de mettre en évidence l'organisation des sous-unités protéiques dans les complexes RNases P et MRP. L'orientation des protéines permet de conclure que certaines protéines semblent stabiliser la structure de l'ARN et d'autres impliquées dans la liaison du substrat.

2.1. Protéines soutenant la structure de la composante ARN chez les eucaryotes

Au cours de l'évolution des bactéries vers les archées jusqu'aux eucaryotes, la sous-unité ARN s'est rétrécie tandis que la quantité de protéines additionnelles s'est accrue. Il a donc été proposé que la principale fonction des protéines additionnelles est de stabiliser le repliement de la composante ARN. Notre modèle suggère que c'est le cas pour les protéines Pop4 et Rpr2. Acquises chez les archées, il y a été mis en évidence une protection du domaine S contre l'action des nucléases en leur présence (Xu *et al.*, 2009). Dans notre modèle, elles sont situées près du bras du module de base en contact direct avec la composante ARN (**Figure 23.A**, panel 2). Ce positionnement est cohérent avec un rôle chaperon des protéines Rpr2 et Pop4 sur le composante ARN. Une étude (Khanova *et al.*, 2012) a par ailleurs récemment validé l'interaction de Pop4 avec l'hélice P9 confirmant notre localisation.

La protéine Pop3 est une autre protéine qui se lierait potentiellement au bras du module de base. Homologue de la protéine ribosomale archée L7Ae (Dlakic, 2005), elle lie les motifs K-turn, notamment celui prédit dans l'hélice P12 de la composante ARN de la RNase P (Rosenblad *et al.*, 2006). Cette interaction place Pop3 dans la région en forme de bras du module de base (**Figure 23.B**, flèche) qui serait assez large pour contenir l'hélice P12 et Pop3. Cependant, la résolution de nos cartes ne nous a pas permis de déterminer l'orientation exacte de Pop3 dans le volume. A cet emplacement, Pop3 serait placée stratégiquement pour médier l'identification et la liaison du substrat dans une position éloignée du site catalytique (Brusca *et al.*, 2001) et pour stabiliser la composante ARN dans une conformation catalytiquement active (Fukuhara *et al.*, 2006; Kosaka *et al.*, 2010).

Les protéines Pop6 et Pop7, spécifiques des eucaryotes, forment un hétérodimère qui lie la structure hélice-boucle-hélice P3 de la composante ARN des RNases P et MRP (Pluk *et al.*, 1999; Perederina *et al.*, 2010b; Khanova *et al.*, 2012). Cette structure semble avoir co-évoluée avec Pop6 et Pop7 chez les eucaryotes. Dans notre modèle, ce sous-complexe se place dans le pied du module de base (**Figure 23.A**, panel 3). L'hétérodimère Pop6-Pop7 entoure l'hélice-boucle-hélice P3 qui pointe vers le module arrière et se positionne directement sous la plaque basale. Cette localisation est similaire à celle de l'hélice P15.2 du noyau bactérien des RNases P de type B (Kazantsev *et al.*, 2005) et de l'hélice P18 du noyau bactérien des RNases de type A (Torres-Larios *et al.*, 2005) qui fait aussi une boucle sous le noyau catalytique ARN. Ainsi, il semble que l'hélice bactérienne P15.2/P18 soit remplacée par la structure hélice-boucle-hélice P3 en association avec l'hétérodimère Pop6-Pop7 chez les eucaryotes. Ce positionnement permet la formation d'un conduit entre le domaine au pied du module de base et le module arrière ce qui fournit un environnement protecteur en limitant l'espace où le substrat est coupé.

2.2. Protéines contribuant à la liaison du substrat

D'autres sous-unités des RNases P et MRP sont moins étroitement liés avec la composante ARN et sont par conséquent plus susceptibles de participer à d'autres fonctions que la stabilisation de la sous-unité ARN.

Les protéines Pop5 et Rpp1, acquises chez les archées, sont des protéines qui formeraient un hétérotétramère (Kawano *et al.*, 2006). Selon notre modèle, le module arrière serait composé uniquement de cet hétérotétramère Pop5-Rpp1 (**Figure 23.A**, panel 1). Notre positionnement du sous-complexe d'archée place les deux protéines Pop5 face à la sous-unité ARN ce qui est cohérent avec l'interaction entre Pop5 et la région conservée CR-IV du domaine catalytique de l'ARN (Perederina *et al.*, 2011). Le complexe Pop5-Rpp1 en se liant

au domaine catalytique du noyau ARN forme des cavités et des conduits entre le module arrière et le module de base, notamment au niveau de la plaque basale et du pied du module. Comme dans le pied du module de base, ces cavités semblent être positionnées de manière stratégique pour contribuer à la coordination de la séquence 5' leader au cours du processus de maturation.

D'autre part, le sous-complexe Pop5-Rpp1 des eucaryotes et des archées se lie directement sur un domaine conservé du domaine catalytique de la composante ARN (Honda *et al.*, 2010; Perederina *et al.*, 2011; Khanova *et al.*, 2012). Ce site de liaison correspond à celui de la protéine C5 de la RNase P bactérienne (Sharin *et al.*, 2005). De plus, Pop5, aPop5 et C5 partagent une structure avec un motif de liaison à l'ARN, le motif RNR (Jovanovic *et al.*, 2002; Perederina *et al.*, 2011). Cela suggère que le complexe Pop5-Rpp1 a une fonction équivalente à celle de la protéine C5 impliquée dans l'activation du domaine catalytique. Le module arrière des RNases P/MRP eucaryotes remplacerait donc la protéine bactérienne mais en fournissant une coordination plus sophistiquée de la séquence 5' leader de l'ARNt au cours de sa maturation.

Chez *S. cerevisiae*, la protéine Pop8 est une protéine apparentée à Pop5 (Rosenblad *et al.*, 2006). Or, une légère asymétrie est observée dans le module arrière de notre modèle (**Figure 19.C**, panel du haut) et nos expériences de marquages aux anticorps montrent que Pop5 et Pop8 sont localisées dans des emplacements similaires, à proximité de Rpp1 (**Figure 21.B**). Il semble donc que chez la levure une des protéines Pop5 soit remplacée par Pop8 dans l'hétérotétramère.

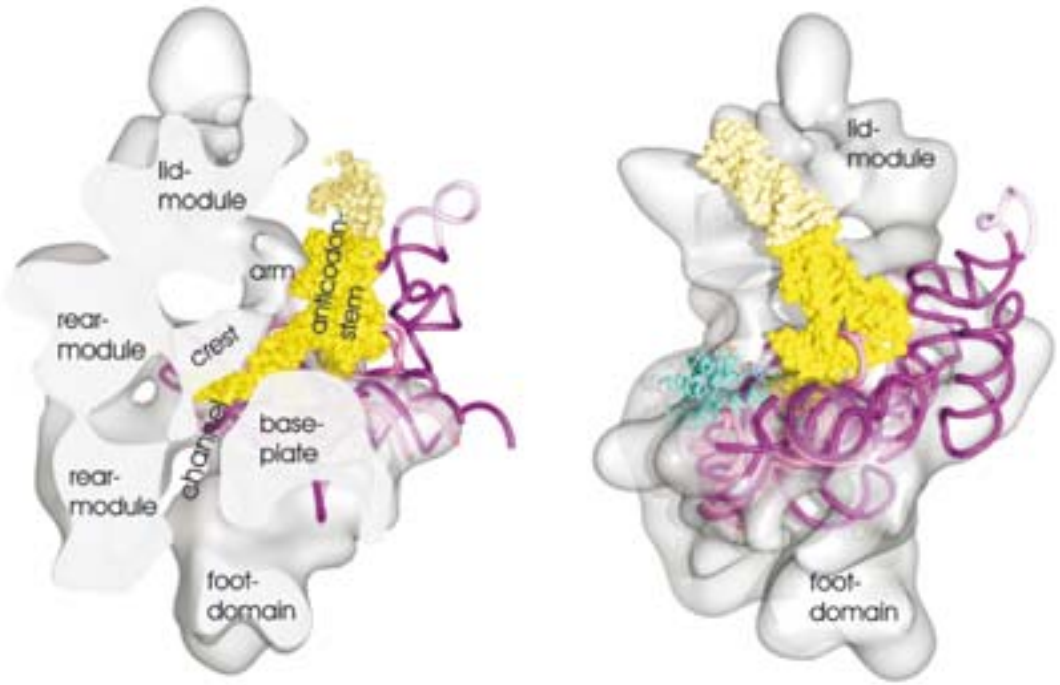


Figure 25. Superposition de la carte de la RNase P avec la structure de la RNase P bactérienne avec un ARNt (Reiter *et al.* , 2010)**.**

La sous-unité ARN est représentée en violet, l'ARNt en jaune et la protéine bactérienne en cyan. Le noyau de l'ARN bactérien a été utilisé pour le placement par superposition avec le modèle de *S. pombe* positionné (en magenta). Ce placement montre la position probable de l'ARNt lié, qui est inclus dans la structure à haute résolution mais absente dans la carte de microscopie électronique. La séquence du complexe bactérien a été modifiée pour favoriser les contacts cristallins au niveau de l'ARNt (jaune clair) et de la composante ARN (rose clair).

Panel de gauche : la carte et la structure superposées sont coupées le long du conduit central de la RNase P. Selon le positionnement de la sous-unité ARN bactérienne sur la carte, l'extrémité 3' de l'ARNt prémature entre dans le conduit.
Panel de droite : La vue est tournée de 90° selon l'axe vertical.

Rear module : module arrière ; *lid module* : module de tête ; *foot-domain* : pied du domaine de base ; *base-plate* : plaque basale du domaine de base ; *crest* : crête ; arm : bras du domaine de base ; *channel* : conduit ; *anticodon stem* : hélice anti-codon de l'ARNt.

3. Implication de l'architecture des complexes RNases P/MRP dans la coordination du substrat

Bien que les RNases P bactériennes et eucaryotes présentent des différences significatives au niveau de leurs sous-unités protéiques et du domaine S de la sous-unité ARN, elles ont en commun leurs domaines ARN catalytiques (Chen *et al.*, 1997) avec les régions conservées CR-I, CR-IV et CR-V qui forment le site catalytique et les régions CR-II et CR-III, responsables de l'identification des boucles T et D de l'ARNt (Krasilnikov *et al.*, 2004; Torres-Larios *et al.*, 2006; Reiter *et al.*, 2010). Etant donné que le substrat commun aux RNases P bactériennes et eucaryotes sont les ARNt prématures, il est probable que l'orientation du noyau catalytique et des régions impliquées dans l'identification du substrat soient comparables chez le complexe bactérien et le complexe eucaryote en dépit de leurs différences. Ainsi, en l'absence de structure à haute résolution du complexe eucaryote, la structure du complexe bactérien peut contribuer à localiser les éléments conservés correspondant dans le complexe RNase P eucaryote.

La superposition de la structure à haute résolution du complexe RNase P bactérien avec son ARNt (Reiter *et al.*, 2010) sur notre carte de microscopie électronique de RNase P eucaryote (**Figure 25**) montre les contributions probables du bras du module de base et du module de tête à la fixation d'un substrat de type ARNt. De plus, l'ARNt dans le complexe tridimensionnel est orienté de façon à ce que l'extrémité 3' de l'ARNt pénètre dans le conduit entre le module arrière et le module de base en direction du module arrière (**Figure 16**, panel de gauche). Par ailleurs, selon ce positionnement de la composante ARN bactérienne sur notre carte, l'entrée du conduit dans la RNase P eucaryote serait éloignée d'environ 2 nm des résidus identifiés sur l'hélice P4 comme constituant le site actif de l'enzyme bactérienne (Reiter *et al.*, 2010). De plus, dans l'enzyme bactérienne, la séquence 5' leader est coordonnée par la protéine unique. Or, dans la carte de la RNase P eucaryote, aucune densité ne semble correspondre à celle de la protéine bactérienne positionnée selon la structure de (Reiter *et al.*, 2010). Ces observations suggèrent que la séquence 5' leader est coordonnée lors de sa maturation d'une manière différente chez les eucaryotes. La séquence 5' leader pourrait, comme l'ARNt mature, entrer dans le conduit et interagir avec les conduits et cavités présents entre le module de base et le module arrière.

A distance du site catalytique, la structure bactérienne n'est pas parfaitement contenue dans la densité de la carte du complexe eucaryote. Néanmoins, il est concevable que le bras du domaine de base, où nous plaçons Pop3, guide l'hélice anticodon de l'ARNt prémature à distance du site catalytique. De plus, le module de tête, constitué de Pop1, restreint la

longueur maximale possible de l'hélice anticodon de l'ARNt (**Figure 16**, panel de droite), ce qui est cohérent avec la capacité de Pop1 à fixer le substrat (Marvin *et al.*, 2011). Ainsi, la poche de fixation du substrat ne reconnaitrait pas uniquement les séquences spécifiques mais identifierait également la forme en L, caractéristique des ARNt.

Dans le complexe bactérien, le domaine S de la composante ARN, formant une partie du bras du module de base de la RNase P eucaryote, se trouve éloigné de l'hélice anticodon de l'ARNt (Reiter *et al.*, 2010) et n'est pas contenu dans les densités de la carte de microscopie électronique (**Figure 25**, panel droite). Il semble donc que le domaine S dans le complexe bactérien ait une conformation plus ouverte que dans le complexe eucaryote. Ce type de structure relâchée est inattendu pour le domaine S, particulièrement si elle ne contribue pas directement à la fixation du substrat ou à la catalyse de sa coupure. Nous supposons que, dans le complexe bactérien, le domaine S a probablement été stabilisé dans une conformation ouverte grâce aux nucléotides additionnels requis pour la cristallisation, localisés à l'extrémité du bras anticodon de l'ARNt et à l'extrémité de l'hélice P12 de la sous-unité ARN. Ainsi, le noyau ARN bactérien pourrait présenter une certaine flexibilité adoptant une conformation ouverte pour faciliter une liaison initiale lâche au substrat et une conformation plus compacte, similaire à celle observée dans le complexe eucaryote, où la fixation du substrat est renforcée lorsque son positionnement est ajusté dans l'enzyme.

Selon notre modèle de RNase P (**Figure 26**), les sous-unités spécifiquement acquises chez les eucaryotes (Pop1, Pop6 et Pop7) sont localisées aux extrémités du complexe, où elles participent à l'identification et la fixation du substrat. Cela est cohérent avec l'hypothèse selon laquelle cette évolution pourrait avoir été motivée par la nécessité de faire face à l'accroissement de la complexité de l'environnement cellulaire. Nous supposons que grâce à l'ajout de ces protéines chez les eucaryotes, l'holoenzyme est alors capable d'identifier un éventail de structures spécifiques, qui couvre une proportion plus large de substrats ARN potentiels que dans le complexe bactérien.

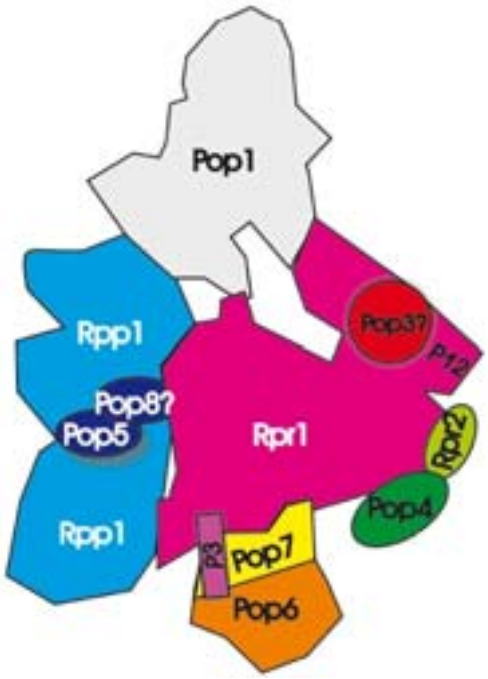


Figure 26. Représentation schématique de l'organisation des sous-unités formant la RNase P.

Les sous-unités pour lesquelles des structures ont pu être alignées sont montrées de la même couleur que dans la Figure 23. La protéine Pop1, dont il n'existe aucune structure résolue, est désignée en gris.

4. Différences architecturales des complexes des RNases P/MRP et spécificité des substrats

Dans le complexe RNase P, l'hétérodimère Pop4-Rpr2, localisé près du bras du module de base, forme une entité qui stabilise et protège le noyau catalytique ARN (Xu *et al.*, 2009). Dans le complexe RNase MRP, la protéine Rpr2 est remplacée par Snm1, une protéine homologue de plus haut poids moléculaire. Par ailleurs, le domaine S différant de manière importante entre les RNases P et MRP (**Figure 6**), une protéine spécifiquement adaptée est nécessaire pour la liaison sur ce domaine. Or, la protéine Snm1 lie Rmp1 (Aspinall *et al.*, 2007), l'autre protéine spécifique de la RNase MRP. Dans notre modèle de la RNase MRP, nous observons une densité supplémentaire près du bras du module de base (**Figure 20**, *) à proximité de la localisation prédite de Rpr2 ce qui parait cohérent avec le volume nécessaire aux protéines Snm1 et Rmp1 dans cet emplacement. Le complexe Pop4-Snm1-Rmp1 serait donc une entité spécifique de la RNase MRP, dédiée au soutien du domaine S du noyau ARN NME1.

De plus, la liaison entre le module arrière et le module de base conduit à la formation de conduits et de cavités propres à chaque complexe. Tandis que la RNase P possède une entrée circulaire d'un conduit traversant le complexe, la RNase MRP présente une fente ouverte sur le côté du domaine S à l'entrée du conduit qui est obstrué dans ce complexe (**Figure 20**, flèche ouverte, rond ouvert et ●). Ainsi, il semble que la voie de sortie de la séquence 5' leader coupée varie entre les deux complexes.

La protéine Pop1 est la plus large sous-unité protéique spécifique des RNases P/MRP eucaryotes (Lygerou *et al.*, 1996b). Elle est importante pour l'assemblage et les fonctions cellulaires des RNases P et MRP (Lygerou *et al.*, 1994; Srisawat *et al.*, 2002). Selon notre modèle, cette sous-unité n'entre pas en contact direct avec la sous-unité ARN mais forme le module de tête et délimite le site de liaison au substrat. Dans la RNase P, le module de tête entre contact avec le bras du module de base, ce qui conduit à la formation d'un site de liaison clos dont la forme ressemble à celle d'un ARNt. Dans la RNase MRP, le module de tête n'entre pas en contact avec le module de base au niveau de son bras (**Figure 19** et **Figure 20**, flèche noire). Cela confère une meilleure flexibilité au site de fixation permettant à la cavité liant le substrat d'accepter un large éventail de substrats. Une action différentielle de Pop1 selon le complexe avait d'ailleurs été mise en évidence lors d'analyse de mutations sur des résidus conservés de Pop1 (Xiao *et al.*, 2006).

5. Corrélation entre l'architecture des complexes P/MRP et les interactions binaires connues

Dans cette étude, nous présentons une carte de l'architecture des complexes RNases P et MRP (**Figure 23**) qui prend en compte la plupart de la densité observée. Bien que notre modèle soit en accord avec certaines des interactions décrites (Pluk *et al.*, 1999; Houser-Scott *et al.*, 2002; Hall *et al.*, 2004; Welting *et al.*, 2004; Welting *et al.*, 2006; Aspinall *et al.*, 2007; Khanova *et al.*, 2012), de nombreuses interactions protéines-protéines et protéines-ARN ne peuvent pas être confirmées. Ces données indiquent que les petites protéines de 20 à 40 kDa auraient au moins 5 contacts protéines-protéines, ce qui n'est pas compatible avec notre modèle dans lequel aucune sous-unité protéique n'est en contact direct avec plus de 3 autres protéines. Cela souligne les difficultés à déduire l'architecture de complexes à partir des seules études d'interactions.

Dans de nombreuses études, il a été mis en évidence que Pop1 entre en contact avec la composante ARN soit au niveau de l'hélice-boucle-hélice P3 (Pluk *et al.*, 1999; Ziehler *et al.*, 2001; Welting *et al.*, 2004) et/ou du domaine conservé CR-IV (Pluk *et al.*, 1999; Khanova *et al.*, 2012) et/ou de l'hélice P12 (Welting *et al.*, 2004). Cela ne correspond pas aux interactions requises dans notre modèle d'un complexe RNase P/MRP fonctionnel. Or, la protéine Pop1 est fortement impliquée dans la maturation du complexe RNase P/MRP, notamment lors de la coupure de la séquence 5' leader de la composante ARN ce qui ne serait pas le cas de Pop3 et Rpr2 (Srisawat *et al.*, 2002; Xiao *et al.*, 2006). Il est donc possible que Pop1 puisse entrer en contact direct avec la sous-unité ARN en divers emplacements. Il pourrait ainsi contribuer à la protection de l'hélice-boucle-hélice P3 en attendant le recrutement du sous-complexe Pop7-Pop6 lors de l'assemblage des RNases P/MRP ou encore à la protection de l'hélice P12 en l'absence de Rpr2 et Pop3.

Enfin, excepté Rpp1 et peut-être Pop5, selon notre modèle, il ne semble qu'aucune autre sous-unité protéique ne se trouve en plusieurs copies dans les RNases P et MRP, contrairement à l'estimation de la stœchiométrie réalisée avec du SYBRO Ruby (Salinas *et al.*, 2005). Ces divergences pourraient s'expliquer par la formation de dimères de complexe apparaissant sous certaines conditions dans la RNase P bactérienne (Barrera *et al.*, 2002; Buck *et al.*, 2005a) ou la formation de super structures dans la cellule.

6. Conclusion et perspectives

En conclusion, notre travail a permis de montrer que les RNases P et MRP eucaryotes s'organisent selon une architecture modulaire commune. Les modules acquis chez les archées et conservés chez les eucaryotes peuvent être interprétés comme gains de l'évolution contribuant à l'identification et la coordination du substrat. Chez les eucaryotes, des changements subtils à l'interface entre les modules de la RNase P et de la RNase MRP altèrent la forme des cavités et du conduit liant le substrat. Ces données sont en accord avec leur spécificité pour des substrats différents et souligne l'importance de la coordination de la séquence 5' leader du substrat.

Pour poursuivre cette étude, il serait intéressant de caractériser la structure des protéines composant ces holoenzymes afin de mieux déterminer les spécificités eucaryotes par rapport aux protéines archées. Le principal obstacle à cette caractérisation concerne l'expression et la purification des protéines individuellement, étant donné que certaines sont difficiles à exprimer, insolubles et uniquement stables transitoirement (Houser-Scott et al., 2002; Aspinall et al., 2007; Perederina et al., 2007). Sur la base de l'architecture modulaire des deux holoenzymes suggérée par notre étude, nous envisageons de co-exprimer les protéines correspondantes aux différents modules : Rpp1-Pop5, Rpr2-Pop4, Smn1-Pop4, Snm1-Rmp1, Smn1-Rmp1-Pop4 et Pop3 avec l'hélice ARN P12. La co-expression en baculo-virus permettrait de se rapprocher des conditions d'expression eucaryotes avec la possibilité de modifications post-traductionnelles afin que les protéines adoptent leur repliement physiologique pour tester leurs conditions de cristallisation.

Pour approfondir la compréhension du mécanisme de reconnaissance du substrat des RNases P et MRP, il sera aussi intéressant de réaliser une carte en microscopie électronique des holoenzymes en présence de leurs substrats respectifs : ARNt prématures pour la RNase P, ARNm CLB2 ou ARN35S pour la RNase MRP. Cela nous permettrait de visualiser les probables remaniements conformationnels liés à la présence du substrat d'une part et de mieux comprendre le rôle des différences structurales entre les RNases P et MRP d'autre part.

Enfin, bien que de nombreuses études décrivent les interactions entre deux protéines, elles ne délivrent aucune information sur la stœchiométrie des sous-unités protéiques dans la RNase P. Une analyse préliminaire des sous-unités protéiques de la RNase MRP de *S. cerevisiae* (Salinas et al., 2005) suggère que la composition en protéine des holoenzymes eucaryotes pourrait être très complexe avec de nombreuses sous-unités présentes en plusieurs exemplaires. Or, le positionnement des différentes protéines archées dans nos cartes de microscopie électronique suggère que seules les protéines Rpp1 et Pop5 seraient présentes en

deux exemplaires dans ces holoenzymes tandis que les autres protéines seraient présentes en un seul exemplaire. Il serait donc intéressant de caractériser la masse moléculaire des RNases P et MRP par une analyse en spectrométrie de masse en condition native de façon à définir la stœchiométrie des différentes sous-unités protéiques dans les RNases P et MRP de *S. cerevisiae*.

ANNEXES

Annexe I : Tableau récapitulatif des souches de *S. cerevisiae* utilisées dans cette étude.

Nom de la souche	Génotype	Description
YSW1	MATa POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1::LEU2 sep1::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112	Schmitt *et al.*, 1994
YSW1-HA-Snm1	MATa His3MX6::PGal1::3HA::Snm1 POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1 ::LEU2 sep1 ::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112	Hipp *et al.*, 2012
YSW1-Rpr2-HA	MATa Rpr2::3HA::His3MX6 POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1::LEU2 sep1::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112	Hipp *et al.*, 2012
YSW1-Rpr2-HA /GFP-Pop1	MATa natNT2::PGal1::yeGFP::Pop1 Rpr2::3HA::His3MX6 POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1::LEU2 sep1::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112	Hipp *et al.*, 2012
YSW1-Rpr2-HA /Rpp1-GFP	MATa Rpp1::yeGFP::hphNT1 Rpr2::3HA::His3MX6 POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1::LEU2 sep1::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112	Hipp *et al.*, 2012
YSW1-Rpr2-HA /Pop5-GFP	MATa Pop5::yeGFP::hphNT1 Rpr2::3HA::His3MX6 POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1::LEU2 sep1::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112	Hipp *et al.*, 2012
YSW1-Rpr2-HA /Pop6-GFP	MATa Pop6::yeGFP::hphNT1 Rpr2::3HA::His3MX6 POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1::LEU2 sep1::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112	Hipp *et al.*, 2012
YSW1-Rpr2-HA /Pop8-GFP	MATa Pop8::yeGFP::hphNT1 Rpr2::3HA::His3MX6 POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1::LEU2 sep1::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112	Hipp *et al.*, 2012
YSW1-GFP-Rpr2-HA	MATa natNT2::PGal1::yeGFP::Rpr2::3HA::His3MX6 POP4::TAPTAG::TRP1ks pep4::LEU2 nuc1::LEU2 sep1::URA3, trp1his3-11,15 can-100 ura3-1 leu2-3,112	Hipp *et al.*, 2012

Annexe II : Amorces utilisées pour réaliser les intégrations génomiques des cassettes 3HA et yeGFP.

Nom de l'amorce	Séquence de l'amorce
Rpr2-HA_F2	CTAATTACAGGACCTACTCGGAGCGGGAGGGTAATCTACTAAATTCTcggatccccgggttaattaa
Rpr2-HA_R1	CCATGTATAGGGGCAAGGGGGTGGGGGGTCGGCTCTTAAAGAAAATCTGGgaattcgagctcgtttaaac
PGAL-HA-Snm1_F4	GAAAACGCTATTCTATGTGAAAATAGAGATGAGTACTACTAGTGGCTAgaattcgagctcgtttaaac
PGAL-HA-Snm1_R3	GGTTGTATTTTTGTCGCAGACTCCGTTCTTGATATTTTTCTGCCTGGTCTTTATTgcactgagcagcgtaatctg
PGAL-yeGFP-Pop1_S4	GATTCTTATTAAGAACTTTCTTTCCACCATTTCCTCTAGACAAACTCCCGCTcatcgatgaattctctgtcg
PGAL-GFP-Pop1_S1	GCCACTTGAACATTTGGCAAGAAAGGTGAGAATTGACCTCATTATAATTACAAATGcgtacgctgcaggtcgac
Rpp1-yeGFP_S2	CGTATTTAATTATCATTTATACATTATATGTTATAGGTTAGCGTATCGTTTTAatcgatgaattcgagctcg
Rpp1-yeGFP_S3	GGATGCTGAACAATTAGGCCATGCTTCCAAAAGACACAAGCCTcgtacgctgcaggtcgac
Pop5-yeGFP_S3	CAAGAAAATTGGAAGGGAAAACGAAAACGAAAACGAGGACGATcgtacgctgcaggtcgac
Pop5-yeGFP_S2	AACTACTGCATATACGTGTACATCTATATATTAATATATTCTAatcgatgaattcgagctcg
Pop6-yeGFP_S3	GATTCAGAGATCATTGATTTAAACCTACATTCATTCACCAAACAAcgtacgctgcaggtcgac
Pop6-yeGFP_S2	GTACAAATGCAAGAAGTAGCTATGGATGCTAAAGAGCTGGTCGATTTAatcgatgaattcgagctcg
Pop8-yeGFP_S3	GGCTGAAAAAAGTAATGGAAGAAGAAGAACAAGACTGTAAATGTATAcgtacgctgcaggtcgac
Pop8-yeGFP_S2	CACTTTCTTCTATGCCCAAGTTTTTTTAAATAATAAACTGATATTCTAatcgatgaattcgagctcg
PGAL-yeGFP-Rpr2_S1	CTACTGTCCTTCCCACATCGCCGACAGAGAAAAGACAAGCAAAAAGATGcgtacgctgcaggtcgac
PGAL-yeGFP-Rpr2_S4	CCATTTTCGTCGATCTCAGGTTTCATCTTTCCTCCGTGTGCTTTCTTGCCcatcgatgaattctctgtcg

Séquence en lettre majuscule : séquence homologue en amont ou en aval du point d'intégration.

Séquence en minuscule : séquence spécifique de la cassette à amplifier caractérisée par les lettres S1/S4 S2/S3 dans les cassettes de (Janke *et al.*, 2004) et F2/R1 F4/R3 dans les cassettes de (Longtine *et al.*, 1998).

ABREVIATIONS

ADN : Acide désoxyribonucléique

ADNc : ADN complémentaire

ARN : Acide ribonucléique

ARNm : ARN messager

ARNr : ARN ribosomique

ARNt : ARN de transfert

CBP : *Calmodulin-Binding Peptide*

Ct : *Cycle threshold*

dNTP : désoxyribonucléoside triphosphate

DMSO : Diméthylsulfoxide

DO : Densité optique

DNase : Désoxyribonucléase

DTT : Dithiothreitol

EMSA : Electrophoretic Mobility Shift Assay

FSC : *Fourier shell correlation,* corrélation des enveloppes de Fourier

GFP : Green Fluorescent Protein

HA : Hemagglutinin Antigen

k_{cat} : Constante catalytique d'une réaction enzymatique

kDa : kilo Dalton

Km : Constante de Michaelis propre à une enzyme

min : Minutes

PAGE : *PolyAcrylamide gel electrophoresis*

pI : Point isoélectrique

pb : Paire de base

PBS : Phosphate buffered saline

PCR : Polymerase Chain Reaction

PEG : Polyéthylèneglycol

pré-ARNt : Précurseur d'ARNt

qPCR : PCR quantitative

rNTP : Ribonucléoside triphosphate

RNase : Ribonuclease

RNP : Ribonucléoprotéine

Rpp : RNase P protein

tr/m : Tour par minute

SDS : Sodium dodecylsulfate

SGD : Saccharomyces Genome Database

snoARN : Petit ARN nucléolaire

snRNA : *small nuclear RNA*

TAP : Double purification d'affinité (*Tandem Affinity Purification*)

TBE : Tris Borate EDTA

UV : Ultra-violet

V-ATPase : ATPase vacuolaire

Vmax : Vitesse maximale d'une réaction enzymatique

yeGFP : Yeast enhanced GFP

YPD : Yeast extract Peptone Dextrose

Autres travaux de recherche

En parallèle de cette étude sur les RNases P et MRP, j'ai mené d'autres travaux de recherche sur des sujets variés qui ont fait l'objet de plusieurs publications :

Publication n°1 :

Diepholz M., Venzke D., Prinz S., **Batisse C.**, Flörchinger B., Rössle M., Svergun D.I., Böttcher B., Féthière J. A different conformation for EGC stator subcomplex in solution and in the assembled yeast V-ATPase: Possible implications for regulatory disassembly. *Structure*, 2008, vol. 16, p. 1789-1798.

J'ai contribué à cette étude en choisissant et en mettant en œuvre une stratégie d'étiquetage de la protéine E de la V-ATPase avec une étiquette Flag ce qui a permis la localisation de cette protéine sur le complexe V-ATPase, étape essentielle à la finalisation de la caractérisation structurale de ce complexe.

Publication n°2 :

Batisse J., **Batisse C.**, Budd A., Böttcher B., Hurt E. Purification of nuclear poly(A)-binding protein Nab2 reveals association with the yeast transcriptome and a messenger ribonucleoprotein (mRNP) core structure. *Journal of Biological Chemistry*, 2009, vol. 284, p. 34911-34917.

J'ai contribué à cette étude des mRNP en conseillant l'auteur principal sur les meilleures approches à utiliser pour purifier les particules afin de les visualiser par microscopie électronique, en préparant les échantillons et en collectant les données pour l'analyse en microscopie électronique qui ont permis pour la première fois la visualisation de mRNP purifiées *in vivo*.

Publication n°3 :

Kühner S., van Noort V., Betts M.J., Leo-Macias A., **Batisse C.**, Rode M., Yamada T., Maier T., Bader S., Beltran-Alvarez P., Castaño-Diez D., Chen W-H., Devos D., Güell M., Norambuena T., Racke I., Rybin V., Schmidt A., Yus E., Aebersold R., Herrmann R., Böttcher B., Frangakis A.S., Russell R.B., Serrano L., Bork P., Gavin A-C. Proteome organization in a genome-reduced bacterium. *Science*, 2009, vol. 326, p. 1235-1240.

Ma contribution à cette étude concerne le choix des complexes à caractériser par microscopie électronique et traitement d'images, le choix et la mise en œuvre des stratégies de purification, la préparation des échantillons et la collecte des données permettant ainsi de participer à la première visualisation de l'organisation d'une cellule bactérienne à partir de

données protéomiques. En outre, j'ai plus particulièrement participé à la rédaction des parties du Matériel et Méthode concernant mes résultats et à la mise en forme de la figure S6.

Publication n°4 :

Ernst S., **Batisse C.**, Zarrabia N., Böttcher B., Börsch M. Regulatory assembly of the vacuolar proton pump VoV1-ATPase in yeast cells by FLIM-FRET. *SPIE Proceedings paper,* 2010, vol. 7569, 75690W.

Pour cette étude, j'ai créé 41 souches de *S. cerevisiae* marquées avec différents fluorophores sur différentes protéines de la V-ATPase dont j'ai testé l'expression par microscopie optique à fluorescence. Ce travail a généré un large éventail de couples de fluorophores permettant d'une part de déterminer le meilleur couple de fluorophores pour réaliser une analyse de FLIM-FRET dans la levure et d'autre part de suivre pour la première fois *in vivo* le processus d'assemblage et de désassemblage des deux sous-unités de la V-ATPase.

Publication n°5 :

Nasertorabi F., **Batisse C.**, Diepholz M., Suck D. and Böttcher B. Insights into the structure of the CCR4-Not complex by electron microscopy. *FEBS Letters*, 2011, vol. 585, p. 2182-2186.

Ma participation à cette étude du complexe de déadénylation CCR4-Not concerne le conseil à l'auteur principal sur différentes approches à utiliser pour purifier les particules, la visualisation du complexe purifié par électrophorèse en conditions natives, la préparation des échantillons et la collecte de données pour l'analyse en microscopie électronique permettant la caractérisation structurale du complexe CCR4-Not. En outre, j'ai activement participé à la mise en forme de la figure 1 et à la discussion dans cet article.

Références bibliographiques du sujet de recherche

Abrahams, J. P., Leslie, A. G., Lutter, R., and Walker, J. E. Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. *Nature*, 1994, Vol 370, n° 6491, p. 621-628.

Alberts, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*, 1998, Vol 92, n° 3, p. 291-294.

Alifano, P., Rivellini, F., Piscitelli, C., Arraiano, C. M., Bruni, C. B., and Carlomagno, M. S. Ribonuclease E provides substrates for ribonuclease P-dependent processing of a polycistronic mRNA. *Genes Dev*, 1994, Vol 8, n° 24, p. 3021-3031.

Altman, S. Isolation of tyrosine tRNA precursor molecules. *Nat New Biol*, 1971, Vol 229, n° 1, p. 19-21.

Altman, S. Ribonuclease P: an enzyme with a catalytic RNA subunit. *Adv Enzymol Relat Areas Mol Biol*, 1989, Vol 62, n° 1-36.

Altman, S., Wesolowski, D., Guerrier-Takada, C., and Li, Y. RNase P cleaves transient structures in some riboswitches. *Proc Natl Acad Sci U S A*, 2005, Vol 102, n° 32, p. 11284-11289.

Amero, C. D., Boomershine, W. P., Xu, Y., and Foster, M. Solution structure of Pyrococcus furiosus RPP21, a component of the archaeal RNase P holoenzyme, and interactions with its RPP29 protein partner. *Biochemistry*, 2008, Vol 47, n° 45, p. 11704-11710.

Anantharaman, V., and Aravind, L. The NYN domains: novel predicted RNAses with a PIN domain-like fold. *RNA Biol*, 2006, Vol 3, n° 1, p. 18-27.

Andrews, A. J., Hall, T. A., and Brown, J. W. Characterization of RNase P holoenzymes from Methanococcus jannaschii and Methanothermobacter thermoautotrophicus. *Biol Chem*, 2001, Vol 382, n° 8, p. 1171-1177.

Aravind, L., Iyer, L. M., and Anantharaman, V. The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. *Genome Biol*, 2003, Vol 4, n° 10, p. R64.

Aspinall, T. V., Gordon, J. M., Bennett, H. J., Karahalios, P., Bukowski, J. P., Walker, S. C., Engelke, D. R., and Avis, J. M. Interactions between subunits of Saccharomyces cerevisiae RNase MRP support a conserved eukaryotic RNase P/MRP architecture. *Nucleic Acids Res*, 2007, Vol 35, n° 19, p. 6439-6450.

Baird, N. J., Fang, X. W., Srividya, N., Pan, T., and Sosnick, T. R. Folding of a universal ribozyme: the ribonuclease P RNA. *Q Rev Biophys*, 2007, Vol 40, n° 2, p. 113-161.

Baird, N. J., Gong, H., Zaheer, S. S., Freed, K. F., Pan, T., and Sosnick, T. R. Extended structures in RNA folding intermediates are due to nonnative interactions rather than electrostatic repulsion. *J Mol Biol*, 2010, Vol 397, n° 5, p. 1298-1306.

Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, 2000, Vol 289, n° 5481, p. 905-920.

Barrera, A., Fang, X., Jacob, J., Casey, E., Thiyagarajan, P., and Pan, T. Dimeric and monomeric Bacillus subtilis RNase P holoenzyme in the absence and presence of pre-tRNA substrates. *Biochemistry*, 2002, Vol 41, n° 43, p. 12986-12994.

Bartkiewicz, M., Gold, H., and Altman, S. Identification and characterization of an RNA molecule that copurifies with RNase P activity from HeLa cells. *Genes Dev*, 1989, Vol 3, n° 4, p. 488-499.

Beebe, J. A., Kurz, J. C., and Fierke, C. A. Magnesium ions are required by Bacillus subtilis ribonuclease P RNA for both binding and cleaving precursor $tRNA_{Asp}$. *Biochemistry*, 1996, Vol 35, n° 32, p. 10493-10505.

Ben-Shem, A., Jenner, L., Yusupova, G., and Yusupov, M. Crystal structure of the eukaryotic ribosome. *Science*, 2010, Vol 330, n° 6008, p. 1203-1209.

Bertrand, E., Houser-Scott, F., Kendall, A., Singer, R. H., and Engelke, D. R. Nucleolar localization of early tRNA processing. *Genes Dev*, 1998, Vol 12, n° 16, p. 2463-2468.

Bjerrum, O. J., Schafer-Nielsen, C. (1986). Analytical Electrophoresis (Veinheim).

Bonafe, L., Dermitzakis, E. T., Unger, S., Greenberg, C. R., Campos-Xavier, B. A., Zankl, A., Ucla, C., Antonarakis, S. E., Superti-Furga, A., and Reymond, A. Evolutionary comparison provides evidence for pathogenicity of RMRP mutations. *PLoS Genet*, 2005, Vol 1, n° 4, p. e47.

Boomershine, W. P., McElroy, C. A., Tsai, H. Y., Wilson, R. C., Gopalan, V., and Foster, M. P. Structure of Mth11/Mth Rpp29, an essential protein subunit of archaeal and eukaryotic RNase P. *Proc Natl Acad Sci U S A*, 2003, Vol 100, n° 26, p. 15398-15403.

Bothwell, A. L., Stark, B. C., and Altman, S. Ribonuclease P substrate specificity: cleavage of a bacteriophage phi80-induced RNA. *Proc Natl Acad Sci U S A*, 1976, Vol 73, n° 6, p. 1912-1916.

Bottcher, B., Wynne, S. A., and Crowther, R. A. Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature*, 1997, Vol 386, n° 6620, p. 88-91.

Bowman, L. H., Goldman, W. E., Goldberg, G. I., Hebert, M. B., and Schlessinger, D. Location of the initial cleavage sites in mouse pre-rRNA. *Mol Cell Biol*, 1983, Vol 3, n° 8, p. 1501-1510.

Brenner, S., and Horne, R. W. A negative staining method for high resolution electron microscopy of viruses. *Biochim Biophys Acta*, 1959, Vol 34, n° 103-110.

Brown, J. W., Haas, E. S., James, B. D., Hunt, D. A., Liu, J. S., and Pace, N. R. Phylogenetic analysis and evolution of RNase P RNA in proteobacteria. *J Bacteriol*, 1991, Vol 173, n° 12, p. 3855-3863.

Brown, J. W., and Pace, N. R. Ribonuclease P RNA and protein subunits from bacteria. *Nucleic Acids Res*, 1992, Vol 20, n° 7, p. 1451-1456.

Brown, J. W., Haas, E. S., and Pace, N. R. Characterization of ribonuclease P RNAs from thermophilic bacteria. *Nucleic Acids Res*, 1993, Vol 21, n° 3, p. 671-679.

Brown, J. W., Nolan, J. M., Haas, E. S., Rubio, M. A., Major, F., and Pace, N. R. Comparative analysis of ribonuclease P RNA using gene sequences from natural microbial populations reveals tertiary structural elements. *Proc Natl Acad Sci U S A*, 1996, Vol 93, n° 7, p. 3001-3006.

Brown, J. W. The Ribonuclease P Database. *Nucleic Acids Res*, 1999, Vol 27, n° 1, p. 314.

Brusca, E. M., True, H. L., and Celander, D. W. Novel RNA-binding properties of Pop3p support a role for eukaryotic RNase P protein subunits in substrate recognition. *J Biol Chem*, 2001, Vol 276, n° 45, p. 42543-42548.

Buck, A. H., Dalby, A. B., Poole, A. W., Kazantsev, A. V., and Pace, N. R. Protein activation of a ribozyme: the role of bacterial RNase P protein. *EMBO J*, 2005a, Vol 24, n° 19, p. 3360-3368.

Buck, A. H., Kazantsev, A. V., Dalby, A. B., and Pace, N. R. Structural perspective on the activation of RNAse P RNA by protein. *Nat Struct Mol Biol*, 2005b, Vol 12, n° 11, p. 958-964.

Cai, T., Reilly, T. R., Cerio, M., and Schmitt, M. E. Mutagenesis of SNM1, which encodes a protein component of the yeast RNase MRP, reveals a role for this ribonucleoprotein endoribonuclease in plasmid segregation. *Mol Cell Biol*, 1999, Vol 19, n° 11, p. 7857-7869.

Cai, T., Aulds, J., Gill, T., Cerio, M., and Schmitt, M. E. The Saccharomyces cerevisiae RNase mitochondrial RNA processing is critical for cell cycle progression at the end of mitosis. *Genetics*, 2002, Vol 161, n° 3, p. 1029-1042.

Cassano, A. G., Anderson, V. E., and Harris, M. E. Analysis of solvent nucleophile isotope effects: evidence for concerted mechanisms and nucleophilic activation by metal coordination in nonenzymatic and ribozyme-catalyzed phosphodiester hydrolysis. *Biochemistry*, 2004, Vol 43, n° 32, p. 10547-10559.

Chamberlain, J. R., Pagan, R., Kindelberger, D. W., and Engelke, D. R. An RNase P RNA subunit mutation affects ribosomal RNA processing. *Nucleic Acids Res*, 1996, Vol 24, n° 16, p. 3158-3166.

Chamberlain, J. R., Lee, Y., Lane, W. S., and Engelke, D. R. Purification and characterization of the nuclear RNase P holoenzyme complex reveals extensive subunit overlap with RNase MRP. *Genes Dev*, 1998, Vol 12, n° 11, p. 1678-1690.

Chang, D. D., and Clayton, D. A. A mammalian mitochondrial RNA processing activity contains nucleus-encoded RNA. *Science*, 1987a, Vol 235, n° 4793, p. 1178-1184.

Chang, D. D., and Clayton, D. A. A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. *EMBO J*, 1987b, Vol 6, n° 2, p. 409-417.

Chang, D. D., and Clayton, D. A. Mouse RNAase MRP RNA is encoded by a nuclear gene and contains a decamer sequence complementary to a conserved region of mitochondrial RNA substrate. *Cell*, 1989, Vol 56, n° 1, p. 131-139.

Charles, J. F., Jaspersen, S. L., Tinker-Kulberg, R. L., Hwang, L., Szidon, A., and Morgan, D. O. The Polo-related kinase Cdc5 activates and is destroyed by the mitotic cyclin destruction machinery in S. cerevisiae. *Curr Biol*, 1998, Vol 8, n° 9, p. 497-507.

Chen, J. L., and Pace, N. R. Identification of the universally conserved core of ribonuclease P RNA. *RNA*, 1997, Vol 3, n° 6, p. 557-560.

Chen, J. L., Nolan, J. M., Harris, M. E., and Pace, N. R. Comparative photocross-linking analysis of the tertiary structures of Escherichia coli and Bacillus subtilis RNase P RNAs. *EMBO J*, 1998, Vol 17, n° 5, p. 1515-1525.

Chen, W. Y., Pulukkunat, D. K., Cho, I. M., Tsai, H. Y., and Gopalan, V. Dissecting functional cooperation among protein subunits in archaeal RNase P, a catalytic ribonucleoprotein complex. *Nucleic Acids Res*, 2010, Vol 38, n° 22, p. 8316-8327.

Chen, W. Y., Xu, Y., Cho, I. M., Oruganti, S. V., Foster, M. P., and Gopalan, V. Cooperative RNP assembly: complementary rescue of structural defects by protein and RNA subunits of archaeal RNase P. *J Mol Biol*, 2011, Vol 411, n° 2, p. 368-383.

Cho, I. M., Lai, L. B., Susanti, D., Mukhopadhyay, B., and Gopalan, V. Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc Natl Acad Sci U S A*, 2010, Vol 107, n° 33, p. 14573-14578.

Chu, S., Archer, R. H., Zengel, J. M., and Lindahl, L. The RNA of RNase MRP is required for normal processing of ribosomal RNA. *Proc Natl Acad Sci U S A*, 1994, Vol 91, n° 2, p. 659-663.

Chu, S., Zengel, J. M., and Lindahl, L. A novel protein shared by RNase MRP and RNase P. *RNA*, 1997, Vol 3, n° 4, p. 382-391.

Coughlin, D. J., Pleiss, J. A., Walker, S. C., Whitworth, G. B., and Engelke, D. R. Genome-wide search for yeast RNase P substrates reveals role in maturation of intron-encoded box C/D small nucleolar RNAs. *Proc Natl Acad Sci U S A*, 2008, Vol 105, n° 34, p. 12218-12223.

Cramer, P., Bushnell, D. A., and Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*, 2001, Vol 292, n° 5523, p. 1863-1876.

Crary, S. M., Niranjanakumari, S., and Fierke, C. A. The protein component of Bacillus subtilis ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-tRNAAsp. *Biochemistry*, 1998, Vol 37, n° 26, p. 9409-9416.

Darr, S. C., Pace, B., and Pace, N. R. Characterization of ribonuclease P from the archaebacterium Sulfolobus solfataricus. *J Biol Chem*, 1990, Vol 265, n° 22, p. 12927-12932.

Dichtl, B., and Tollervey, D. Pop3p is essential for the activity of the RNase MRP and RNase P ribonucleoproteins in vivo. *EMBO J*, 1997, Vol 16, n° 2, p. 417-429.

Diepholz, M., Venzke, D., Prinz, S., Batisse, C., Florchinger, B., Rossle, M., Svergun, D. I., Bottcher, B., and Fethiere, J. A different conformation for EGC stator subcomplex in solution and in the assembled yeast V-ATPase: possible implications for regulatory disassembly. *Structure*, 2008, Vol 16, n° 12, p. 1789-1798.

Dlakic, M. 3D models of yeast RNase P/MRP proteins Rpp1p and Pop3p. *RNA*, 2005, Vol 11, n° 2, p. 123-127.

Doudna, J. A., and Lorsch, J. R. Ribozyme catalysis: not different, just worse. *Nat Struct Mol Biol*, 2005, Vol 12, n° 5, p. 395-402.

Eder, P. S., Kekuda, R., Stolc, V., and Altman, S. Characterization of two scleroderma autoimmune antigens that copurify with human ribonuclease P. *Proc Natl Acad Sci U S A*, 1997, Vol 94, n° 4, p. 1101-1106.

Ellis, J. C., Barnes, J., and Brown, J. W. Is Alba an RNase P subunit? *RNA Biol*, 2007, Vol 4, n° 3, p. 169-172.

Esakova, O., Perederina, A., Quan, C., Schmitt, M. E., and Krasilnikov, A. S. Footprinting analysis demonstrates extensive similarity between eukaryotic RNase P and RNase MRP holoenzymes. *RNA*, 2008, Vol 14, n° 8, p. 1558-1567.

Esakova, O., and Krasilnikov, A. S. Of proteins and RNA: the RNase P/MRP family. *RNA*, 2010, Vol 16, n° 9, p. 1725-1747.

Esakova, O., Perederina, A., Quan, C., Berezin, I., and Krasilnikov, A. S. Substrate recognition by ribonucleoprotein ribonuclease MRP. *RNA*, 2011, Vol 17, n° 2, p. 356-364.

Evans, D., Marquez, S. M., and Pace, N. R. RNase P: interface of the RNA and protein worlds. *Trends Biochem Sci*, 2006, Vol 31, n° 6, p. 333-341.

Fan, H., Goodier, J. L., Chamberlain, J. R., Engelke, D. R., and Maraia, R. J. 5' processing of tRNA precursors can Be modulated by the human La antigen phosphoprotein. *Mol Cell Biol*, 1998, Vol 18, n° 6, p. 3201-3211.

Fernandez-Tornero, C., Bottcher, B., Riva, M., Carles, C., Steuerwald, U., Ruigrok, R. W., Sentenac, A., Muller, C. W., and Schoehn, G. Insights into transcription initiation and termination from the electron microscopy structure of yeast RNA polymerase III. *Mol Cell*, 2007, Vol 25, n° 6, p. 813-823.

Forster, A. C., and Altman, S. External guide sequences for an RNA enzyme. *Science*, 1990, Vol 249, n° 4970, p. 783-786.

Forti, F., Sabbattini, P., Sironi, G., Zangrossi, S., Deho, G., and Ghisotti, D. Immunity determinant of phage-plasmid P4 is a short processed RNA. *J Mol Biol*, 1995, Vol 249, n° 5, p. 869-878.

Frank, D. N., Adamidi, C., Ehringer, M. A., Pitulle, C., and Pace, N. R. Phylogenetic-comparative analysis of the eukaryal ribonuclease P RNA. *RNA*, 2000, Vol 6, n° 12, p. 1895-1904.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol*, 1996, Vol 116, n° 1, p. 190-199.

Fukuhara, H., Kifusa, M., Watanabe, M., Terada, A., Honda, T., Numata, T., Kakuta, Y., and Kimura, M. A fifth protein subunit Ph1496p elevates the optimum temperature for the ribonuclease P activity from Pyrococcus horikoshii OT3. *Biochem Biophys Res Commun*, 2006, Vol 343, n° 3, p. 956-964.

Garber, R. L., and Altman, S. In vitro processing of B. mori transfer RNA precursor molecules. *Cell*, 1979, Vol 17, n° 2, p. 389-397.

Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, 2002, Vol 415, n° 6868, p. 141-147.

Giegé, R., Florentz, C., and Dreher, T. W. The TYMV tRNA-like structure. *Biochimie*, 1993, Vol 75, n° 7, p. 569-582.

Gill, T., Cai, T., Aulds, J., Wierzbicki, S., and Schmitt, M. E. RNase MRP cleaves the CLB2 mRNA to promote cell cycle progression: novel method of mRNA degradation. *Mol Cell Biol*, 2004, Vol 24, n° 3, p. 945-953.

Gill, T., Aulds, J., and Schmitt, M. E. A specialized processing body that is temporally and asymmetrically regulated during the cell cycle in Saccharomyces cerevisiae. *J Cell Biol*, 2006, Vol 173, n° 1, p. 35-45.

Gimple, O., and Schön, A. In vitro and in vivo processing of cyanelle tmRNA by RNase P. *Biol Chem*, 2001, Vol 382, n° 10, p. 1421-1429.

Gobert, A., Gutmann, B., Taschner, A., Gossringer, M., Holzmann, J., Hartmann, R. K., Rossmanith, W., and Giegé, P. A single Arabidopsis organellar protein has RNase P activity. *Nat Struct Mol Biol*, 2010, Vol 17, n° 6, p. 740-744.

Golas, M. M., Sander, B., Will, C. L., Luhrmann, R., and Stark, H. Molecular architecture of the multiprotein splicing factor SF3b. *Science*, 2003, Vol 300, n° 5621, p. 980-984.

Gopalan, V., Baxevanis, A. D., Landsman, D., and Altman, S. Analysis of the functional role of conserved residues in the protein subunit of ribonuclease P from Escherichia coli. *J Mol Biol*, 1997, Vol 267, n° 4, p. 818-829.

Gopalan, V. Uniformity amid diversity in RNase P. *Proc Natl Acad Sci U S A*, 2007, Vol 104, n° 7, p. 2031-2032.

Gossringer, M., Kretschmer-Kazemi Far, R., and Hartmann, R. K. Analysis of RNase P protein (rnpA) expression in Bacillus subtilis utilizing strains with suppressible rnpA expression. *J Bacteriol*, 2006, Vol 188, n° 19, p. 6816-6823.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, 1983, Vol 35, n° 3 Pt 2, p. 849-857.

Guerrier-Takada, C., Eder, P. S., Gopalan, V., and Altman, S. Purification and characterization of Rpp25, an RNA-binding protein subunit of human ribonuclease P. *RNA*, 2002, Vol 8, n° 3, p. 290-295.

Gutmann, B., Gobert, A., and Giegé, P. PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. *Genes Dev*, 2012, Vol 26, n° 10, p. 1022-1027.

Haas, E. S., Banta, A. B., Harris, J. K., Pace, N. R., and Brown, J. W. Structure and evolution of ribonuclease P RNA in Gram-positive bacteria. *Nucleic Acids Res*, 1996, Vol 24, n° 23, p. 4775-4782.

Haas, E. S., and Brown, J. W. Evolutionary variation in bacterial RNase P RNAs. *Nucleic Acids Res*, 1998, Vol 26, n° 18, p. 4093-4099.

Hada, K., Nakashima, T., Osawa, T., Shimada, H., Kakuta, Y., and Kimura, M. Crystal structure and functional analysis of an archaeal chromatin protein Alba from the hyperthermophilic archaeon Pyrococcus horikoshii OT3. *Biosci Biotechnol Biochem*, 2008, Vol 72, n° 3, p. 749-758.

Haeusler, R. A., Pratt-Hyatt, M., Good, P. D., Gipson, T. A., and Engelke, D. R. Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes. *Genes Dev*, 2008, Vol 22, n° 16, p. 2204-2214.

Hall, T. A., and Brown, J. W. The ribonuclease P family. *Methods Enzymol*, 2001, Vol 341, n° 56-77.

Hall, T. A., and Brown, J. W. Archaeal RNase P has multiple protein subunits homologous to eukaryotic nuclear RNase P proteins. *RNA*, 2002, Vol 8, n° 3, p. 296-306.

Hall, T. A., and Brown, J. W. Interactions between RNase P protein subunits in archaea. *Archaea*, 2004, Vol 1, n° 4, p. 247-254.

Hamma, T., and Ferre-D'Amare, A. R. Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 A resolution. *Structure*, 2004, Vol 12, n° 5, p. 893-903.

Hansen, A., Pfeiffer, T., Zuleeg, T., Limmer, S., Ciesiolka, J., Feltens, R., and Hartmann, R. K. Exploring the minimal substrate requirements for trans-cleavage by RNase P holoenzymes from Escherichia coli and Bacillus subtilis. *Mol Microbiol*, 2001, Vol 41, n° 1, p. 131-143.

Harauz, G., and van Heel, M. Exact filters for general geometry 3-dimensional reconstruction. *Optik*, 1986, Vol 73, n° 146-156.

Harris, J. K., Haas, E. S., Williams, D., Frank, D. N., and Brown, J. W. New insight into RNase P RNA structure from comparative analysis of the archaeal RNA. *RNA*, 2001, Vol 7, n° 2, p. 220-232.

Harris, M. E., Nolan, J. M., Malhotra, A., Brown, J. W., Harvey, S. C., and Pace, N. R. Use of photoaffinity crosslinking and molecular modeling to analyze the global architecture of ribonuclease P RNA. *EMBO J*, 1994, Vol 13, n° 17, p. 3953-3963.

Hartmann, E., and Hartmann, R. K. The enigma of ribonuclease P evolution. *Trends Genet*, 2003, Vol 19, n° 10, p. 561-569.

Hartmann, R. K., Heinrich, J., Schlegl, J., and Schuster, H. Precursor of C4 antisense RNA of bacteriophages P1 and P7 is a substrate for RNase P of Escherichia coli. *Proc Natl Acad Sci U S A*, 1995, Vol 92, n° 13, p. 5822-5826.

Hashimoto, C., and Steitz, J. A. Sequential association of nucleolar 7-2 RNA with two different autoantigens. *J Biol Chem*, 1983, Vol 258, n° 3, p. 1379-1382.

Henry, Y., Wood, H., Morrissey, J. P., Petfalski, E., Kearsey, S., and Tollervey, D. The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. *EMBO J*, 1994, Vol 13, n° 10, p. 2452-2463.

Hermanns, P., Bertuch, A. A., Bertin, T. K., Dawson, B., Schmitt, M. E., Shaw, C., Zabel, B., and Lee, B. Consequences of mutations in the non-coding RMRP RNA in cartilage-hair hypoplasia. *Hum Mol Genet*, 2005, Vol 14, n° 23, p. 3723-3740.

Hipp, K., Galani, K., Batisse, C., Prinz, S., and Bottcher, B. Modular architecture of eukaryotic RNase P and RNase MRP revealed by electron microscopy. *Nucleic Acids Res*, 2012, Vol 40, n° 7, p. 3275-3288.

Hirose, Y., Nakashima, E., Ohashi, H., Mochizuki, H., Bando, Y., Ogata, T., Adachi, M., Toba, E., Nishimura, G., and Ikegawa, S. Identification of novel RMRP mutations and specific founder haplotypes in Japanese patients with cartilage-hair hypoplasia. *J Hum Genet*, 2006, Vol 51, n° 8, p. 706-710.

Holzmann, J., Frank, P., Loffler, E., Bennett, K. L., Gerner, C., and Rossmanith, W. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell*, 2008, Vol 135, n° 3, p. 462-474.
Holzmann, J., and Rossmanith, W. tRNA recognition, processing, and disease: hypotheses around an unorthodox type of RNase P in human mitochondria. *Mitochondrion*, 2009, Vol 9, n° 4, p. 284-288.
Honda, T., Kakuta, Y., Kimura, K., Saho, J., and Kimura, M. Structure of an archaeal homolog of the human protein complex Rpp21-Rpp29 that is a key core component for the assembly of active ribonuclease P. *J Mol Biol*, 2008, Vol 384, n° 3, p. 652-662.
Honda, T., Hara, T., Nan, J., Zhang, X., and Kimura, M. Archaeal homologs of human RNase P protein pairs Pop5 with Rpp30 and Rpp21 with Rpp29 work on distinct functional domains of the RNA subunit. *Biosci Biotechnol Biochem*, 2010, Vol 74, n° 2, p. 266-273.
Hopper, A. K., and Phizicky, E. M. tRNA transfers to the limelight. *Genes Dev*, 2003, Vol 17, n° 2, p. 162-180.
Houser-Scott, F., Xiao, S., Millikin, C. E., Zengel, J. M., Lindahl, L., and Engelke, D. R. Interactions among the protein and RNA subunits of Saccharomyces cerevisiae nuclear RNase P. *Proc Natl Acad Sci U S A*, 2002, Vol 99, n° 5, p. 2684-2689.
Hsieh, J., Walker, S. C., Fierke, C. A., and Engelke, D. R. Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release. *RNA*, 2009, Vol 15, n° 2, p. 224-234.
Hsieh, J., Koutmou, K. S., Rueda, D., Koutmos, M., Walter, N. G., and Fierke, C. A. A divalent cation stabilizes the active conformation of the B. subtilis RNase P x pre-tRNA complex: a role for an inner-sphere metal ion in RNase P. *J Mol Biol*, 2010, Vol 400, n° 1, p. 38-51.
Intine, R. V., Sakulich, A. L., Koduru, S. B., Huang, Y., Pierstorff, E., Goodier, J. L., Phan, L., and Maraia, R. J. Control of transfer RNA maturation by phosphorylation of the human La antigen on serine 366. *Mol Cell*, 2000, Vol 6, n° 2, p. 339-348.
Jaag, H. M., Lu, Q., Schmitt, M. E., and Nagy, P. D. Role of RNase MRP in viral RNA degradation and RNA recombination. *J Virol*, 2011, Vol 85, n° 1, p. 243-253.
Jacobson, M. R., Cao, L. G., Taneja, K., Singer, R. H., Wang, Y. L., and Pederson, T. Nuclear domains of the RNA subunit of RNase P. *J Cell Sci*, 1997, Vol 110 (Pt 7), n° 829-837.
James, B. D., Olsen, G. J., Liu, J. S., and Pace, N. R. The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. *Cell*, 1988, Vol 52, n° 1, p. 19-26.
Janke, C., Magiera, M. M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., and Knop, M. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast*, 2004, Vol 21, n° 11, p. 947-962.
Jarrous, N., Eder, P. S., Guerrier-Takada, C., Hoog, C., and Altman, S. Autoantigenic properties of some protein subunits of catalytically active complexes of human ribonuclease P. *RNA*, 1998, Vol 4, n° 4, p. 407-417.
Jarrous, N., Eder, P. S., Wesolowski, D., and Altman, S. Rpp14 and Rpp29, two protein subunits of human ribonuclease P. *RNA*, 1999a, Vol 5, n° 2, p. 153-157.
Jarrous, N., Wolenski, J. S., Wesolowski, D., Lee, C., and Altman, S. Localization in the nucleolus and coiled bodies of protein subunits of the ribonucleoprotein ribonuclease P. *J Cell Biol*, 1999b, Vol 146, n° 3, p. 559-572.
Jarrous, N., and Altman, S. Human ribonuclease P. *Methods Enzymol*, 2001a, Vol 342, n° 93-100.

Jarrous, N., Reiner, R., Wesolowski, D., Mann, H., Guerrier-Takada, C., and Altman, S. Function and subnuclear distribution of Rpp21, a protein subunit of the human ribonucleoprotein ribonuclease P. *RNA*, 2001b, Vol 7, n° 8, p. 1153-1164.

Jarrous, N. Human ribonuclease P: subunits, function, and intranuclear localization. *RNA*, 2002, Vol 8, n° 1, p. 1-7.

Jiang, T., and Altman, S. Protein-protein interactions with subunits of human nuclear RNase P. *Proc Natl Acad Sci U S A*, 2001a, Vol 98, n° 3, p. 920-925.

Jiang, T., Guerrier-Takada, C., and Altman, S. Protein-RNA interactions in the subunits of human nuclear RNase P. *RNA*, 2001b, Vol 7, n° 7, p. 937-941.

Jovanovic, M., Sanchez, R., Altman, S., and Gopalan, V. Elucidation of structure-function relationships in the protein subunit of bacterial RNase P using a genetic complementation approach. *Nucleic Acids Res*, 2002, Vol 30, n° 23, p. 5065-5073.

Jung, Y. H., and Lee, Y. RNases in ColE1 DNA metabolism. *Mol Biol Rep*, 1995, Vol 22, n° 2-3, p. 195-200.

Kakuta, Y., Ishimatsu, I., Numata, T., Kimura, K., Yao, M., Tanaka, I., and Kimura, M. Crystal structure of a ribonuclease P protein Ph1601p from Pyrococcus horikoshii OT3: an archaeal homologue of human nuclear ribonuclease P protein Rpp21. *Biochemistry*, 2005, Vol 44, n° 36, p. 12086-12093.

Karwan, R., Bennett, J. L., and Clayton, D. A. Nuclear RNase MRP processes RNA at multiple discrete sites: interaction with an upstream G box is required for subsequent downstream cleavages. *Genes Dev*, 1991, Vol 5, n° 7, p. 1264-1276.

Kassenbrock, C. K., Gao, G. J., Groom, K. R., Sulo, P., Douglas, M. G., and Martin, N. C. 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*, 1995, Vol 15, n° 9, p. 4763-4770.

Kawano, S., Nakashima, T., Kakuta, Y., Tanaka, I., and Kimura, M. Crystal structure of protein Ph1481p in complex with protein Ph1877p of archaeal RNase P from Pyrococcus horikoshii OT3: implication of dimer formation of the holoenzyme. *J Mol Biol*, 2006, Vol 357, n° 2, p. 583-591.

Kazantsev, A. V., Krivenko, A. A., Harrington, D. J., Carter, R. J., Holbrook, S. R., Adams, P. D., and Pace, N. R. High-resolution structure of RNase P protein from Thermotoga maritima. *Proc Natl Acad Sci U S A*, 2003, Vol 100, n° 13, p. 7497-7502.

Kazantsev, A. V., Krivenko, A. A., Harrington, D. J., Holbrook, S. R., Adams, P. D., and Pace, N. R. Crystal structure of a bacterial ribonuclease P RNA. *Proc Natl Acad Sci U S A*, 2005, Vol 102, n° 38, p. 13392-13397.

Kazantsev, A. V., and Pace, N. R. Bacterial RNase P: a new view of an ancient enzyme. *Nat Rev Microbiol*, 2006, Vol 4, n° 10, p. 729-740.

Kazantsev, A. V., Krivenko, A. A., and Pace, N. R. Mapping metal-binding sites in the catalytic domain of bacterial RNase P RNA. *RNA* 2009, Vol 15, n° 2, p. 266-276.

Khanova, E., Esakova, O., Perederina, A., Berezin, I., and Krasilnikov, A. S. Structural organizations of yeast RNase P and RNase MRP holoenzymes as revealed by UV-crosslinking studies of RNA-protein interactions. *RNA* 2012, Vol 18, n° 4, p. 720-728.

Kifusa, M., Fukuhara, H., Hayashi, T., and Kimura, M. Protein-protein interactions in the subunits of ribonuclease P in the hyperthermophilic archaeon Pyrococcus horikoshii OT3. *Biosci Biotechnol Biochem*, 2005, Vol 69, n° 6, p. 1209-1212.

Kikovska, E., Svard, S. G., and Kirsebom, L. A. Eukaryotic RNase P RNA mediates cleavage in the absence of protein. *Proc Natl Acad Sci U S A*, 2007, Vol 104, n° 7, p. 2062-2067.

Kim, J. J., Kilani, A. F., Zhan, X., Altman, S., and Liu, F. The protein cofactor allows the sequence of an RNase P ribozyme to diversify by maintaining the catalytically active structure of the enzyme. *RNA*, 1997, Vol 3, n° 6, p. 613-623.

Kirsebom, L. A., Baer, M. F., and Altman, S. Differential effects of mutations in the protein and RNA moieties of RNase P on the efficiency of suppression by various tRNA suppressors. *J Mol Biol*, 1988, Vol 204, n° 4, p. 879-888.
Kirsebom, L. A., and Svard, S. G. Base pairing between Escherichia coli RNase P RNA and its substrate. *EMBO J*, 1994, Vol 13, n° 20, p. 4870-4876.
Kirsebom, L. A., and Vioque, A. RNase P from bacteria. Substrate recognition and function of the protein subunit. *Mol Biol Rep*, 1995, Vol 22, n° 2-3, p. 99-109.
Kirsebom, L. A. RNase P RNA mediated cleavage: substrate recognition and catalysis. *Biochimie*, 2007, Vol 89, n° 10, p. 1183-1194.
Kiss, T., and Filipowicz, W. Evidence against a mitochondrial location of the 7-2/MRP RNA in mammalian cells. *Cell*, 1992a, Vol 70, n° 1, p. 11-16.
Kiss, T., Marshallsay, C., and Filipowicz, W. 7-2/MRP RNAs in plant and mammalian cells: association with higher order structures in the nucleolus. *EMBO J*, 1992b, Vol 11, n° 10, p. 3737-3746.
Klein, D. J., Schmeing, T. M., Moore, P. B., and Steitz, T. A. The kink-turn: a new RNA secondary structure motif. *EMBO J*, 2001, Vol 20, n° 15, p. 4214-4221.
Ko, J. H., and Altman, S. OLE RNA, an RNA motif that is highly conserved in several extremophilic bacteria, is a substrate for and can be regulated by RNase P RNA. *Proc Natl Acad Sci U S A*, 2007, Vol 104, n° 19, p. 7815-7820.
Kole, R., and Altman, S. Reconstitution of RNase P activity from inactive RNA and protein. *Proc Natl Acad Sci U S A*, 1979, Vol 76, n° 8, p. 3795-3799.
Kole, R., Baer, M. F., Stark, B. C., and Altman, S. E. coli RNAase P has a required RNA component. *Cell*, 1980, Vol 19, n° 4, p. 881-887.
Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K., and Inokuchi, H. A tRNA-like structure is present in 10Sa RNA, a small stable RNA from Escherichia coli. *Proc Natl Acad Sci U S A*, 1994, Vol 91, n° 20, p. 9223-9227.
Kosaka, S., Hada, K., Nakashima, T., and Kimura, M. Structural changes in ribonuclease P RNA in the hyperthermophilic archaeon Pyrococcus horikoshii OT3 induced on interaction with proteins. *Biosci Biotechnol Biochem*, 2010, Vol 74, n° 2, p. 394-396.
Koutmou, K. S., Zahler, N. H., Kurz, J. C., Campbell, F. E., Harris, M. E., and Fierke, C. A. Protein-precursor tRNA contact leads to sequence-specific recognition of 5' leaders by bacterial ribonuclease P. *J Mol Biol*, 2010, 396, n° 1, p. 195-208.
Kouzuma, Y., Mizoguchi, M., Takagi, H., Fukuhara, H., Tsukamoto, M., Numata, T., and Kimura, M. Reconstitution of archaeal ribonuclease P from RNA and four protein components. *Biochem Biophys Res Commun*, 2003, Vol 306, n° 3, p. 666-673.
Krasilnikov, A. S., and Mondragon, A. On the occurrence of the T-loop RNA folding motif in large RNA molecules. *RNA*, 2003a, Vol 9, n° 6, p. 640-643.
Krasilnikov, A. S., Yang, X., Pan, T., and Mondragon, A. Crystal structure of the specificity domain of ribonuclease P. *Nature*, 2003b, Vol 421, n° 6924, p. 760-764.
Krasilnikov, A. S., Xiao, Y., Pan, T., and Mondragon, A. Basis for structural diversity in homologous RNAs. *Science*, 2004, Vol 306, n° 5693, p. 104-107.
Krehan, M., Heubeck, C., Menzel, N., Seibel, P., and Schon, A. RNase MRP RNA and RNase P activity in plants are associated with a Pop1p containing complex. *Nucleic Acids Res*, 2012, Vol.
Kufel, J., Allmang, C., Verdone, L., Beggs, J. D., and Tollervey, D. Lsm proteins are required for normal processing of pre-tRNAs and their efficient association with La-homologous protein Lhp1p. *Mol Cell Biol*, 2002, Vol 22, n° 14, p. 5248-5256.
Kurz, J. C., Niranjanakumari, S., and Fierke, C. A. Protein component of Bacillus subtilis RNase P specifically enhances the affinity for precursor-tRNAAsp. *Biochemistry*, 1998, Vol 37, n° 8, p. 2393-2400.

Kurz, J. C., and Fierke, C. A. The affinity of magnesium binding sites in the Bacillus subtilis RNase P x pre-tRNA complex is enhanced by the protein subunit. *Biochemistry*, 2002, Vol 41, n° 30, p. 9545-9558.

LaGrandeur, T. E., Huttenhofer, A., Noller, H. F., and Pace, N. R. Phylogenetic comparative chemical footprint analysis of the interaction between ribonuclease P RNA and tRNA. *EMBO J*, 1994, Vol 13, n° 17, p. 3945-3952.

Lai, L. B., Chan, P. P., Cozen, A. E., Bernick, D. L., Brown, J. W., Gopalan, V., and Lowe, T. M. Discovery of a minimal form of RNase P in Pyrobaculum. *Proc Natl Acad Sci U S A*, 2010, Vol 107, n° 52, p. 22493-22498.

Larimer, F. W., and Stevens, A. Disruption of the gene XRN1, coding for a 5'----3' exoribonuclease, restricts yeast cell growth. *Gene*, 1990, Vol 95, n° 1, p. 85-90.

Lee, D. Y., and Clayton, D. A. Initiation of mitochondrial DNA replication by transcription and R-loop processing. *J Biol Chem*, 1998, Vol 273, n° 46, p. 30614-30621.

Lee, J. Y., Evans, C. F., and Engelke, D. R. Expression of RNase P RNA in Saccharomyces cerevisiae is controlled by an unusual RNA polymerase III promoter. *Proc Natl Acad Sci U S A*, 1991a, Vol 88, n° 16, p. 6986-6990.

Lee, J. Y., Rohlman, C. E., Molony, L. A., and Engelke, D. R. Characterization of RPR1, an essential gene encoding the RNA component of Saccharomyces cerevisiae nuclear RNase P. *Mol Cell Biol*, 1991b, Vol 11, n° 2, p. 721-730.

Lee, Y., Kindelberger, D. W., Lee, J. Y., McClennen, S., Chamberlain, J., and Engelke, D. R. Nuclear pre-tRNA terminal structure and RNase P recognition. *RNA*, 1997, Vol 3, n° 2, p. 175-185.

Leschziner, A. (2010). The orthogonal tilt reconstruction method, In Methods Enzymol, G. Jensen, ed., p. 237-262.

Li, D., Willkomm, D. K., and Hartmann, R. K. Minor changes largely restore catalytic activity of archaeal RNase P RNA from Methanothermobacter thermoautotrophicus. *Nucleic Acids Res*, 2009, Vol 37, n° 1, p. 231-242.

Li, X., Zaman, S., Langdon, Y., Zengel, J. M., and Lindahl, L. Identification of a functional core in the RNA component of RNase MRP of budding yeasts. *Nucleic Acids Res*, 2004, Vol 32, n° 12, p. 3703-3711.

Li, Y., and Altman, S. Partial reconstitution of human RNase P in HeLa cells between its RNA subunit with an affinity tag and the intact protein components. *Nucleic Acids Res*, 2002, Vol 30, n° 17, p. 3706-3711.

Li, Z., Pandit, S., and Deutscher, M. P. 3' exoribonucleolytic trimming is a common feature of the maturation of small, stable RNAs in Escherichia coli. *Proc Natl Acad Sci U S A*, 1998, Vol 95, n° 6, p. 2856-2861.

Lindahl, L., Fretz, S., Epps, N., and Zengel, J. M. Functional equivalence of hairpins in the RNA subunits of RNase MRP and RNase P in Saccharomyces cerevisiae. *RNA* 2000, Vol 6, n° 5, p. 653-658.

Lindahl, L., Bommankanti, A., Li, X., Hayden, L., Jones, A., Khan, M., Oni, T., and Zengel, J. M. RNase MRP is required for entry of 35S precursor rRNA into the canonical processing pathway. *RNA*, 2009, Vol 15, n° 7, p. 1407-1416.

Liu, F., and Altman, S. Differential evolution of substrates for an RNA enzyme in the presence and absence of its protein cofactor. *Cell*, 1994, Vol 77, n° 7, p. 1093-1100.

Longtine, M. S., McKenzie, A., 3rd, Demarini, D. J., Shah, N. G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J. R. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast*, 1998, Vol 14, n° 10, p. 953-961.

Loria, A., and Pan, T. Domain structure of the ribozyme from eubacterial ribonuclease P. *RNA*, 1996, Vol 2, n° 6, p. 551-563.

Loria, A., and Pan, T. Recognition of the T stem-loop of a pre-tRNA substrate by the ribozyme from Bacillus subtilis ribonuclease P. *Biochemistry*, 1997, Vol 36, n° 21, p. 6317-6325.
Loria, A., Niranjanakumari, S., Fierke, C. A., and Pan, T. Recognition of a pre-tRNA substrate by the Bacillus subtilis RNase P holoenzyme. *Biochemistry*, 1998, Vol 37, n° 44, p. 15466-15473.
Loria, A., and Pan, T. The cleavage step of ribonuclease P catalysis is determined by ribozyme-substrate interactions both distal and proximal to the cleavage site. *Biochemistry*, 1999, Vol 38, n° 27, p. 8612-8620.
Loria, A., and Pan, T. The 3' substrate determinants for the catalytic efficiency of the Bacillus subtilis RNase P holoenzyme suggest autolytic processing of the RNase P RNA in vivo. *RNA*, 2000, Vol 6, n° 10, p. 1413-1422.
Lu, Q., Wierzbicki, S., Krasilnikov, A. S., and Schmitt, M. E. Comparison of mitochondrial and nucleolar RNase MRP reveals identical RNA components with distinct enzymatic activities and protein components. *RNA*, 2010, Vol 16, n° 3, p. 529-537.
Ludtke, S. J., Baldwin, P. R., and Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol*, 1999, Vol 128, n° 1, p. 82-97.
Lumelsky, N., and Altman, S. Selection and characterization of randomly produced mutants in the gene coding for M1 RNA. *J Mol Biol*, 1988, Vol 202, n° 3, p. 443-454.
Lygerou, Z., Mitchell, P., Petfalski, E., Seraphin, B., and Tollervey, D. The POP1 gene encodes a protein component common to the RNase MRP and RNase P ribonucleoproteins. *Genes Dev*, 1994, Vol 8, n° 12, p. 1423-1433.
Lygerou, Z., Allmang, C., Tollervey, D., and Seraphin, B. Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP in vitro. *Science*, 1996a, Vol 272, n° 5259, p. 268-270.
Lygerou, Z., Pluk, H., van Venrooij, W. J., and Seraphin, B. hPop1: an autoantigenic protein subunit shared by the human RNase P and RNase MRP ribonucleoproteins. *EMBO J*, 1996b, Vol 15, n° 21, p. 5936-5948.
Maida, Y., Yasukawa, M., Furuuchi, M., Lassmann, T., Possemato, R., Okamoto, N., Kasim, V., Hayashizaki, Y., Hahn, W. C., and Masutomi, K. An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. *Nature*, 2009, Vol 461, n° 7261, p. 230-235.
Mann, H., Ben-Asouli, Y., Schein, A., Moussa, S., and Jarrous, N. Eukaryotic RNase P: role of RNA and protein subunits of a primordial catalytic ribonucleoprotein in RNA-based catalysis. *Mol Cell*, 2003, Vol 12, n° 4, p. 925-935.
Maraia, R. J., and Intine, R. V. Recognition of nascent RNA by the human La antigen: conserved and divergent features of structure and function. *Mol Cell Biol*, 2001, Vol 21, n° 2, p. 367-379.
Marquez, S. M., Harris, J. K., Kelley, S. T., Brown, J. W., Dawson, S. C., Roberts, E. C., and Pace, N. R. Structural implications of novel diversity in eucaryal RNase P RNA. *RNA*, 2005, Vol 11, n° 5, p. 739-751.
Marquez, S. M., Chen, J. L., Evans, D., and Pace, N. R. Structure and function of eukaryotic Ribonuclease P RNA. *Mol Cell*, 2006, Vol 24, n° 3, p. 445-456.
Martin, A. N., and Li, Y. RNase MRP RNA and human genetic diseases. *Cell Res*, 2007, Vol 17, n° 3, p. 219-226.
Marvin, M. C., Walker, S. C., Fierke, C. A., and Engelke, D. R. Binding and cleavage of unstructured RNA by nuclear RNase P. *RNA*, 2011, Vol 17, n° 8, p. 1429-1440.
Marvin, M. C. and Engelke, D. R. Broadening the Mission of an RNA Enzyme. *J Cell Biochem.,* 2009, vol 108, n°6, p.1244-51.

Massire, C., Jaeger, L., and Westhof, E. Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. *J Mol Biol*, 1998, Vol 279, n° 4, p. 773-793.

Mobley, E. M., and Pan, T. Design and isolation of ribozyme-substrate pairs using RNase P-based ribozymes containing altered substrate binding sites. *Nucleic Acids Res*, 1999, Vol 27, n° 21, p. 4298-4304.

Mondragon, A. (2010). Structural Studies of Ribonuclease P. *Protein Reviews*, In Ribonuclease P, F. Liu, and S. Altman, eds., p. 63-78.

Morales, M. J., Wise, C. A., Hollingsworth, M. J., and Martin, N. C. Characterization of yeast mitochondrial RNase P: an intact RNA subunit is not essential for activity in vitro. *Nucleic Acids Res*, 1989, Vol 17, n° 17, p. 6865-6881.

Morales, M. J., Dang, Y. L., Lou, Y. C., Sulo, P., and Martin, N. C. A 105-kDa protein is required for yeast mitochondrial RNase P activity. *Proc Natl Acad Sci U S A*, 1992, Vol 89, n° 20, p. 9875-9879.

Nakashima, E., Tran, J. R., Welting, T. J., Pruijn, G. J., Hirose, Y., Nishimura, G., Ohashi, H., Schurman, S. H., Cheng, J., Candotti, F., *et al.* Cartilage hair hypoplasia mutations that lead to RMRP promoter inefficiency or RNA transcript instability. *Am J Med Genet A*, 2007, Vol 143A, n° 22, p. 2675-2681.

Niranjanakumari, S., Stams, T., Crary, S. M., Christianson, D. W., and Fierke, C. A. Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA. *Proc Natl Acad Sci U S A*, 1998, Vol 95, n° 26, p. 15212-15217.

Niranjanakumari, S., Day-Storms, J. J., Ahmed, M., Hsieh, J., Zahler, N. H., Venters, R. A., and Fierke, C. A. Probing the architecture of the B. subtilis RNase P holoenzyme active site by cross-linking and affinity cleavage. *RNA*, 2007, Vol 13, n° 4, p. 521-535.

Nogales, E., and Grigorieff, N. Molecular Machines: putting the pieces together. *J Cell Biol*, 2001, Vol 152, n° 1, p. F1-10.

Numata, T., Ishimatsu, I., Kakuta, Y., Tanaka, I., and Kimura, M. Crystal structure of archaeal ribonuclease P protein Ph1771p from Pyrococcus horikoshii OT3: an archaeal homolog of eukaryotic ribonuclease P protein Rpp29. *RNA*, 2004, Vol 10, n° 9, p. 1423-1432.

Odell, L., Huang, V., Jakacka, M., and Pan, T. Interaction of structural modules in substrate binding by the ribozyme from Bacillus subtilis RNase P. *Nucleic Acids Res*, 1998, Vol 26, n° 16, p. 3717-3723.

Oh, B. K., and Pace, N. R. Interaction of the 3'-end of tRNA with ribonuclease P RNA. *Nucleic Acids Res*, 1994, Vol 22, n° 20, p. 4087-4094.

Pace, N. R., and Smith, D. Ribonuclease P: function and variation. *J Biol Chem*, 1990, Vol 265, n° 7, p. 3587-3590.

Pagan-Ramos, E., Lee, Y., and Engelke, D. R. A conserved RNA motif involved in divalent cation utilization by nuclear RNase P. *RNA,* 1996, Vol 2, n° 11, p. 1100-1109.

Paluh, J. L., and Clayton, D. A. Schizosaccharomyces pombe RNase MRP RNA is homologous to metazoan RNase MRP RNAs and may provide clues to interrelationships between RNase MRP and RNase P. *Yeast*, 1995, Vol 11, n° 13, p. 1249-1264.

Pan, T. Higher order folding and domain analysis of the ribozyme from Bacillus subtilis ribonuclease P. *Biochemistry*, 1995, Vol 34, n° 3, p. 902-909.

Pannone, B. K., Kim, S. D., Noe, D. A., and Wolin, S. L. Multiple functional interactions between components of the Lsm2-Lsm8 complex, U6 snRNA, and the yeast La protein. *Genetics*, 2001, Vol 158, n° 1, p. 187-196.

Pannucci, J. A., Haas, E. S., Hall, T. A., Harris, J. K., and Brown, J. W. RNase P RNAs from some Archaea are catalytically active. *Proc Natl Acad Sci U S A*, 1999, Vol 96, n° 14, p. 7803-7808.

Parker, R., and Sheth, U. P bodies and the control of mRNA translation and degradation. *Mol Cell*, 2007, Vol 25, n° 5, p. 635-646.

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

Perederina, A., Esakova, O., Koc, H., Schmitt, M. E., and Krasilnikov, A. S. Specific binding of a Pop6/Pop7 heterodimer to the P3 stem of the yeast RNase MRP and RNase P RNAs. *RNA*, 2007, Vol 13, n° 10, p. 1648-1655.

Perederina, A., Esakova, O., Quan, C., Khanova, E., and Krasilnikov, A. S. Eukaryotic ribonucleases P/MRP: the crystal structure of the P3 domain. *EMBO J*, 2010a, Vol 29, n° 4, p. 761-769.

Perederina, A., and Krasilnikov, A. S. The P3 domain of eukaryotic RNases P/MRP: making a protein-rich RNA-based enzyme. *RNA Biol*, 2010b, Vol 7, n° 5, p. 534-539.

Perederina, A., Khanova, E., Quan, C., Berezin, I., Esakova, O., and Krasilnikov, A. S. Interactions of a Pop5/Rpp1 heterodimer with the catalytic domain of RNase MRP. *RNA*, 2011, Vol 17, n° 10, p. 1922-1931.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, 2004, Vol 25, n° 13, p. 1605-1612.

Pfeiffer, T., Tekos, A., Warnecke, J. M., Drainas, D., Engelke, D. R., Seraphin, B., and Hartmann, R. K. Effects of phosphorothioate modifications on precursor tRNA processing by eukaryotic RNase P enzymes. *J Mol Biol*, 2000, Vol 298, n° 4, p. 559-565.

Piccinelli, P., Rosenblad, M. A., and Samuelsson, T. Identification and analysis of ribonuclease P and MRP RNA in a broad range of eukaryotes. *Nucleic Acids Res*, 2005, Vol 33, n° 14, p. 4485-4495.

Pluk, H., van Eenennaam, H., Rutjes, S. A., Pruijn, G. J., and van Venrooij, W. J. RNA-protein interactions in the human RNase MRP ribonucleoprotein complex. *RNA*, 1999, Vol 5, n° 4, p. 512-524.

Potuschak, T., Rossmanith, W., and Karwan, R. RNase MRP and RNase P share a common substrate. *Nucleic Acids Res*, 1993, Vol 21, n° 14, p. 3239-3243.

Pulukkunat, D. K., and Gopalan, V. Studies on Methanocaldococcus jannaschii RNase P reveal insights into the roles of RNA and protein cofactors in RNase P catalysis. *Nucleic Acids Res*, 2008, Vol 36, n° 12, p. 4172-4180.

Qin, H., Sosnick, T. R., and Pan, T. Modular construction of a tertiary RNA structure: the specificity domain of the Bacillus subtilis RNase P RNA. *Biochemistry*, 2001, Vol 40, n° 37, p. 11202-11210.

Qu, X., Smith, G. J., Lee, K. T., Sosnick, T. R., Pan, T., and Scherer, N. F. Single-molecule nonequilibrium periodic Mg2+-concentration jump experiments reveal details of the early folding pathways of a large RNA. *Proc Natl Acad Sci U S A*, 2008, Vol 105, n° 18, p. 6602-6607.

Radermacher, M., Wagenknecht, T., Verschoor, A., and Frank, J. Three-dimensional structure of the large ribosomal subunit from Escherichia coli. *EMBO J*, 1987, Vol 6, n° 4, p. 1107-1114.

Reich, C., Olsen, G. J., Pace, B., and Pace, N. R. Role of the protein moiety of ribonuclease P, a ribonucleoprotein enzyme. *Science*, 1988, Vol 239, n° 4836, p. 178-181.

Reimer, G., Raska, I., Scheer, U., and Tan, E. M. Immunolocalization of 7-2-ribonucleoprotein in the granular component of the nucleolus. *Exp Cell Res*, 1988, Vol 176, n° 1, p. 117-128.

Reiner, R., Ben-Asouli, Y., Krilovetzky, I., and Jarrous, N. A role for the catalytic ribonucleoprotein RNase P in RNA polymerase III transcription. *Genes Dev*, 2006, Vol 20, n° 12, p. 1621-1635.

Reiner, R., Krasnov-Yoeli, N., Dehtiar, Y., and Jarrous, N. Function and assembly of a chromatin-associated RNase P that is required for efficient transcription by RNA polymerase I. *PLoS One*, 2008, Vol 3, n° 12, p. e4072.

Reiter, N. J., Osterman, A., Torres-Larios, A., Swinger, K. K., Pan, T., and Mondragon, A. Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*, 2010, Vol 468, n° 7325, p. 784-789.

Ridanpaa, M., van Eenennaam, H., Pelin, K., Chadwick, R., Johnson, C., Yuan, B., vanVenrooij, W., Pruijn, G., Salmela, R., Rockas, S.*, et al.* Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. *Cell*, 2001, Vol 104, n° 2, p. 195-203.

Robbens, S., Derelle, E., Ferraz, C., Wuyts, J., Moreau, H., and Van de Peer, Y. The complete chloroplast and mitochondrial DNA sequence of Ostreococcus tauri: organelle genomes of the smallest eukaryote are examples of compaction. *Mol Biol Evol*, 2007, Vol 24, n° 4, p. 956-968.

Rosenblad, M. A., Lopez, M. D., Piccinelli, P., and Samuelsson, T. Inventory and analysis of the protein subunits of the ribonucleases P and MRP provides further evidence of homology between the yeast and human enzymes. *Nucleic Acids Res*, 2006, Vol 34, n° 18, p. 5145-5156.

Rossmanith, W. Of P and Z: mitochondrial tRNA processing enzymes. *Biochim Biophys Acta*, 2012, Vol 1819, n° 9-10, p. 1017-1026.

Rozhdestvensky, T. S., Tang, T. H., Tchirkova, I. V., Brosius, J., Bachellerie, J. P., and Huttenhofer, A. Binding of L7Ae protein to the K-turn of archaeal snoRNAs: a shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea. *Nucleic Acids Res*, 2003, Vol 31, n° 3, p. 869-877.

Rubin, G. M. Three forms of the 5.8-S ribosomal RNA species in Saccharomyces cerevisiae. *Eur J Biochem*, 1974, Vol 41, n° 1, p. 197-202.

Rueda, D., Hsieh, J., Day-Storms, J. J., Fierke, C. A., and Walter, N. G. The 5' leader of precursor tRNAAsp bound to the Bacillus subtilis RNase P holoenzyme has an extended conformation. *Biochemistry*, 2005, Vol 44, n° 49, p. 16130-16139.

Salgado-Garrido, J., Bragado-Nilsson, E., Kandels-Lewis, S., and Seraphin, B. Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J*, 1999, Vol 18, n° 12, p. 3451-3462.

Salinas, K., Wierzbicki, S., Zhou, L., and Schmitt, M. E. Characterization and purification of Saccharomyces cerevisiae RNase MRP reveals a new unique protein component. *J Biol Chem*, 2005, Vol 280, n° 12, p. 11352-11360.

Samanta, M. P., Tongprasit, W., Sethi, H., Chin, C. S., and Stolc, V. Global identification of noncoding RNAs in Saccharomyces cerevisiae by modulating an essential RNA processing pathway. *Proc Natl Acad Sci U S A*, 2006, Vol 103, n° 11, p. 4192-4197.

Sambrook, J., Russell, D.W. (2001). Molecular cloning : a laboratory manual, 3rd edn (New York).

Savino, T. M., Bastos, R., Jansen, E., and Hernandez-Verdun, D. The nucleolar antigen Nop52, the human homologue of the yeast ribosomal RNA processing RRP1, is recruited at late stages of nucleologenesis. *J Cell Sci*, 1999, Vol 112 (Pt 12), n° 1889-1900.

Schedl, P., and Primakoff, P. Mutants of Escherichia coli thermosensitive for the synthesis of transfer RNA. *Proc Natl Acad Sci U S A*, 1973, Vol 70, n° 7, p. 2091-2095.
Scheres, S. H., Gao, H., Valle, M., Herman, G. T., Eggermont, P. P., Frank, J., and Carazo, J. M. Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Methods*, 2007, Vol 4, n° 1, p. 27-29.
Schiestl, R. H., and Gietz, R. D. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet*, 1989, Vol 16, n° 5-6, p. 339-346.
Schmitt, M. E., and Clayton, D. A. Yeast site-specific ribonucleoprotein endoribonuclease MRP contains an RNA component homologous to mammalian RNase MRP RNA and essential for cell viability. *Genes Dev*, 1992, Vol 6, n° 10, p. 1975-1985.
Schmitt, M. E., Bennett, J. L., Dairaghi, D. J., and Clayton, D. A. Secondary structure of RNase MRP RNA as predicted by phylogenetic comparison. *Faseb J*, 1993a, Vol 7, n° 1, p. 208-213.
Schmitt, M. E., and Clayton, D. A. Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in Saccharomyces cerevisiae. *Mol Cell Biol*, 1993b, Vol 13, n° 12, p. 7935-7941.
Schmitt, M. E., and Clayton, D. A. Characterization of a unique protein component of yeast RNase MRP: an RNA-binding protein with a zinc-cluster domain. *Genes Dev*, 1994, Vol 8, n° 21, p. 2617-2628.
Schmitz-Linneweber, C., and Small, I. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci*, 2008, Vol 13, n° 12, p. 663-670.
Seif, E. R., Forget, L., Martin, N. C., and Lang, B. F. Mitochondrial RNase P RNAs in ascomycete fungi: lineage-specific variations in RNA secondary structure. *RNA*, 2003, Vol 9, n° 9, p. 1073-1083.
Shadel, G. S., Buckenmeyer, G. A., Clayton, D. A., and Schmitt, M. E. Mutational analysis of the RNA component of Saccharomyces cerevisiae RNase MRP reveals distinct nuclear phenotypes. *Gene*, 2000, Vol 245, n° 1, p. 175-184.
Sharin, E., Schein, A., Mann, H., Ben-Asouli, Y., and Jarrous, N. RNase P: role of distinct protein cofactors in tRNA substrate recognition and RNA-based catalysis. *Nucleic Acids Res*, 2005, Vol 33, n° 16, p. 5120-5132.
Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem*, 1996, Vol 68, n° 5, p. 850-858.
Shi, H., and Moore, P. B. The crystal structure of yeast phenylalanine tRNA at 1.93 A resolution: a classic structure revisited. *RNA*, 2000, Vol 6, n° 8, p. 1091-1105.
Sidote, D. J., and Hoffman, D. W. NMR structure of an archaeal homologue of ribonuclease P protein Rpp29. *Biochemistry*, 2003, Vol 42, n° 46, p. 13541-13550.
Sidote, D. J., Heideker, J., and Hoffman, D. W. Crystal structure of archaeal ribonuclease P protein aRpp29 from Archaeoglobus fulgidus. *Biochemistry*, 2004, Vol 43, n° 44, p. 14128-14138.
Smith, D., Burgin, A. B., Haas, E. S., and Pace, N. R. Influence of metal ions on the ribonuclease P reaction. Distinguishing substrate binding from catalysis. *J Biol Chem*, 1992, Vol 267, n° 4, p. 2429-2436.
Smith, J. K., Hsieh, J., and Fierke, C. A. Importance of RNA-protein interactions in bacterial ribonuclease P structure and catalysis. *Biopolymers*, 2007, Vol 87, n° 5-6, p. 329-338.
Smith, S. D., Banerjee, N., and Sitz, T. O. Gene heterogeneity: a basis for alternative 5.8S rRNA processing. *Biochemistry*, 1984, Vol 23, n° 16, p. 3648-3652.
Spitzfaden, C., Nicholson, N., Jones, J. J., Guth, S., Lehr, R., Prescott, C. D., Hegg, L. A., and Eggleston, D. S. The structure of ribonuclease P protein from Staphylococcus aureus reveals a unique binding site for single-stranded RNA. *J Mol Biol*, 2000, Vol 295, n° 1, p. 105-115.

Srisawat, C., and Engelke, D. R. Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. *RNA*, 2001, Vol 7, n° 4, p. 632-641.

Srisawat, C., Houser-Scott, F., Bertrand, E., Xiao, S., Singer, R. H., and Engelke, D. R. An active precursor in assembly of yeast nuclear ribonuclease P. *RNA,* 2002, Vol 8, n° 10, p. 1348-1360.

Stahley, M. R., and Strobel, S. A. Structural evidence for a two-metal-ion mechanism of group I intron splicing. *Science*, 2005, Vol 309, n° 5740, p. 1587-1590.

Stams, T., Niranjanakumari, S., Fierke, C. A., and Christianson, D. W. Ribonuclease P protein structure: evolutionary origins in the translational apparatus. *Science*, 1998, Vol 280, n° 5364, p. 752-755.

Stark, B. C., Kole, R., Bowman, E. J., and Altman, S. Ribonuclease P: an enzyme with an essential RNA component. *Proc Natl Acad Sci U S A*, 1978, Vol 75, n° 8, p. 3717-3721.

Stohl, L. L., and Clayton, D. A. Saccharomyces cerevisiae contains an RNase MRP that cleaves at a conserved mitochondrial RNA sequence implicated in replication priming. *Mol Cell Biol*, 1992, Vol 12, n° 6, p. 2561-2569.

Stolc, V., and Altman, S. Rpp1, an essential protein subunit of nuclear RNase P required for processing of precursor tRNA and 35S precursor rRNA in Saccharomyces cerevisiae. *Genes Dev*, 1997, Vol 11, n° 21, p. 2926-2937.

Stolc, V., Katz, A., and Altman, S. Rpp2, an essential protein subunit of nuclear RNase P, is required for processing of precursor tRNAs and 35S precursor rRNA in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A*, 1998, Vol 95, n° 12, p. 6716-6721.

Stribinskis, V., Gao, G. J., Sulo, P., Dang, Y. L., and Martin, N. C. 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*, 1996, Vol 16, n° 7, p. 3429-3436.

Stribinskis, V., Gao, G. J., Ellis, S. R., and Martin, N. C. 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*, 2001, Vol 158, n° 2, p. 573-585.

Stribinskis, V., Heyman, H. C., Ellis, S. R., Steffen, M. C., and Martin, N. C. Rpm2p, a component of yeast mitochondrial RNase P, acts as a transcriptional activator in the nucleus. *Mol Cell Biol*, 2005, Vol 25, n° 15, p. 6546-6558.

Stribinskis, V., and Ramos, K. S. Rpm2p, a protein subunit of mitochondrial RNase P, physically and genetically interacts with cytoplasmic processing bodies. *Nucleic Acids Res*, 2007, Vol 35, n° 4, p. 1301-1311.

Sun, F. J., and Caetano-Anolles, G. The ancient history of the structure of ribonuclease P and the early origins of Archaea. *BMC Bioinformatics*, 2010, Vol 11, n° 153.

Sun, L., Campbell, F. E., Zahler, N. H., and Harris, M. E. Evidence that substrate-specific effects of C5 protein lead to uniformity in binding and catalysis by RNase P. *EMBO J*, 2006, Vol 25, n° 17, p. 3998-4007.

Sun, L., and Harris, M. E. Evidence that binding of C5 protein to P RNA enhances ribozyme catalysis by influencing active site metal ion affinity. *RNA*, 2007, Vol 13, n° 9, p. 1505-1515.

Takagi, H., Watanabe, M., Kakuta, Y., Kamachi, R., Numata, T., Tanaka, I., and Kimura, M. Crystal structure of the ribonuclease P protein Ph1877p from hyperthermophilic archaeon Pyrococcus horikoshii OT3. *Biochem Biophys Res Commun*, 2004, Vol 319, n° 3, p. 787-794.

Tallsjo, A., and Kirsebom, L. A. Product release is a rate-limiting step during cleavage by the catalytic RNA subunit of Escherichia coli RNase P. *Nucleic Acids Res*, 1993, Vol 21, n° 1, p. 51-57.

Terada, A., Honda, T., Fukuhara, H., Hada, K., and Kimura, M. Characterization of the archaeal ribonuclease P proteins from Pyrococcus horikoshii OT3. *J Biochem*, 2006, Vol 140, n° 2, p. 293-298.
Thiel, C. T., Horn, D., Zabel, B., Ekici, A. B., Salinas, K., Gebhart, E., Ruschendorf, F., Sticht, H., Spranger, J., Muller, D.*, et al.* Severely incapacitating mutations in patients with extreme short stature identify RNA-processing endoribonuclease RMRP as an essential cell growth regulator. *Am J Hum Genet*, 2005, Vol 77, n° 5, p. 795-806.
Thiel, C. T., Mortier, G., Kaitila, I., Reis, A., and Rauch, A. Type and level of RMRP functional impairment predicts phenotype in the cartilage hair hypoplasia-anauxetic dysplasia spectrum. *Am J Hum Genet*, 2007, Vol 81, n° 3, p. 519-529.
Thomas, B. C., Chamberlain, J., Engelke, D. R., and Gegenheimer, P. Evidence for an RNA-based catalytic mechanism in eukaryotic nuclear ribonuclease P. *RNA*, 2000, Vol 6, n° 4, p. 554-562.
Thompson, M., Haeusler, R. A., Good, P. D., and Engelke, D. R. Nucleolar clustering of dispersed tRNA genes. *Science*, 2003, Vol 302, n° 5649, p. 1399-1401.
Toor, N., Keating, K. S., Taylor, S. D., and Pyle, A. M. Crystal structure of a self-spliced group II intron. *Science*, 2008, Vol 320, n° 5872, p. 77-82.
Topper, J. N., and Clayton, D. A. Characterization of human MRP/Th RNA and its nuclear gene: full length MRP/Th RNA is an active endoribonuclease when assembled as an RNP. *Nucleic Acids Res*, 1990, Vol 18, n° 4, p. 793-799.
Topper, J. N., Bennett, J. L., and Clayton, D. A. A role for RNAase MRP in mitochondrial RNA processing. *Cell*, 1992, Vol 70, n° 1, p. 16-20.
Toro, I., Thore, S., Mayer, C., Basquin, J., Seraphin, B., and Suck, D. RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. *EMBO J*, 2001, Vol 20, n° 9, p. 2293-2303.
Torres-Larios, A., Swinger, K. K., Krasilnikov, A. S., Pan, T., and Mondragon, A. Crystal structure of the RNA component of bacterial ribonuclease P. *Nature*, 2005, Vol 437, n° 7058, p. 584-587.
Torres-Larios, A., Swinger, K. K., Pan, T., and Mondragon, A. Structure of ribonuclease P--a universal ribozyme. *Curr Opin Struct Biol*, 2006, Vol 16, n° 3, p. 327-335.
Towbin, H., Staehelin, T., and Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A*, 1979, Vol 76, n° 9, p. 4350-4354.
Tranguch, A. J., Kindelberger, D. W., Rohlman, C. E., Lee, J. Y., and Engelke, D. R. Structure-sensitive RNA footprinting of yeast nuclear ribonuclease P. *Biochemistry*, 1994, Vol 33, n° 7, p. 1778-1787.
Tsai, H. Y., Pulukkunat, D. K., Woznick, W. K., and Gopalan, V. Functional reconstitution and characterization of Pyrococcus furiosus RNase P. *Proc Natl Acad Sci U S A*, 2006, Vol 103, n° 44, p. 16147-16152.
Turmel, M., Lemieux, C., Burger, G., Lang, B. F., Otis, C., Plante, I., and Gray, M. W. The complete mitochondrial DNA sequences of Nephroselmis olivacea and Pedinomonas minor. Two radically different evolutionary patterns within green algae. *Plant Cell*, 1999, Vol 11, n° 9, p. 1717-1730.
Ulbrich, C., Diepholz, M., Bassler, J., Kressler, D., Pertschy, B., Galani, K., Bottcher, B., and Hurt, E. Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. *Cell*, 2009, Vol 138, n° 5, p. 911-922.
Valentine, R. C. Contrast enhancement in the electron microscopy of viruses. *Adv Virus Res*, 1961, Vol 8, n° 287-318.

van Eenennaam, H., Pruijn, G. J., and van Venrooij, W. J. hPop4: a new protein subunit of the human RNase MRP and RNase P ribonucleoprotein complexes. *Nucleic Acids Res*, 1999, Vol 27, n° 12, p. 2465-2472.

van Eenennaam, H., Lugtenberg, D., Vogelzangs, J. H., van Venrooij, W. J., and Pruijn, G. J. hPop5, a protein subunit of the human RNase MRP and RNase P endoribonucleases. *J Biol Chem*, 2001, Vol 276, n° 34, p. 31635-31641.

Van Eenennaam, H., Vogelzangs, J. H., Lugtenberg, D., Van Den Hoogen, F. H., Van Venrooij, W. J., and Pruijn, G. J. Identity of the RNase MRP- and RNase P-associated Th/To autoantigen. *Arthritis Rheum*, 2002, Vol 46, n° 12, p. 3266-3272.

van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R., and Schatz, M. A new generation of the IMAGIC image processing system. *J Struct Biol*, 1996, Vol 116, n° 1, p. 17-24.

Vincent, R. D., Hofmann, T. J., and Zassenhaus, H. P. Sequence and expression of NUC1, the gene encoding the mitochondrial nuclease in Saccharomyces cerevisiae. *Nucleic Acids Res*, 1988, Vol 16, n° 8, p. 3297-3312.

Walker, S. C., Aspinall, T. V., Gordon, J. M., and Avis, J. M. Probing the structure of Saccharomyces cerevisiae RNase MRP. *Biochem Soc Trans*, 2005, Vol 33, n° Pt 3, p. 479-481.

Walker, S. C., and Engelke, D. R. Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit Rev Biochem Mol Biol*, 2006, Vol 41, n° 2, p. 77-102.

Walker, S. C., Marvin, M. C., and Engelke, D. R. (2010). Eukaryote RNase P and RNase MRP. *Protein Reviews*, In Ribonuclease P, F. Liu, and S. Altman, eds., p. 173-202.

Wang, L., Haeusler, R. A., Good, P. D., Thompson, M., Nagar, S., and Engelke, D. R. Silencing near tRNA genes requires nucleolar localization. *J Biol Chem*, 2005, Vol 280, n° 10, p. 8637-8639.

Wang, M. J., Davis, N. W., and Gegenheimer, P. Novel mechanisms for maturation of chloroplast transfer RNA precursors. *EMBO J*, 1988, Vol 7, n° 6, p. 1567-1574.

Warnecke, J. M., Held, R., Busch, S., and Hartmann, R. K. Role of metal ions in the hydrolysis reaction catalyzed by RNase P RNA from Bacillus subtilis. *J Mol Biol*, 1999, Vol 290, n° 2, p. 433-445.

Waugh, D. S., and Pace, N. R. Complementation of an RNase P RNA (rnpB) gene deletion in Escherichia coli by homologous genes from distantly related eubacteria. *J Bacteriol*, 1990, Vol 172, n° 11, p. 6316-6322.

Wegscheid, B., Condon, C., and Hartmann, R. K. Type A and B RNase P RNAs are interchangeable in vivo despite substantial biophysical differences. *EMBO Rep*, 2006, Vol 7, n° 4, p. 411-417.

Welting, T. J., van Venrooij, W. J., and Pruijn, G. J. Mutual interactions between subunits of the human RNase MRP ribonucleoprotein complex. *Nucleic Acids Res*, 2004, Vol 32, n° 7, p. 2138-2146.

Welting, T. J., Kikkert, B. J., van Venrooij, W. J., and Pruijn, G. J. Differential association of protein subunits with the human RNase MRP and RNase P complexes. *RNA*, 2006, Vol 12, n° 7, p. 1373-1382.

Welting, T. J., Peters, F. M., Hensen, S. M., van Doorn, N. L., Kikkert, B. J., Raats, J. M., van Venrooij, W. J., and Pruijn, G. J. Heterodimerization regulates RNase MRP/RNase P association, localization, and expression of Rpp20 and Rpp25. *RNA*, 2007, Vol 13, n° 1, p. 65-75.

Welting, T. J., Mattijssen, S., Peters, F. M., van Doorn, N. L., Dekkers, L., van Venrooij, W. J., Heus, H. A., Bonafe, L., and Pruijn, G. J. Cartilage-hair hypoplasia-associated mutations in the RNase MRP P3 domain affect RNA folding and ribonucleoprotein assembly. *Biochim Biophys Acta*, 2008, Vol 1783, n° 3, p. 455-466.

Westhof, E., Wesolowski, D., and Altman, S. Mapping in three dimensions of regions in a catalytic RNA protected from attack by an Fe(II)-EDTA reagent. *J Mol Biol*, 1996, Vol 258, n° 4, p. 600-613.
Westhof, E., and Massire, C. Structural biology. Evolution of RNA architecture. *Science*, 2004, Vol 306, n° 5693, p. 62-63.
Wilson, R. C., Bohlen, C. J., Foster, M. P., and Bell, C. E. Structure of Pfu Pop5, an archaeal RNase P protein. *Proc Natl Acad Sci U S A*, 2006, Vol 103, n° 4, p. 873-878.
Wilusz, J. E., Freier, S. M., and Spector, D. L. 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell*, 2008, Vol 135, n° 5, p. 919-932.
Wolin, S. L., and Matera, A. G. The trials and travels of tRNA. *Genes Dev*, 1999, Vol 13, n° 1, p. 1-10.
Wolin, S. L., and Cedervall, T. The La protein. *Annu Rev Biochem*, 2002, Vol 71, n° 375-403.
Woodhams, M. D., Stadler, P. F., Penny, D., and Collins, L. J. RNase MRP and the RNA processing cascade in the eukaryotic ancestor. *BMC Evol Biol*, 2007, Vol 7 Suppl 1, n° S13.
Xiao, S., Scott, F., Fierke, C. A., and Engelke, D. R. Eukaryotic ribonuclease P: a plurality of ribonucleoprotein enzymes. *Annu Rev Biochem*, 2002, Vol 71, n° 165-189.
Xiao, S., Day-Storms, J. J., Srisawat, C., Fierke, C. A., and Engelke, D. R. Characterization of conserved sequence elements in eukaryotic RNase P RNA reveals roles in holoenzyme assembly and tRNA processing. *RNA*, 2005, Vol 11, n° 6, p. 885-896.
Xiao, S., Hsieh, J., Nugent, R. L., Coughlin, D. J., Fierke, C. A., and Engelke, D. R. Functional characterization of the conserved amino acids in Pop1p, the largest common protein subunit of yeast RNases P and MRP. *RNA*, 2006, Vol 12, n° 6, p. 1023-1037.
Xu, Y., Amero, C. D., Pulukkunat, D. K., Gopalan, V., and Foster, M. P. Solution structure of an archaeal RNase P binary protein complex: formation of the 30-kDa complex between Pyrococcus furiosus RPP21 and RPP29 is accompanied by coupled protein folding and highlights critical features for protein-protein and protein-RNA interactions. *J Mol Biol*, 2009, Vol 393, n° 5, p. 1043-1055.
Yang, L., and Altman, S. A noncoding RNA in Saccharomyces cerevisiae is an RNase P substrate. *RNA*, 2007, Vol 13, n° 5, p. 682-690.
Yoo, C. J., and Wolin, S. L. The yeast La protein is required for the 3' endonucleolytic cleavage that matures tRNA precursors. *Cell*, 1997, Vol 89, n° 3, p. 393-402.
Yuan, Y., Singh, R., and Reddy, R. Rat nucleolar 7-2 RNA is homologous to mouse mitochondrial RNase mitochondrial RNA-processing RNA. *J Biol Chem*, 1989, Vol 264, n° 25, p. 14835-14839.
Yuan, Y., and Altman, S. Substrate recognition by human RNase P: identification of small, model substrates for the enzyme. *EMBO J*, 1995, Vol 14, n° 1, p. 159-168.
Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H., and Noller, H. F. Crystal structure of the ribosome at 5.5 A resolution. *Science*, 2001, Vol 292, n° 5518, p. 883-896.
Zhu, Y., Stribinskis, V., Ramos, K. S., and Li, Y. Sequence analysis of RNase MRP RNA reveals its origination from eukaryotic RNase P RNA. *RNA*, 2006, Vol 12, n° 5, p. 699-706.
Ziehler, W. A., Yang, J., Kurochkin, A. V., Sandusky, P. O., Zuiderweg, E. R., and Engelke, D. R. Structural analysis of the P10/11-P12 RNA domain of yeast RNase P RNA and its interaction with magnesium. *Biochemistry*, 1998, Vol 37, n° 10, p. 3549-3557.
Ziehler, W. A., Day, J. J., Fierke, C. A., and Engelke, D. R. Effects of 5' leader and 3' trailer structures on pre-tRNA processing by nuclear RNase P. *Biochemistry*, 2000, Vol 39, n° 32, p. 9909-9916.

Ziehler, W. A., Morris, J., Scott, F. H., Millikin, C., and Engelke, D. R. An essential protein-binding domain of nuclear RNase P RNA. *RNA*, 2001, Vol 7, n° 4, p. 565-575.

Zuker, M., Mathews, D. H., and Turner, D. H. (1999). Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide In RNA Biochemistry and Biotechnology . J. Barciszewski and B. F. C. Clark, ed. (Dordrecht, NL: Kluwer Academic Publishers), p. 11-43.

Claire Batisse

Caractérisation biochimique et structurale des RNases P et MRP chez la levure *Saccharomyces cerevisiae*

Résumé

La RNase P est une endoribonucléase responsable de la maturation de l'extrémité 5' des ARNt prématures. Holoenzyme très conservée, elle est constituée d'une composante ARN formant le noyau catalytique et d'une composante protéique dont le nombre de sous-unités est variable : une protéine chez les bactéries, 5 chez les archées et d'au moins 9 chez les eucaryotes. Les eucaryotes possèdent également une autre endoribonucléase, la RNase MRP dont la composition est proche de la RNase P tant au niveau ribonucléique que protéique mais avec une spécificité de substrat propre.

Dans cette étude, nous proposons une méthode originale et spécifique pour purifier la RNase P et la RNase MRP de *S. cerevisiae*. Grâce à la microscopie électronique et au traitement d'images, nous avons déterminé la première structure de ces deux holoenzymes à une résolution d'environ 1.5 nm. Ces structures révèlent une architecture modulaire commune où les protéines stabilisent la composante ARN et contribuent à l'édification de cavités et de conduits. Les spécificités structurales sont localisées en des positions stratégiques pour l'identification et la coordination du substrat.

Mots-clé : RNase P, RNase MRP, Microscopie électronique, Assemblage modulaire, Maturation ARN

Résumé en anglais

Ribonuclease P (RNase P) is an endoribonuclease that cleaves the 5'-leader sequence of pre-tRNAs. RNase P is conserved between all taxonomic kingdoms and consists of a catalytic RNA subunit and protein components of variable size, from one protein in bacteria to 5 proteins in archae and at least 9 proteins in eukaryotic cells. In addition to RNase P, eukaryotes possess the RNase MRP which has a related RNA core and shares 8 proteins subunits with RNase P but with its own substrate specificity.

Here, we propose an original method to purify specifically RNase P and RNase MRP from *S. cerevisiae*. Using electron microscopy and image processing, we solved the first structure of these two holoenzymes at a resolution of about 1.5 nm. We showed that eukaryotic RNase P and RNase MRP have a modular architecture, where proteins stabilize the RNA fold and contribute to cavities, channels and chambers between the modules.Structural features are located at strategic positions for substrate recognition by shape and coordination of the cleaved-off sequence.

Keywords: RNase P/ RNase MRP/ electron microscopy/ modular assembly/ RNA-processing