HAL
open science

# Etude comparative de couples ARNt/aminoacyl-ARNt synthétases chez la levure et la mitochondrie humaine.

Aurélie Fender

## To cite this version:

Aurélie Fender. Etude comparative de couples ARNt/aminoacyl-ARNt synthétases chez la levure et la mitochondrie humaine.. Biologie cellulaire. Université Louis Pasteur - Strasbourg I, 2005. Français. NNT : . tel-00139137

**HAL Id: tel-00139137**

**https://theses.hal.science/tel-00139137v1**

Submitted on 29 Mar 2007

**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.

UNIVERSITE LOUIS PASTEUR
STRASBOURG I

ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE

# THESE

Présentée pour l'obtention du grade de

## DOCTEUR DE L'UNIVERSITE LOUIS PASTEUR

*Discipline :* Sciences du Vivant
*Spécialité :* Aspects Moléculaires et Cellulaires de la Biologie

Par

Aurélie FENDER

# Etude comparative de couples ARNt/aminoacyl-ARNt synthétases chez la levure et la mitochondrie humaine

Soutenue le 18 novembre 2005 devant la Commission d'examen

| | |
|---|---|
| Prof Catherine Florentz | Directeur de thèse |
| Dr Anne-Catherine Dock-Bregeon | Rapporteur interne |
| Dr Marc Mirande | Rapporteur externe |
| Prof Mario Mörl | Rapporteur externe |
| Dr Richard Giegé | Examinateur |

C’est avec plaisir que je tiens à remercier toutes les personnes qui ont permis l’accomplissement de cette thèse.

Tout d’abord, je remercie Madame A.-C. Dock-Bregeon, Chargée de recherche à l’Institut de Génétique et de Biologie Moléculaire et Cellulaire (Strasbourg), Messieurs M. Mirande, Directeur de recherche au Laboratoire d’Enzymologie et Biochimie Structurales (UPR 9063-CNRS)(Gif-sur-Yvette), M. Mörl, Professeur à l’Université de Leipzig, pour l’honneur qu’ils me font en jugeant mon travail.

Je remercie aussi Eric Westhof, Professeur à l’Université Louis Pasteur et Directeur de notre UPR, pour l’intérêt qu’il a porté à mon travail et pour les conseils dont il m’a fait profiter au cours de ma thèse (choix de post-doc…)

J’aimerais également remercier Richard Giegé, Directeur de recherche au CNRS, pour l’intérêt qu’il a porté à mon travail tout au long de ma thèse et maintenant en acceptant de le juger. Je tiens à vous dire que j’admire votre passion et votre dévouement pour la science. Merci aussi pour les discussions sur l’« évolution ».

J’exprime toute ma gratitude à Catherine Florentz, directrice de ma thèse et Professeur à l’Université Louis Pasteur, pour m’avoir donné l’opportunité de faire une thèse dans son équipe. Merci pour tous les « Brainstorming » que nous avons eus ensemble, pour ton investissement, ton soutien et surtout ton énergie.

Je tiens à remercier Pascale Romby, Directeur de recherche au CNRS, pour son aide et son soutien dans la recherche de post-doc.

Un énorme merci à Marie, mon coach personnel. Tu auras été présente jusqu’au bout, et je t’en suis extrêmement reconnaissante. Merci pour ton soutien, ta disponibilité sans limite, et ton investissement. Ton aide aura été précieuse à mes yeux et je crois que je n’arriverais pas à dire combien tu m’as apporté.

Merci à Joern pour m’avoir fait choisir ce laboratoire. Je tiens à te remercier pour ton soutien et toutes les discussions que nous avons eu ensemble. Ton anniversaire à l’Oktober Fest restera longtemps gravé dans ma mémoire.

Merci à toutes les personnes des labos 443-447, pour leur bonne humeur, les astuces de manips, les conseils, les pots… Merci également à l'ensemble des personnes de l'UPR qui m'ont apporté leurs connaissances, leurs conseils, merci pour votre amabilité.

Un merci tout particulier aux joyeux compagnons du midi : Frénégondine, Sir Michel, Borice et « le concombre de mer masqué ». Merci pour toutes les grandes discussions philosophiques et récréatives, et surtout les fous rires… N'oubliez pas que nous sommes tous liés dans notre entreprise…

Merci à ma famille pour leur soutien et pour tout le reste…

Et pour finir merci à Jean, le plus méritant dans l'histoire… Merci de partager ta vie à la mienne et merci pour tout le bien que tu m'apportes…

# Abréviations

| | |
|---|---|
| A | adénine |
| aa | acide aminé |
| aaRS | aminoacyl-ARNt synthétase |
| ADN | acide désoxyribonucléique |
| Arg | arginine |
| ArgRS | Arginyl-ARNt synthétase |
| ARN | acide ribonucléique |
| ARNm | acide ribonucléique méssager |
| ARNt | acide ribonucléique de transfert |
| Asp | aspartate |
| AspRS | Aspartyl-ARNt synthétase |
| ATP | adénosine triphosphate |
| BET | bromure d'éthidium |
| BSA | albumine de sérum bovin |
| C | cytidine |
| Cam | chloramphénicol |
| CIP | phosphatase intestinale de veau |
| Ci | Curie |
| Da | Dalton |
| DO | densité optique |
| DTE | dithioérythritol |
| DTT | dithiothréitol |
| EDTA | ethylène diamine tétraacétate |
| G | guanine |
| g | gramme |
| *g* | multiple de la gravitation naturelle de la terre |
| h | heure |
| Hepes | acide N-2 hydroxy éthyl pipérazine N'-2 éthane sulfonique |
| IPTG | Isopropyl-b-D-1-thiogalactopyranoside |
| Kan | kanamycine |
| kb | kilo bases |
| $k_{cat}$ | constante catalytique |
| $K_M$ | constante de Michaelis-Menten |
| µ | micro |
| M | molaire |
| mA | milliampère |
| min | minute |
| mt | mitochondrial/mitochondriaux |
| nm | nanomètre |
| PAGE | électrophorèse sur gel de polyacrylamide |
| PM | poids moléculaire |
| SDS | dodécyl sulfate de sodium |
| TBE | tris borate EDTA |
| TCA | acide trichloroacétique |
| Tris | tris(hydroxyméthyl)aminométhane |
| U | unité |

| | |
|---|---|
| UV | ultra-violet |
| V | volt |
| p/v | poids à volume |
| v/v | volume à volume |

**INTRODUCTION** ..... **6**

Préambule ..... 6

A. L'AMINOACYLATION : ETAPE CLE DE LA TRADUCTION ..... 8

*1. Les ARNt* ..... *9*

1.1. Biosynthèse et maturation ..... 9

1.2. Structures ..... 10

1.3. Rôles alternatifs ..... 12

1.4. Pseudo-ARNt ..... 13

*2. Les aminoacyl-ARNt synthétases* ..... *17*

2.1. Enzymes de classe I ..... 18

2.2. Enzymes de classe II ..... 19

2.3. Organisation modulaire et motifs structuraux ..... 20

2.4. Aminoacyl-ARNt synthétases particulières ..... 21

2.5. Complexe multisynthétasique ..... 22

2.6. Mécanisme d'édition ..... 24

2.7. Fonctions alternatives ..... 24

2.8. Paralogues d'aminoacyl-ARNt synthétases ..... 26

2.9. Les aminoacyl-ARNt synthétases comme cibles thérapeutiques ..... 27

*3. Le second code génétique* ..... *28*

3.1. Identité des ARNt ..... 29

Approches expérimentales ..... 30

Transplantation ..... 31

Localisation des éléments d'identité ..... 32

Identité ambiguë et antidéterminants ..... 33

Subtilité de l'identité ..... 34

3.2. Identité des aminoacyl-ARNt synthétases ..... 34

3.3. Interaction des ARNt/aaRS ..... 35

3.4. Aminoacylation de fragments d'ARN et systèmes d'aminoacylation artificiels ..... 37

*4. Aspects évolutifs* ..... *39*

4.1. Le paradigme actuel : de la soupe primordiale aux trois règnes du vivant ..... 40

4.2. Émergence des systèmes d'aminoacylation du monde des ARN ..... 42

4.3. Évolution des ARNt ..... 43

4.4. Évolution des aminoacyl-ARNt synthétases ..... 45

4.5. Conservation des éléments d'identité ou barrières d'espèces ..... 47

B. LE SYSTEME D'ASPARTYLATION DANS DIVERS ORGANISMES ..... 48

*1. Identité des ARNt$^{Asp}$* ..... *48*

*2. Structure/fonction de l'AspRS* ..... *51*

*3. Interaction ARNt$^{Asp}$/AspRS* ..... *57*

C. LES SYSTEMES D'AMINOACYLATION MITOCHONDRIAUX DE MAMMIFERES ..... 58

*1. Généralités* ..... *59*

1.1. Origine endosymbiotique ..... 59

1.2. Structure et fonction de la mitochondrie ..... 60

1.3. Expression du génome ..... 61

*2. Les systèmes d'aminoacylation mitochondriaux* ..... *65*

2.1. Les ARNt mitochondriaux ..... 65

Maturation ..... 66

Structures secondaires et tertiaires ..... 68

Rôle alternatif ..... 69

Mutations des ARNt : polymorphisme versus pathologie ..... 70

2.2. Les aminoacyl-ARNt synthétases mitochondriales ..... 71

Séquences et structures ..... 71

Propriétés d'aminoacylation ..... 73

Fonctions alternatives ..... 74

D. OBJECTIFS DE LA THESE ..... 75

**CHAPITRE 1 :** Etude des systèmes d'aminoacylation arginine et aspartate chez la levure ..... **78**

A. PARTICULARITES D'UNE FAMILLE ISOACCEPTRICE ..... 80

article n°1 : Idiosyncrasies fonctionnelles d'ARNt isoaccepteurs dans les systèmes d'aminoacylation spécifique et non spécifique

B. RELATION FONCTIONNELLE ET EVOLUTIVE ENTRE LES SYSTEMES ARGININE ET ASPARTATE DE LEVURE. 86

article n°2 : Réminiscence d'une ancienne spécificité dans un ARNt contemporain

**CHAPITRE 2 :** Etude du système d'aspartylation mitochondrial humain ..... **92**

A. VERS LE JEU COMPLET D'AMINOACYL-ARNT SYNTHETASES MITOCHONDRIALES HUMAINES : CARACTERISATION DE L'ASPRS ET DE LA TYRRS (article n°3) ..... 94
B. ELEMENTS D'IDENTITE DANS UN ARNT MITOCHONDRIAL ..... 106
article n°4: Evolution reverse d'un système d'aminoacylation mitochondrial
*1. Identité des ARNt$^{Asp}$ mt de mammifères ..... 110*
1.1. La perte du résidu 73 comme élément d'identité est spécifique des ARNt$^{Asp}$ mt de mammifères ..... 110
1.2. Eléments d'identité complémentaires ? ..... 111
1.3. ARNt$^{Asp}$ et mutations pathogéniques ..... 113
*2. Identité des autres systèmes mitochondriaux de mammifères ..... 114*
*3. Adapatation de l'AspRS mt humaine à son ARNt$^{Asp}$ ..... 119*
C. UNICITE DE RECONNAISSANCE DE L'ARNT$^{ASP}$ MT HUMAIN ..... 122
*1. Aminoacylations croisées ou barrières d'espèces ..... 122*
*2. Non reconnaissance d'un ARNt dégénéré ? ..... 123*
*3. Nécessité des éléments d'identité aspartate d'*E. coli ..... *124*
*4. Implication de l'architecture centrale de l'ARNt ..... 128*
*5. Recherche d'éléments antidéterminants ..... 129*
*6. Effet additif des mutations ..... 132*
*7. Conclusion et perspectives ..... 132*
D. ESSAIS DE CRISTALLISATION ..... 134
*1. Approche préliminaire et stratégies ..... 135*
*2. Recherche de conditions de cristallisation ..... 136*
*3. Echange de tampon et effet sur la solubilité de l'AspRS ..... 137*
*4. Conditions d'expression et de purification ..... 138*
*5. Conclusions et perspectives ..... 140*

**CONCLUSIONS & PERSPECTIVES ..... 142**

A. SYSTEMES DE LEVURE ..... 142
B. SYSTEMES MITOCHONDRIAUX DE MAMMIFERES ..... 144

**MATERIEL & METHODES ..... 150**

A. MATERIEL ..... 150
*1. Produits chimiques ..... 150*
*2. Enzymes commerciales ..... 150*
*3. Enzymes non commerciales ..... 151*
*4. Souches cellulaires ..... 151*
4.1. Souches d'*E. coli* ..... 151
4.2. Souche de *S. cerevisiae* ..... 152
*5. Plasmides ..... 152*
B. MILIEUX DE CULTURE ..... 153
C. TECHNIQUES COURANTES ..... 154
*1. Préparation de bactéries compétentes et transformation ..... 154*
Méthode au $CaCl_2$ ..... 154
Préparation et transformation de bactéries électrocompétentes ..... 154
*2. Etude des acides nucléiques ..... 155*
2.1. Détermination de la concentration d'une solution d'acides nucléiques ..... 155
2.2. Extraction phénolique ..... 155
2.3. Précipitation des acides nucléiques ..... 156
2.4. Amplification d'ADN plasmidique ..... 156
2.5. Migration des acides nucléiques sur gel ..... 156
Electrophorèse sur gel d'agarose ..... 156
Electrophorèse sur gel de polyacrylamide ..... 157
2.6. Elution des ARN à partir d'un gel de polyacrylamide ..... 157
Elution passive ..... 157
Electroélution ..... 157
2.7. Northern blot ..... 158
*3. Etude des protéines ..... 160*
3.1. Détermination de la concentration d'une solution de protéines ..... 160
3.2. Précipitation des protéines ..... 160
3.3. Electrophorèse en conditions dénaturantes (SDS-PAGE) ..... 160
D. PREPARATION D'ARNT PAR CLONAGE ET TRANSCRIPTION *IN VITRO* ..... 161
*1. Principe ..... 161*
*2. Construction d'un gène synthétique ..... 162*

*3. Mutagénèse dirigée* ... *162*
*4. Clonage dans pTFMa* ... *163*
*5. Sélection de clones positifs* ... *164*
*6. Transcription* in vitro ... *164*
*7. Purification de transcrits sur gel préparatifs* ... *165*
*8. Analyse de l'extrémité 3' des ARNt transcrits* ... *165*
E. CLONAGE, SUREXPRESSION ET PURIFICATION DE L'ASPRS MITOCHONDRIALE HUMAINE ... 167
*1. Clonage du gène à partir d'une banque d'ADNc humains* ... *167*
*2. Expression et purification* ... *168*
F. AMINOACYLATION *IN VITRO* ... 169
*1. Détermination des taux de charge et des constantes cinétiques* ... *169*
*2. Conditons optimales d'aminoacylation* ... *170*
2.1. Conditions optimales pour l'ArgRS de levure ... 170
2.2. Conditions optimales pour l'AspRS de levure ... 171
2.3. Conditions optimales pour l'AspRS mt humaine ... 171
G. ETUDE DE L'AMINOACYLATION *IN VIVO* CHEZ LA LEVURE ... 171
*1. Construction du gène muté dans le vecteur pRS314* ... *171*
*2. Transformation de la levure, sélection blanc/rouge* ... *172*
*3. Analyse d'ARNt par northern blot en condition acide* ... *172*
3.1. Préparation des ARN cellulaires en conditions acides ... 172
3.2. Séparation des ARN sur gel acide ... 173
3.3. Transfert sur membrane de nylon ... 173
3.4. Marquage des sondes en 5' ... 173
3.5. Hybridation des sondes radioactives et révélation ... 174
H. PRODUCTION D'ANTICORPS ET TEST ELISA ... 175
*1. Préparation de l'antigène* ... *175*
*2. Déroulement des injections* ... *175*
*3. Test Elisa* ... *176*
I. ESSAIS DE CRISTALLISATION ... 178
*1. Généralités* ... *178*
*2. Principe général de la cristallisation* ... *179*
*3. Principe de la cristallisation par diffusion de vapeur* ... *180*
*4. Préparation du matériel biologique* ... *181*

**REFERENCES ... 186**

# INTRODUCTION

## Préambule

La traduction de l'information génétique en protéines se place au cœur de l'expression de toute forme de vie. Elle met en oeuvre une orchestration fine, précise et spécifique de très nombreuses macromolécules incluant protéines et ARN. Les ARN de transfert (ARNt) et les aminoacyl-ARNt synthétases (aaRS), enzymes qui catalysent la fixation d'un acide aminé sur l'extrémité de l'ARNt, occupent une place cruciale dans ce processus en contribuant à la sélection et au transfert des acides aminés spécifiques qui seront incorporés dans les protéines naissantes au niveau du ribosome.

Les règles de reconnaissance entre aaRS et ARNt, (notamment les signaux essentiels la promouvant) ont été largement étudiées et l'on connaît aujourd'hui les « jeux d'éléments d'identité » permettant la spécificité d'aminoacylation pour de nombreux systèmes procaryotiques et dans une moindre mesure pour des systèmes eucaryotiques. Cependant, on est encore loin de comprendre toutes les particularités et finesses des systèmes d'aminoacylation, la nature des relations entre les différents systèmes, les questions qui touchent à leur évolution,.… Par ailleurs, les systèmes d'aminoacylation de certains organismes restent largement inexplorés.

Dans ce travail, les originalités et subtilités des mécanismes sous-jacents à la spécificité d'aminoacylation des ARNt par les aaRS tant au sein des systèmes classiques que de systèmes atypiques ont été explorés. Tout d'abord, les règles de l'identité au sein des systèmes classiques d'aminoacylation spécifiques de l'arginine et de l'aspartate chez la levure *Saccharomyces cerevisiae* ont été affinées. L'étude conjointe de ces deux systèmes nous a permis d'approcher de façon originale l'identité au sein d'une famille d'ARNt isoaccepteurs. Au vu de la relation fonctionnelle entre ces systèmes (démontrée précédemment au laboratoire), l'évolution des gènes des $ARNt^{Asp}$ et $ARNt^{Arg}$ a été analysée.

En parallèle, des systèmes atypiques ont été explorés. Ceux-ci concernent les ARNt et aaRS de mitochondries de mammifères et en particulier le système aspartate de la mitochondrie humaine. Les connaissances sur les ARNt mitochondriaux (mt) et leurs partenaires sont encore limitées. Cependant, ces systèmes sont originaux car leurs ARNt, codés par le génome mt, possèdent des structures dégénérées, alors que les aaRS mt, codées

par le génome nucléaire, ressemblent aux synthétases classiques. Ceci laisse présager de nouvelles règles de reconnaissance entre ces deux molécules et en particulier de nouveaux signaux d'identité. Le choix du couple spécifique de l'acide aspartique a été largement orienté par le fait que les systèmes correspondants de levure et bactériens (*E. coli* et *T. thermophilus*) sont bien connus. Il sera donc possible de réaliser une analyse comparative des systèmes d'aspartylation et ainsi d'aborder la question de leur évolution.

L'Introduction de cette thèse est partagée en 4 parties. Tout d'abord, les différents aspects concernant l'aminoacylation et les molécules mises en jeu, ARNt et aminoacyl-ARNt synthétases, sont présentés en détail. Cette partie se divise en 4 grandes sections, avec les deux premières consacrées à la présentation des ARNt et des aminoacyl-ARNt synthétases tant d'un point de vue fonctionnel que structural, illustrant directement la diversité de ces molécules et justifiant largement l'intérêt qui leur porté. La section suivante concerne les mécanismes sous-jacents à la spécificité de reconnaissance des ARNt par les aaRS, avec un accent particulier sur l'identité, mieux connue pour les ARNt, comme garant de la spécificité d'aminoacylation et par conséquent de la fidélité de la traduction. Le mode d'interaction des couples aaRS/ARNt et l'ingénierie des systèmes d'aminoacylation sont également développés dans cette section. Finalement, la 4$^{ème}$ section résume les connaissances actuelles quant à l'origine et l'évolution des systèmes d'aminoacylation.

Une deuxième partie de l'Introduction est dévolue au système d'aminoacylation spécifique de l'aspartate dans des systèmes modèles eucaryotiques et procaryotiques étudiés jusqu'à présent.

Dans une troisième partie, les systèmes d'aminoacylation du compartiment mitochondrial chez les mammifères sont présentés et replacés dans le contexte de la mitochondrie. Ceci inclut les connaissances de base sur les mitochondries, l'origine de ces organelles, les particularités du génome mitochondrial et ses mécanismes d'expression. Un état des lieux des connaissances concernant les ARNt et les aaRS des mitochondries de mammifères, et notamment de la mitochondrie humaine, est établi. Les particularités de structures et de séquences des ARNt par rapport aux molécules canoniques sont décrites.

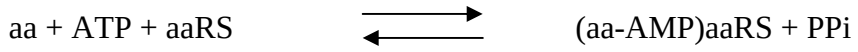
La dernière partie de l'Introduction présente les objectifs détaillés de cette thèse.

# A. L’aminoacylation : étape clé de la traduction

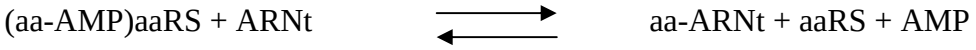
La traduction de l’information génétique est une étape fondamentale dans l’expression de toute forme de vie. Elle est la conversion d’un langage nucléique, composé d’une succession de triplets (codons) portés d’abord par l’ADN puis par l’ARN messager (ARNm), en un langage protéique. Cette étape se fait au niveau du ribosome où les ARN de transfert (ARNt), en tant que molécules adaptatrices, associent à chaque codon complémentaire à leur anticodon l’acide aminé (aa) correspondant. La fidélité de la traduction repose ainsi essentiellement sur le déchiffrage correct du code génétique (reconnaissance codon/anticodon) et la fixation spécifique de l’acide aminé sur l’ARNt catalysée par l’aminoacyl-ARNt synthétase (aaRS) correspondante. Nous allons nous intéresser plus particulièrement à cette deuxième étape.

Pour chacun des acides aminés, il existe généralement une aaRS et un ARNt ou une famille d’ARNt isoaccepteurs. Les caractéristiques structurales et fonctionnelles des deux partenaires de la réaction d’aminoacylation sont largement étudiées (Cusack, 1997 ; Giegé et al., 1998 ; Martinis et al., 1999 ; Francklyn et al., 2002 ; Giegé & Frugier, 2003). Il est établi en particulier, que les ARNt classiques ont des structures secondaires en forme de feuille de trèfle et tertiaires en forme de « L » constitué d’une branche acceptrice de l’acide aminé et d’une branche anticodon. Par ailleurs, les vingt aaRS sont des enzymes modulaires, avec un domaine catalytique et un domaine de liaison à l’anticodon, qui se distinguent les unes des autres par leur structure tertiaire et leur composition oligomérique.

La réaction d’aminoacylation se déroule en deux étapes. Dans un premier temps, l’aaRS active l’acide aminé qui lui est spécifique en aminoacyl-adénylate (aa-AMP), en présence de magnésium et d’ATP. Certaines aaRS nécessitent la fixation de l’ARNt pour catalyser cette première réaction. Il s’agit de la GluRS, GlnRS, ArgRS et de la LysRS de classe I.




Puis l’acide aminé activé est transféré sur l’extrémité 3’ acceptrice de l’ARNt.



# 1. Les ARNt

Dès 1955, F. Crick prédisait l'existence de molécules adaptatrices faisant le lien entre le langage nucléique porté par l'ARNm et la chaîne polypeptidique naissante au niveau du ribosome (Crick, 1955). Découverts seulement trois ans après par (Hoagland et al., 1958), les ARNt sont devenus des molécules phares et actuellement environ 5800 séquences de gènes d'ARNt de 111 organismes différents (eucaryotes, archaea, bactéries, organelles) sont répertoriées dans la banque d'ARNt (Sprinzl & Vassilenko, 2005 ; http://www.tRNA.uni-bayreuth.de). Ce nombre est en croissance continue grâce aux nombreux projets de séquençage de génome en cours. De plus, le séquençage au niveau ARN, a permis d'identifier plus de 80 modifications post-transcriptionnelles (Björk, 1995 ; McCloskey & Crain, 1998 ; Sprinzl & Vassilenko, 2005). Si la séquence primaire du premier ARNt séquencé ($ARNt^{Ala}$, (Holley et al., 1965)) permettait déjà de prédire une structure secondaire en feuille de trèfle, il a fallu attendre 1974 pour obtenir la première image tridimensionnelle d'un ARNt grâce à la résolution de la structure cristallographique de l'$ARNt^{Phe}$ de levure (Kim et al., 1974 ; Robertus et al., 1974).

Outre les fonctions et structures des ARNt, leur histoire évolutive intéresse un nombre croissant de biologistes. En effet, l'accessibilité à de très nombreuses séquences d'ARNt de tous les règnes du vivant a généré un engouement tout particulier pour l'étude phylogénétique de ces gènes.

## 1.1. Biosynthèse et maturation

Chez la bactérie, les gènes d'ARNt sont généralement organisés en opéron avec d'autres gènes d'ARNt et/ou d'ARNr et/ou d'ARNm. Chez les eucaryotes, les gènes d'ARNt sont isolés et sont transcrits par l'ARN polymérase III (les ARN polymérases I et II synthétisant les ARNr et ARNm, respectivement). Une fois transcrits, les pre-ARNt passent par six étapes majeures de maturation (Hopper & Phizicky, 2003) : (i) le clivage en 5' par la RNase P (Harris & Christian, 2003 ; Xiao et al., 2005), (ii) la digestion en 3' par des endonucléases et des exonucléases (Deutscher, 1995 ; Mörl & Marchfelder, 2001 ; Schiffer et al., 2002), (iii) l'addition de la séquence CCA en 3' pour les ARNt des eucaryotes, de

beaucoup de bactéries, de quelques archaebactéries et des organelles (Levinger et al., 2004; Weiner, 2004 ), (iv) l'épissage des introns de certains ARNt d'eucaryotes et de quelques archaebactéries (Abelson et al., 1998), (v) la modification (Ikeuchi Y, 2005), et (vi) l'édition (modification de la séquence codée) (Placido et al., 2005).

Récemment, il a été montré que certains ARNt (Glu, His, Trp et iMet) chez l'archaea *Nanoarchaeum equitans* sont reconstitués à partir de deux gènes (Randau et al., 2005). La séparation entre les moitiés de gène est localisée juste après le nucléotide 37, un nucléotide après l'anticodon (position classique pour les introns). Après transcription, les deux ARN s'hybrideraient par leur région 5' et 3' riches en GC et complémentaires, et les ARNt seraient libérés suite à un clivage en *trans* des régions additionnelles et ligation.

Plusieurs groupes ont observé l'aminoacylation des ARNt dans le noyau, chez la levure et dans les cellules de vertébrés (Lund & Dahlberg, 1998 ; Sarkar et al., 1999 ; Grosshans et al., 2000). Très critiqué au moment de la parution des différents articles, aujourd'hui de nouvelles découvertes relancent le sujet. Les ARNt de levure sont épissés dans le cytoplasme (Yoshihisa et al., 2003) et récemment deux groupes indépendants (dont celui de Yoshihisa) ont mis en évidence que ces ARNt cytoplasmiques matures peuvent être redirigés dans le noyau (Shaheen & Hopper, 2005; Takano et al., 2005). Deux types d'import des ARNt dans le noyau sont décrits, l'un constitutif qui permettrait le contrôle de l'intégrité des ARNt (spécialement après épissage) et l'autre en condition de carence en acides aminés qui permettrait de séquestrer les ARNt dans le noyau et de réduire la synthèse de protéines. Ce phénomène n'a pas encore été exploré dans d'autres organismes. Notons que l'épissage cytoplasmique des ARNt semble spécifique à la levure (Hopper & Phizicky, 2003).

## 1.2. Structures

La majorité des séquences connues d'ARNt à ce jour, comprises entre 75 et 95 nucléotides, adoptent une structure secondaire canonique en feuille de trèfle comportant les différents domaines structuraux et fonctionnels de l'ARNt. Illustrée dans la Figure 1, la structure secondaire de l'ARNt est composée de 4 régions appariées en hélice (les bras accepteur de l'acide aminé avec une extrémité NCCA 3' sortante, D, de l'anticodon et T), de trois régions simple brin (les boucles D, de l'anticodon, et T, et la région variable).

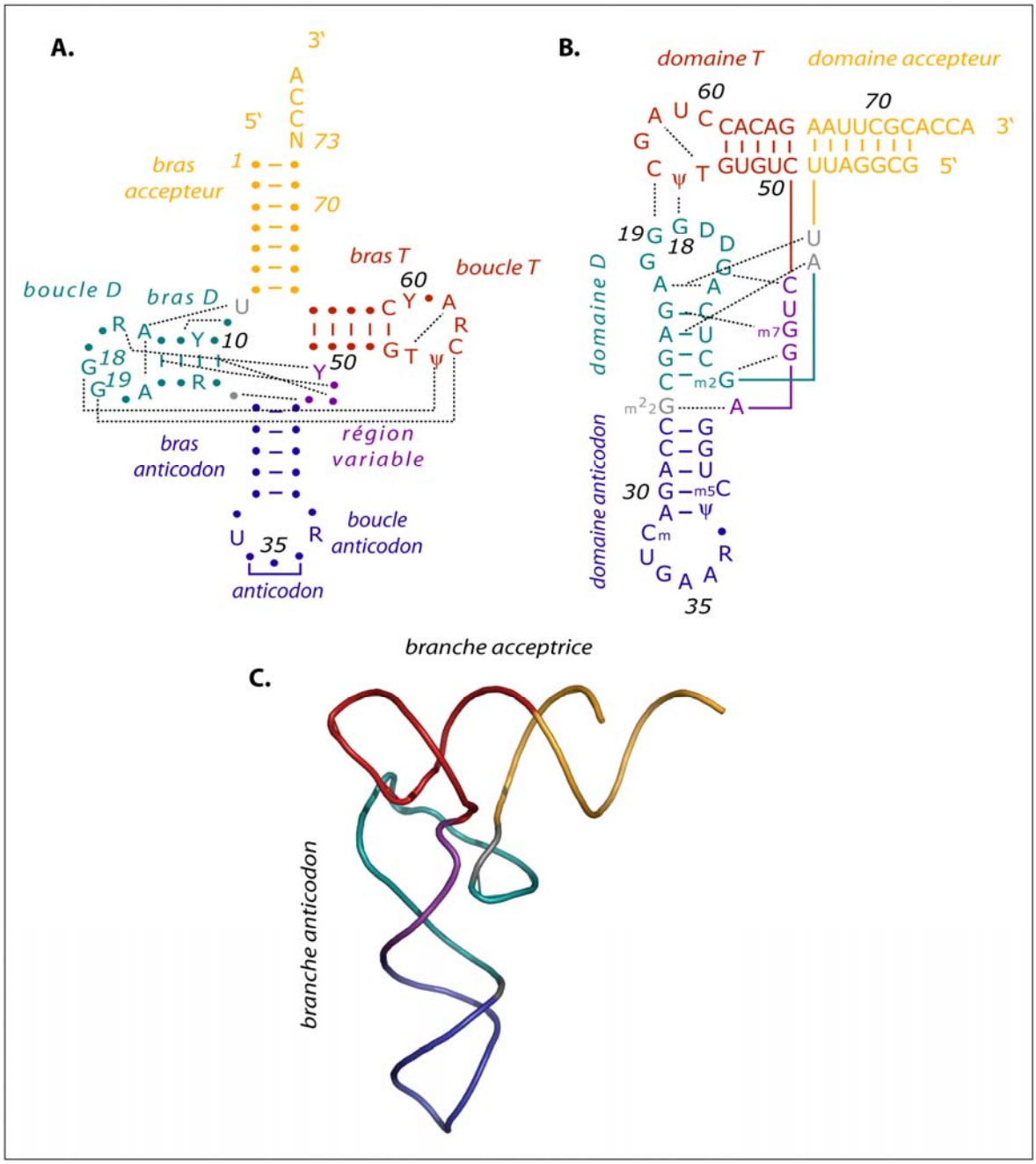



**Figure 1. Structures secondaire (A) et tertiaires (B&C) d'ARNt classiques.** Les différents domaines structuraux sont indiqués. Les interactions à longue distance permettant le repliement tertiaire sont représentées par des traits en pointillé. **A.** Séquence consensus des ARNt canoniques. **B.** Séquence de l'ARNt$^{Phe}$ de levure. **C.** Structure cristallographique de l'ARNt$^{Phe}$ de levure (Jovine et al., 2000). La numérotation ainsi que la nomenclature des bases modifiées correspondent à celles établies par (Sprinzl & Vassilenko, 2005). Y et R indiquent des pyrimidines et des purines, respectivement.

La comparaison des séquences a permis l'identification de résidus conservés (U8, A14, G18, G19, A21, U33, G53, T54, ψ55, C56, A58, C61, C74, C75 et A76) et semi-conservés (Y11, R15, R24, Y32, R37, Y48, R57 et Y60) dans presque tous les ARNt, les ARNt d'organelles faisant exceptions (voir partie C.2.1.). Les structures cristallographiques

d'ARNt montrent que la plupart de ces résidus sont impliqués dans des interactions tertiaires, qui, conjointement aux interactions secondaires, imposent un repliement tridimensionnel en forme de « L » composé de la branche acceptrice de l'acide aminé (résultat de l'empilement des bras accepteur et T), de la zone charnière et de la branche de l'anticodon (empilement des bras D et de l'anticodon) (Figure 1). L'extrémité acceptrice et la boucle anticodon sont approximativement distantes de 80 Å. Bien que la structure tridimensionnelle globale soit conservée, certaines particularités rendent unique chaque structure d'ARNt. L'angle entre la branche acceptrice et la branche anticodon varie de 90° à 110°. La région variable peut former un bras supplémentaire (cas des $ARN^{Tyr}$ procaryotiques, l'$ARNt^{Ser}$ et $ARNt^{Leu}$, (Rich & RajBhandary, 1976)). La composition nucléotidique et en modifications post-transcriptionnelles a un rôle au niveau de la flexibilité, ou bien encore l'interaction tertiaire inhabituelle G15•G48 de la paire dite de Levitt participe à l'échafaudage propre de l'$ARNt^{Cys}$. Un certain nombre de ces particularités sont impliquées dans la reconnaissance par l'aaRS spécifique.

## 1.3. Rôles alternatifs

Les ARNt sont des molécules retrouvées dans tous les règnes du vivant. Abondants dans la cellule, en plus d'être des adaptateurs moléculaires, ils peuvent être impliqués dans divers mécanismes. Citons les exemples de la régulation de gènes, l'initiation de la rétro-transcription ou la synthèse de l'hème de la chlorophylle.

Chez *Bacillus subtilis* et d'autres bactéries Gram+, un mécanisme d'antiterminaison régule l'expression de la plupart des aaRS (16 gènes) et d'autres enzymes impliquées dans la biosynthèse d'acides aminés (revue par (Putzer et al., 1995 ; Ryckelynck et al., 2005)). La région régulatrice, longue de ~300 nucléotides, est composée d'une série de structures en tige boucle suivie d'une séquence conservée « T-box » en amont d'une structure en tige boucle terminatrice. En cas de carence en acide aminé, l'ARNt libre interagit via son anticodon et son extrémité CCA-3' non chargée au niveau d'un codon spécifique et de la T-box de la zone régulatrice respectivement, et induit une conformation alternative antiterminatrice. Dans ce cas, la transcription du gène est poursuivie et le ribosome progresse.

Les rétrovirus et les rétro-transposons à LTR (« Long Terminal Repeat ») utilisent des ARNt cellulaires comme amorce de la rétro-transcription pendant leurs cycles de réplication.

Dans le cas des rétrovirus, l'ARNt amorce est sélectivement empaqueté dans les particules virales ou pseudo-virales, où il interagit avec une séquence complémentaire à son extrémité 3' présente en 5' du génome viral à ARN simple brin et appelé PBS (« Primer Binding Site »). Le PBS des rétro-transposons peut également être complémentaire à une région interne de l'ARNt amorce. Des interactions additionnelles entre l'ARNt et d'autres régions du rétro-ARN (hors PBS) ont été mis en évidence par exemple dans le cas du virus de l'immunodéficience humaine de type 1 (VIH-1) ou du rétro-transposon de levure Ty1. L'ARNt amorce sélectionné varie selon les rétrovirus et rétro-transposons. Par exemple, le virus murin Moloney leukemia utilise l'ARNt$^{Pro}$, le virus de la leucose aviaire l'ARNt$^{Trp}$, les lentivirus (tels que le VIH ou le virus de l'immunodéficience du singe) choisissent préférentiellement l'ARNt$_3^{Lys}$. Enfin, les rétro-transposons Ty1 et Ty2 de levure utilisent l'ARNt$_i^{Met}$. L'ensemble de ces informations est revu par (Marquet et al., 1995 ; Le Grice, 2003).

Dans les organismes photosynthétiques, la première étape de la biosynthèse de la chlorophylle nécessite l'ARNt$^{Glu}$ aminoacylé par la GluRS. L'acide glutamique est réduit de façon ARNt-dépendante par une réductase en glutamate 1-semialdéhyde. Les autres étapes ne nécessitent plus l'ARNt$^{Glu}$ et conduisent à la formation des molécules porphyriniques (Jahn et al., 1992 ; Verkamp et al., 1995).

## 1.4. Pseudo-ARNt

La Figure 2 donne des exemples de repliement d'ARN naturels dont la structure tridimensionnelle mime celles des ARNt canoniques. Toutes ces molécules ont la capacité d'être reconnues par des aaRS, voire d'être aminoacylés. Parmi les dérivés d'ARNt, citons l'ARN transfer-messager (tm), les structures en pseudo-ARNt retrouvées en 3' de certains ARN génomiques de virus de plante (par exemple TYMV, TMV ou BMV) et dans les régions régulatrices d'ARNm. Notons aussi l'ARNt accepteur de la pyrrolysine (22$^{ème}$ acide aminé) dont la structure atypique se rapproche des ARNt mitochondriaux (Théobald-Dietrich et al., 2004). De plus, un nouveau type de dérivé d'ARNt a été récemment mis en évidence (Ataide et al., 2005) (voir plus loin).

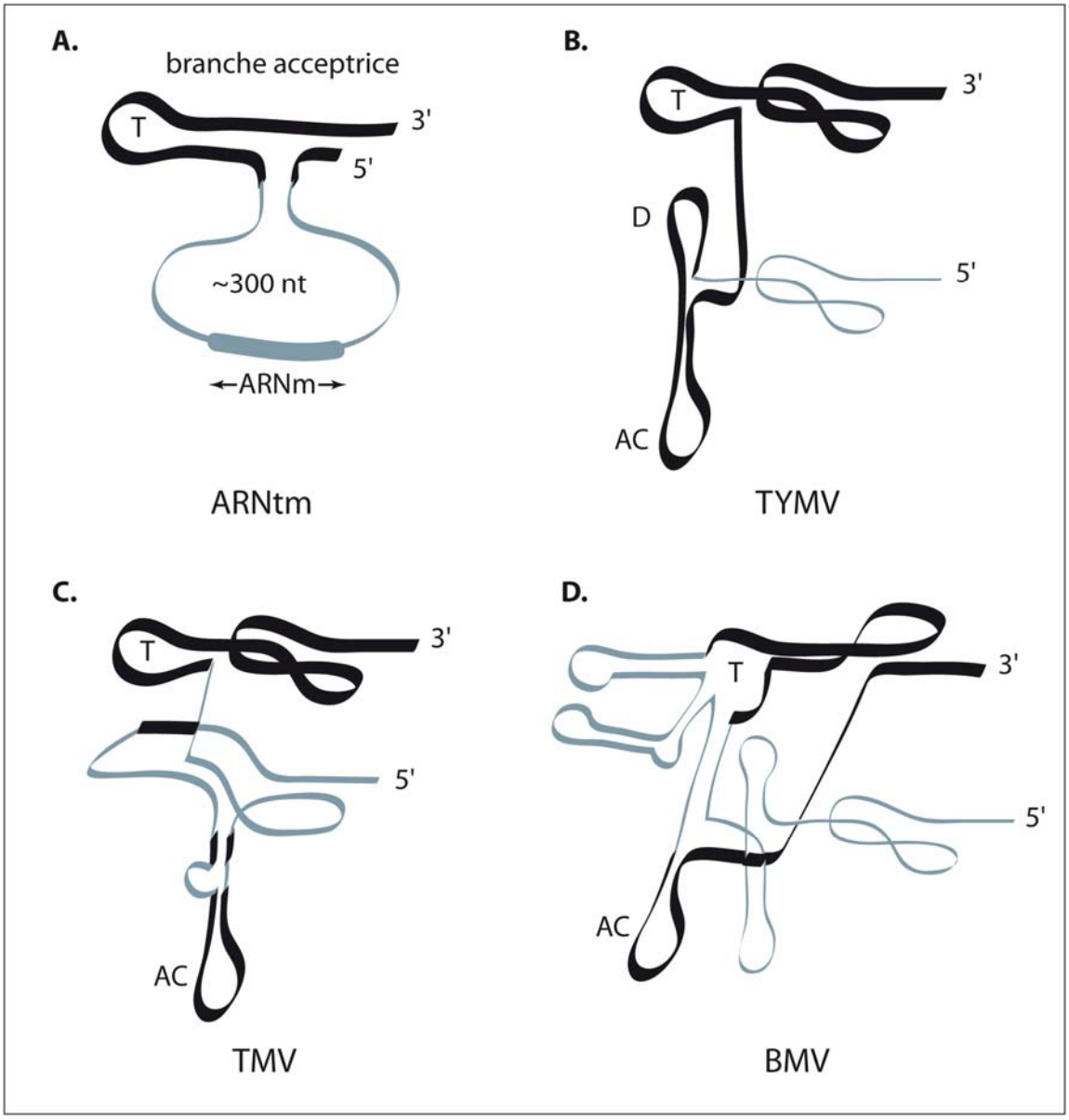



**Figure 2. Repliements schématiques de pseudo-ARNt.** Les domaines mimant l'ARNt sont indiqués en noir (T pour domaine T, D pour domaine D et AC pour domaine anticodon de l'ARNt). **A.** Représentation schématique de l'ARN trans-messager. **B-D.** Pseudo-ARNt retrouvées en 3' des ARN génomiques des virus de la mosaïque jaune du navet (TYMV, **B.**), de la mosaïque du tabac (TMV, **C.**) et de la mosaïque du brome (BMV, **D.**). Les structures en pseudo-nœuds sont schématisées par un huit. Figure adaptée de (Giegé & Frugier, 2003).

L'ARNtm, constitué de ~260 à 430 nucléotides, est conservé dans le règne des bactéries. Il est impliqué dans la dégradation des protéines aberrantes synthétisées à partir d'ARNm tronqués et dépourvus de codons stop (Atkins & Gesteland, 1996 ; Jentsch, 1996 ; Keiler et al., 1996). L'ARNtm est une molécule chimère mimant la branche acceptrice de l'ARNt$^{Ala}$ (Komine et al., 1994) et un ARNm codant pour un peptide signal de dégradation.

Sa partie ARNt lui permet d'être aminocylé par l'AlaRS et d'entrer au niveau du ribosome en pause pour incorporer une alanine à la chaîne polypeptidique en cours de synthèse. Puis, sa partie messager est lue et permet l'ajout d'un signal de dégradation et la terminaison de la synthèse de la protéine aberrante. La structure pseudo-ARNt implique les extrémités 5' et 3' de l'ARN et ressemble fortement au bras accepteur et au domaine T de l'ARNt$^{Ala}$. Les domaines D et anticodon ne sont pas conservés. La partie messager est insérée dans la structure ARNt. Dans le cas des $\alpha$-protéobactéries, l'ARNtm mature résulte de l'association de deux ARN distincts dérivant d'un seul gène (Keiler et al., 2000).

Des structures en pseudo-ARNt sont retrouvées dans les régions régulatrices de certains ARNm telles que la région régulatrice de l'opéron histidine d'*E. coli* (Ames et al., 1983), les régions régulant l'expression des ARNm de la MetRS (Dardel et al., 1990) ou de la ThrRS d'*E. coli* (Springer et al., 1989). Dans ce dernier cas, la partie en amont du gène codant pour la ThRS se replie sous la forme de deux bras anticodon, de séquences similaires à celle de l'ARNt$^{Thr}$. La fixation de la ThRS, empêche la fixation du ribosome et ainsi inhibe la traduction de son propre gène, en cas de pénurie en ARN$^{Thr}$ libres dans la cellule (Romby et al., 1996 ; Caillet et al., 2003). Alternativement un fort taux d'ARNt$^{Thr}$ libres induit le déplacement de l'enzyme liée à la région régulatrice et permet la traduction.

L'ARN génomique d'un certain nombre de virus de plantes comporte une structure en pseudo-ARNt à l'extrémité 3' (revue par (Fechter et al., 2001)). Toutes ces structures présentent un bras accepteur grâce à la présence d'un pseudo-nœud qui permet un repliement tridimensionnel globalement proche du L des ARNt canoniques. Les pseudo-ARNt peuvent être aminoacylés par l'HisRS (TMV), la ValRS (TYMV) et la TyrRS BMV). La plus forte similarité de séquence avec un ARNt canonique est retrouvée dans les structures acceptrices de la valine (par exemple le virus de la mosaïque jaune du navet ou TYMV abrégé de la traduction anglaise). Ces structures semblent jouer un rôle dans le mécanisme de leur propre réplication (revue par (Fechter et al., 2001)). Récemment, il a été montré que la région pseudo-ARNt de TYMV, indispensable pour la survie du virus, est impliquée dans un nouveau mécanisme d'initiation interne de la traduction chez les eucaryotes (Barends et al., 2003). Le pseudo-ARNt valylé induit l'initiation interne de la synthèse de la polyprotéine (codée par le génome viral), et dans ce cas, le premier acide aminé incorporé est la valine.

Un nouveau pseudo-ARNt vient d'être mis en évidence chez le pathogène *Bacillus cereus* (Ataide et al., 2005). La séquence de ce pseudo-ARNt, que les auteurs ont nommé « ARNt$^{Other}$ », peut adopter une structure secondaire en feuille de trèfle bien qu'il présente de nombreuses déviations par rapport aux ARNt canoniques. La Figure 3 représente deux

repliements secondaires prédictifs de cette molécule Ces structures diffèrent de celles des ARNt canoniques par (i) un nombre inférieur de paires de bases (6 au lieu de 7) dans le bras accepteur, 4 nucléotides au lieu de 2 dans la région connectrice entre les bras D et T et une boucle anticodon de 9 résidus au lieu de 7 (structure de gauche), ou (ii) un nombre élevé de mésappariements (structure de droite). L'ARNt$^{Other}$, bien que possédant un anticodon Trp, est le substrat des deux LysRS associées (LysRS-1 de classe I et LysRS-2 de classe II qui coexistent dans *B. cereus*, voir §A.2.4). La nature de l'acide aminé incorporé *in vivo* n'est pas connue, cependant l'ARNt$^{Other}$ est chargé par la lysine *in vitro*. L'aminoacylation par la TrpRS n'est pas exclue, d'autant plus que *B. cereus* possède une TrpRS particulière, similaire à celle de *Deinococcus radiodurans* qui peut utiliser des dérivés du tryptophane (Buddha & Crane, 2005). Pour le moment, la fonction de l'ARNt$^{Other}$ n'est pas élucidée mais son aminoacylation pourrait avoir un rôle dans la synthèse protéique en phase stationnaire de croissance, phase durant laquelle il est exprimé de façon prédominante.

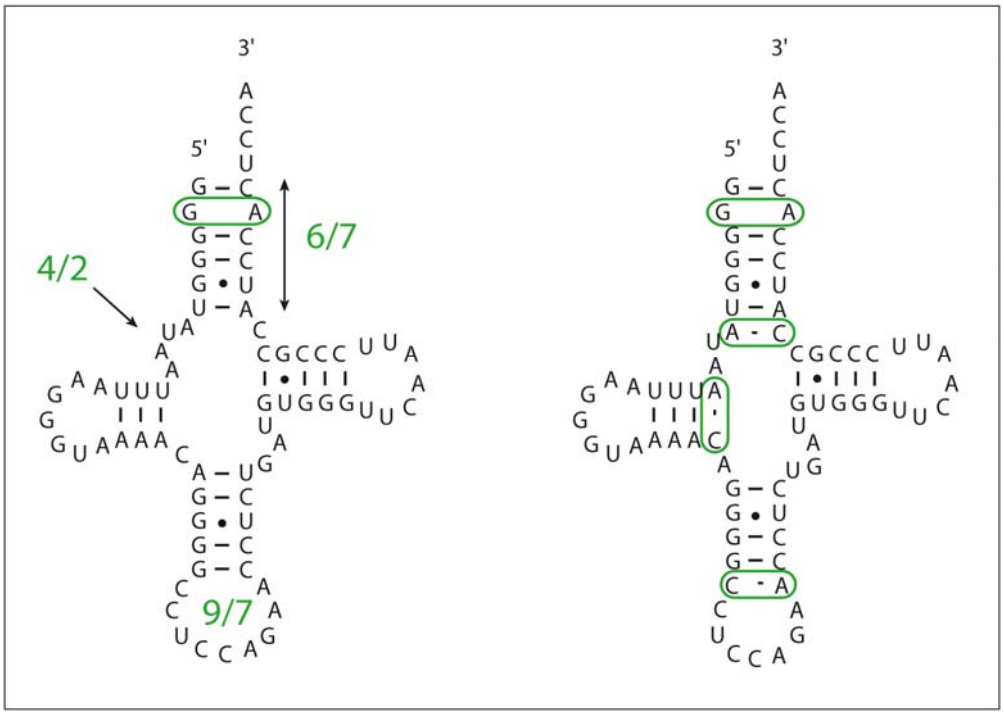



**Figure 3. Structures secondaires prédictives de l'ARNt$^{Other}$ de *B. cereus*.** Les différentes déviations structurales par rapport aux ARNt canoniques sont indiquées en vert. Les paires de bases non-canoniques sont encerclées et le rapport x/y donne le nombre de nucléotides d'un domaine de l'ARNt bizarre (x) par rapport à celui canonique (y). La figure est adaptée de (Ataide et al., 2005).

## 2. Les aminoacyl-ARNt synthétases

Les aaRS, largement étudiées (Cusack, 1997 ; Martinis et al., 1999 ; Francklyn et al., 2002), constituent une famille de protéines fascinantes car elles occupent une place dans de nombreux de mécanismes. En effet, bien que leur fonction première soit l'aminoacylation des ARNt, ces enzymes ubiquitaires prouvent au fur et à mesure des découvertes leur rôle essentiel dans d'autres fonctions cellulaires telles que l'épissage, la modification d'ARN, la synthèse d'acide aminé ou la transduction du signal. Elles sont aussi considérées comme les premières protéines apparues, faisant le lien entre le monde des ARN et celui des protéines. L'analyse phylogénétique de leurs séquences sert de modèle pour la compréhension de l'évolution des macromolécules du vivant (Woese et al., 2000). Ces derniers temps, elles sont aussi l'objet de recherche visant à découvrir de nouvelles drogues contre des pathogènes résistants. Toutes ces particularités font de l'étude des aaRS un champ d'exploration immense.

Les aaRS sont des enzymes modulaires de tailles et de structures oligomériques très hétérogènes, qui assurent leur fonction d'aminoacylation grâce à deux types de sites catalytiques structuralement très différents. L'ensemble des données au niveau des séquences, des structures et de la biochimie a permis leur répartition en deux classes de 10 membres chacune et en six sous-classes (Cusack et al., 1990 ; Eriani et al., 1990 ; Arnez & Moras, 1997). Les principales caractéristiques spécifiques à chaque classe, décrites dans ce qui suit, sont également résumées dans le Tableau 1.

| | Classe I | | Classe II | |
|---|---|---|---|---|
| **Principales caractéristiques** | | | | |
| Motifs conservés | | | | |
| Liaison au substrat | HIGH | | Motif 2: …F**R**xE | |
| | KMSKS | | Motif 3: …(Gx)3E**R**… | |
| Dimérisation | | | Motif 1: …**P**… | |
| Repliement du site actif | Feuillet β parallèle (repliement de Rossman) | | Feuillet β anti-parallèle | |
| Conformation de l'ATP | Étendue | | Repliée | |
| Liaison à l'ARNt | | | | |
| Bras accepteur | Sillon mineur | | Sillon majeur | |
| Boucle variable | Face au solvant | | Face à la protéine | |
| Site d'aminoacylation | 2'OH | | 3'OH | |
| | **Sous-classe Ia** | | **Sous-classe IIa** | |
| | ValRS | α | SerRS | α2 |
| | IleRS | α | ThrRS | α2 |
| | LeuRS* | α | ProRS | α2 |
| | MetRS* | α2 | GlyRS* | α2β2 |
| | CysRS | α | HisRS | α2 |
| | **Sous-classe Ib** | | **Sous-classe IIb** | |
| | TyrRS | α2 | AspRS | α2 |
| | TrpRS | α2 | LysRS§ | α2 |
| | **Sous-classe Ic** | | AsnRS | α2 |
| | GluRS | α | **Sous classe IIc** | |
| | GlnRS | α | PheRS* | α2β2 |
| | ArgRS | α | AlaRS* | α4 |

**Tableau 1. Principales caractéristiques des aaRS et répartition selon leur classe.** La structure oligomérique est donnée pour *E. coli*. *Synthétases pour lesquelles la structure oligomérique n'est pas conservée au cours de l'évolution. §Une LysRS de classe I a été mise en évidence dans certains organismes (Ibba et al., 1997).

## 2.1. Enzymes de classe I

Les enzymes de classe I sont essentiellement monomériques et possèdent typiquement un site catalytique replié en domaine de Rossman, identifié comme un motif de liaison aux nucléotides (Rossmann et al., 1974). Ce repliement est basé sur une série de brins β parallèles reliés entre eux par des hélices α. Les aaRS de classe I sont caractérisées par la présence des deux séquences signatures HIGH (His-Ile-Gly-His) et KMSKS (Lys-Met-Ser-Lys-Ser) qui

font partie du site de liaison à l'ATP. Elles fixent préférentiellement l'acide aminé sur l'extrémité 2' OH de l'ARNt et interagissent au niveau du petit sillon de l'hélice du bras accepteur de l'ARNt, exposant ainsi la face de la région variable au solvant (Arnez & Moras, 1997). Cependant, les TyrRS bactériennes font figure d'exception en interagissant avec le grand sillon de l'hélice acceptrice de l'ARNt, tel que le font les aaRS de classe II (Fechter et al., 2000 ; Yaremchuk et al., 2002).

Des caractéristiques supplémentaires communes ont permis une subdivision de ces synthétases en sous-classes. Cette seconde classification se base sur des modules communs et la nature de l'acide aminé chargé. La première sous-classe, Ia, regroupe les aaRS chargeant les acides aminés aliphatiques et soufrés. La sous-classe Ib est composée des aaRS fixant des acides aminés aromatiques. La troisième, Ic, comprend les aaRS activant des acides aminés hydrophiles et nécessitant l'ARNt pour l'étape d'activation.

Le classement de l'ArgRS est problématique. Son domaine catalytique est proche de celui des enzymes du sous-groupe Ib, alors que son domaine de liaison de l'anticodon est similaire à celui des aaRS de la sous-classe Ia (Klipcan & Safro, 2004).

## 2.2. Enzymes de classe II

Les enzymes de classe II sont multimériques, essentiellement dimériques. Elles partagent trois motifs plus ou moins conservés, les motifs 1, 2 et 3. Le motif 1 participe à la dimérisation et est caractérisé par la présence d'une proline. Les motifs 2 et 3 font partie du site actif et sont caractérisés chacun par un résidu arginine dans un contexte précis. Le site catalytique se replie selon un feuillet β anti-parallèle entouré d'hélices α. Les aaRS de classe II fixent l'acide aminé préférentiellement sur le groupement 3'OH de l'ARNt et interagissent avec le grand sillon de l'hélice acceptrice de l'ARNt. Cependant, la PheRS fixe l'acide aminé d'après le mode des aaRS de classe I, sur le groupement 2' OH.

Tout comme les aaRS de classe I, les enzymes de classe II sont subdivisées en trois sous-classes (Arnez & Moras, 1997). La classe IIa comprend celles partageant des homologies de séquences dans leur partie C-terminale et chargeant des acides aminés petits et polaires. La sous-classe IIb comprend les aaRS ayant le domaine de liaison à l'anticodon en N-terminal et contenant une structure appelée OB (Oligonucleotide/Oligosaccharide Binding) fold. La troisième sous-classe, IIc, regroupe les aaRS tétramériques.

## 2.3. Organisation modulaire et motifs structuraux

De façon générale, les aaRS sont constituées d'un module catalytique supposé le plus ancien, auquel se seraient ajoutés au cours de l'évolution d'autres modules tels que celui de liaison à l'anticodon (voir §A.4.), permettant d'augmenter la spécificité d'interaction avec l'ARNt. De plus, certaines aaRS posssèdent des modules leur conférant des fonctions complémentaires (édition) ou totalement dissociées de l'aminoacylation (fonctions alternatives, définies plus loin §A.2.7).

À ce jour, il existe au moins une structure cristallographique pour chacune des 20 spécificités d'aaRS, libre ou en complexe avec certains de leurs substrats (acide aminé, ATP, ARNt, analogue de l'adénylate, etc…). Les structures cristallographiques et les alignements de séquences ont permis de mettre en évidence la grande complexité architecturale des aaRS. Cependant, un certain nombre de domaines sont conservés dans plusieurs aaRS de spécificités différentes (revue par (Wolf et al., 1999)).

Le domaine « DALR », désigné ainsi d'après un motif caractéristique d'acides aminés conservés (également nommé Add-2 dans l'ArgRS, (Cavarelli et al., 1998)) est le plus répandu dans la séquence des aaRS, après les motifs catalytiques. Composé d'hélices α, il est impliqué dans la reconnaissance de l'anticodon et est présent dans sept enzymes de classe I (Arg, Cys, Lys, Met, Val, Ile, Leu) et dans la sous unité β de la GlyRS bactérienne (enzyme de classe II).

Chez un certain nombre d'aaRS de mammifères de classes I (Trp, Met, Gln), de classe II (His, Gly), et dans l'enzyme bipartite GluProRS, on trouve un motif conservé structuré en hélice α « coiled-coil », unique ou répété. Il pourrait être impliqué dans la reconnaissance de l'ARNt ou dans la formation du complexe mutisynthétasique (Cavarelli et al., 1998).

Certaines aaRS assurent une fonction complémentaire à l'aminocylation, à savoir l'édition d'acides aminés incorrectement activés ou transférés à l'ARNt, grâce à la présence de domaine d'édition. Pour exemple, le domaine CP1 (« Connecting Peptide 1 ») est conservé parmi les enzymes de classe I LeuRS, ValRS et IleRS.

Un domaine homologue au domaine EMAP-II (« Endothelial and Monocyte-Activating Polypeptide II »), structuré en tonneau β et retrouvé dans les cytokines, est présent dans les TyrRS eucaryotiques, dans la sous unité β de la PheRS et dans la MetRS d'archaea et de quelques bactéries. Ce domaine a un rôle dans la régulation de l'angiogenèse en tant que stimulateur (voir plus loin §2.7).

Le domaine OB fold est présent dans les AspRS, LysRS-2 et AsnRS au niveau du module d'interaction avec l'anticodon de l'ARNt. Ce domaine a été nommé ainsi puisqu'il est retrouvé dans un certain nombre de protéines qui interagissent avec des oligonucléotides ou oligosaccharides (Murzin, 1993). Il se replie selon un feuillet β de 5 brins avec une hélice α insérée entre les brins 3 et 4. Le même repliement est également observé dans d'autres protéines liant l'ARN telles que le facteur d'initiation IF1 d'*E. coli* (Murzin, 1993) et la protéine SmpB d'*Aquifex aeolicus* qui interagit avec l'ARNtm (Dong et al., 2002).

## 2.4. Aminoacyl-ARNt synthétases particulières

**La LysRS de classe I.** La plupart des archae ne possède pas de gènes codant pour la LysRS connue de classe II, retrouvées chez les eucaryotes et la plupart des bactéries. Pourtant l'activité de lysinylation est détectée dans ces organismes. Cette activité est assurée par une LysRS dont la séquence comporte le domaine de Rossman, caractéristique du site catalytique des enzymes de classe I (Ibba et al., 1997). L'appartenance de la LysRS aux deux classes ébranle le concept de conservation de la classification au cours de l'évolution du vivant, respecté par ailleurs par toutes les autre aaRS connues à ce jour. Notons que dans certains organismes, les deux types de LysRS coexistent (Polycarpo et al., 2004 ; Ataide et al., 2005).

**La PylRS.** La pyrrolysine (Pyl) est définie comme étant le 22$^{ème}$ acide aminé impliqué dans la synthèse protéique (Atkins & Gesteland, 2002). Il est incorporé dans certains gènes d'archaea méthanogènes en réponse au codon stop UAG. La pyrrolysine a été découverte pour la première fois dans le site actif de la méthylamine méthyltransférase de *Methanosarcina barkeri*. Son incorporation serait étendue à toutes les enzymes de ce type (Burke et al., 1998 ; Paul et al., 2000) et s'accompagne de la présence de gènes *pylT*, codant pour un ARNt$^{Pyl}$, et *pylS*. Des études récentes ont montré que les produits de ces deux gènes suffisent à l'incorporation de Pyl dans la synthèse protéique chez *E. coli*, lorsque le milieu est supplémenté avec cet acide aminé (Blight et al., 2004). Il a aussi été montré que le produit du gène *pylS* est capable d'activer et de transférer directement la Pyl sur l'ARNt$^{Pyl}$ (Blight et al., 2004 ; Polycarpo et al., 2004) et serait donc une nouvelle aaRS, la PylRS. Les travaux de Polycarpo et al., 2004 montrent également qu'un mécanisme indirect de lysynylation de

l'ARNt$^{Pyl}$ par les deux LysRS (de classe I et II) complexées, coexisterait dans les cellules de *M. barkeri* : le résidu Lys chargé sur l'ARNt$^{Pyl}$ serait ensuite converti en Pyl.

**La SepRS.** Dans un certain nombre d'archaebactéries méthanogènes, le gène codant pour la CysRS est absent. Une nouvelle aaRS de classe II, la *O*-phosphoséryl-ARNt synthétase (SepRS), a été découverte (Sauerwald et al., 2005). Elle est impliquée dans la formation de Cys-ARNt$^{Cys}$ en deux étapes. La SepRS transfère de façon spécifique l'*O*-phosphosérine sur l'ARNt$^{Cys}$, qui est ensuite convertie en Cys-ARNt$^{Cys}$ grâce à la synthétase SepCysS. Ce mécanisme rappelle celui impliqué dans la formation de Sec-ARNt$^{Sec}$ (Hubert et al., 1996) ou les voies indirectes d'asparaginylation et de glutaminylation (Schön & Söll, 1988 ; Curnow et al., 1996 ; Becker & Kern, 1998) qui ne seront pas détaillées ici.

## 2.5. Complexe multisynthétasique

Chez la plupart des animaux, de la drosophile à l'Homme, on retrouve un complexe macromoléculaire composé de neuf aaRS (IleRS, LeuRS, MetRS, GlnRS, ArgRS, LysRS, AspRS et l'enzyme bipartite GluProRS) et de trois co-facteurs non enzymatiques (p18, p38 et p43) (revue par exemple par (Lee et al., 2004 ; Park et al., 2005b). Des études de microscopie électronique récentes (Wolfe et al., 2005) confirment le modèle selon lequel le complexe serait subdivisé en 3 domaines formant un Y, le facteur p43 se trouvant au centre du complexe. A la base du Y se retrouvent les GluProRS, IleRS et LeuRS, et dans chacune des branches soit les AspRS, MetRS et GlnRS, soit les LysRS et ArgRS.

### Les facteurs non-enzymatiques

Le facteur p43 occupe une place centrale dans le complexe. Homologue du facteur Arc1p de levure, qui se lie à la GluRS et la MetRS et stimule l'activité de la MetRS (Simos et al., 1998 ; Deinert et al., 2001), il pourrait jouer le rôle de donneur d'ARNt pour les aaRS (Norcum & Warrington, 2000). Le domaine en partie C-terminale liant les ARNt est identique à la cytokine EMAPII, libérée en conditions d'apoptose (Quevillon et al., 1997). Une étude *in vitro* a montré que le clivage de p43 par la caspase 7 libère ce domaine EMAPII. Le peptide obtenu est capable d'induire la migration de phagocytes mononucléaires humains (Shalak et

al., 2001). Ainsi, p43 aurait un rôle de commutateur induisant la croissance ou la mort cellulaire, selon les conditions.

La protéine p38 a essentiellement un rôle structural au sein du complexe multisynthétasique. Sa déplétion suffit à la désintégration du complexe (Kim et al., 2002). Cependant, elle est impliquée dans d'autres fonctions cellulaires. Par exemple, p38 est un substrat de parkine (ou E3 ubiquitine-ligase) impliquée dans l'ubiquitination et la dégradation par le protéasome des protéines cibles (Corti et al., 2003).

Le facteur p18 est la plus petite protéine du complexe. Son rôle n'est pas très bien connu, mais sa similarité de séquence avec les sous-unités β et γ du facteur d'élongation EF-1H (Quevillon & Mirande, 1996) suggère qu'il pourrait être impliqué dans la connexion de l'aminoacylation à la machinerie traductionnelle. Il a aussi été montré que ce facteur serait impliqué dans des processus de réparation de l'ADN (Park et al., 2005a).

### Rôles putatifs du complexe multisynthétasique

Les raisons précises de la formation d'un complexe stable multisynthétasique dans le cytoplasme des cellules d'eucaryotes supérieurs restent mystérieuses. Des études biochimiques *in vitro* ont montré que l'association des aaRS en complexes n'est pas essentiel pour leur fonction catalytique (par exemple l'ArgRS de hamster ou l'IleRS humaine sont actives isolées, (Shiba et al., 1994b; Guigou et al., 2004 )). Cependant la conservation de ce complexe chez les animaux multicellulaires suggère une fonction essentielle. Une des hypothèses majeures est que l'association des enzymes, autour de facteurs, permettrait une traduction efficace grâce à un flux d'ARNt contrôlé et localisé. Par exemple, le facteur p43 capable d'interagir avec les ARNt de façon aspécifique pourrait stimuler l'aminoacylation en stabilisant le complexe aaRS/ARNt. Dans ce sens, l'interaction de p43 stimulerait l'activité de l'ArgRS (Park et al., 1999). Cependant, une étude plus récente infirme ces résultats (Guigou et al., 2004). Le rôle stimulateur de p43 n'est toutefois pas exclu pour d'autres enzymes. Une récente étude théorique suggère que l'association des aaRS permettrait de mieux diriger les aminoacyl-ARNt vers le ribosome (Wolfson & Knight, 2005). Il avait été montré que l'affinité de l'interaction entre l'aminoacyl-ARNt et le facteur d'élongation bactérien EF-Tu dépend directement de la nature de l'acide aminé et de la séquence de l'ARNt (LaRiviere et al., 2001). L'association des aaRS avec le facteur p43 pourrait compenser des affinités réduites des aminoacyl-ARNt pour le facteur d'élongation eucaryotique eEF-1A (Wolfson & Knight, 2005).

Toutefois, l'implication des facteurs non-enzymatiques et des aaRS dans des processus cellulaires variés, suggère un autre rôle pour l'association des différentes protéines. Ainsi, le complexe multisynthétasique pourrait être un réservoir contrôlant les fonctions non canoniques de ses composants.

## 2.6. Mécanisme d'édition

Dans certains cas, le module catalytique des aaRS ne suffit pas à discriminer correctement entre des acides aminés proches ou plus petits. L'acquisition de module additionnel, dit de correction ou d'édition, a permis aux aaRS bénéficiaires d'éliminer les acides aminés non spécifiques (Jakubowski & Goldman, 1992). Notons, que l'édition ici n'est pas à confondre avec l'édition des ARN consistant à modifier la séquence codée. Les aaRS de classe I IleRS, ValRS et LeuRS possèdent toutes les trois le même domaine CP1 (« connecting peptide 1 »), inséré dans le domaine de Rossman, et qui leur permet d'éliminer respectivement la valine, la thréonine et d'autres précurseurs comme la norvaline (Nureki et al., 1998). Les enzymes de classe II AlaRS et ThrRS possèdent un domaine additionnel au cœur catalytique, distinct du domaine CP1 des enzymes de classe I. Il permet d'éliminer la sérine et la glycine dans le cas de l'AlaRS (Tsui & Fersht, 1981 ; Beebe et al., 2003) et la sérine en ce qui concerne la ThrRS (Dock-Bregeon et al., 2004). Un domaine d'édition supplémentaire, distinct des deux autres, a été décrit chez la ProRS de type bactérien et n'est pas retrouvé dans les ProRS de type eucaryotiques. Inséré dans le cœur catalytique, il permet l'élimination de l'alanine (Beuning & Musier-Forsyth, 2000 ; Wong et al., 2002). De façon intéressante, des protéines homologues aux domaines d'édition des ProRS et AlaRS sont capables de réaliser la même fonction, mais de façon autonome (Ahel et al., 2003 ; An & Musier-Forsyth, 2004).

## 2.7. Fonctions alternatives

Parmi les différentes fonctions alternatives associées aux aaRS (Martinis et al., 1999 ; Francklyn et al., 2002), j'ai choisi d'en développer 4 plus particulièrement.

**Activité cytokine.** Les TrpRS et TyrRS humaines jouent un rôle dans la régulation de l'angiogenèse, respectivement en tant qu'inhibiteur et stimulateur (revue par (Yang et al., 2004)). Les deux aaRS acquièrent leur nouvelle fonction après clivage protéolytique, libérant ainsi des fragments actifs. Le fragment actif de la TrpRS, appelé mini-TrpRS, est tronqué dans sa partie N-terminale et est obtenu soit par clivage protéolytique soit par épissage alternatif. Les deux fragments actifs de la TyrRS sont la mini-TyrRS (partie N-terminale) et C-TyrRS (partie C-terminale). Alors que la mini-TyrRS possède une séquence ELR spécifique de certaines interleukines, la C-TyrRS possède un domaine cytokine homologue au motif EMAP II. De récents travaux ont montré l'implication de la LysRS humaine dans l'induction de la réponse immunitaire via l'activation des monocytes et macrophages (Park et al., 2005c). Sécrétée des cellules suite à l'induction par la cytokine TNF-α (« Tumor Necrosis Factor »), la LysRS se lie aux monocytes et aux macrophages de façon à augmenter leur migration et la production de TNF-α. Alors que les TyrRS et TrpRS possèdent une activité cytokine sous une forme tronquée, c'est la LysRS dans son ensemble qui assure cette fonction additionnelle.

**Epissage.** Les TyrRS mitochondriales de *Neurospora crassa* et *Podospora anserina*, et la LeuRS mitochondriale de *S. cerevisiae* interviennent dans l'épissage auto-catalytique des introns de groupe I. Les aaRS facilitent le repliement des ARN sous la forme active. Alors que seul le domaine d'insertion CP1 de la LeuRS est responsable de la liaison à l'intron (Rho et al., 2002), la TyrRS interagit en tant que telle, par un mode d'interaction différent de celui avec l'ARNt (Paukstelis et al., 2005).

**Régulation.** La ThrRS d'*E. coli* régule l'expression de son propre gène (Romby et al., 1996 ; Caillet et al., 2003). Elle est capable d'interagir avec l'opérateur de son propre ARNm dont le repliement et la séquence miment deux branches anticodon d'ARNt$^{Thr}$. Cette interaction empêche l'accessibilité par le ribosome et ainsi la traduction de l'ARNm. Un autre exemple concerne l'AspRS de levure. qui est aussi capable de réguler l'expression de son propre ARNm (Frugier et al., 2005). Les auteurs proposent que le surplus d'AspRS, non séquestré par l'ARNt$^{Asp}$, est importé dans le noyau où l'enzyme interagit avec son propre messager et en inhibe l'accumulation. L'enzyme bipartite GluProRS est impliquée dans la répression de la traduction (Sampath et al., 2004). La forme fusionnée des deux aaRS n'est retrouvée que chez les eucaryotes supérieurs (Fett & Knippers, 1991) et fait partie du

complexe multisynthétasique (voir §2.5). Suite à la phosphorylation par l'interféron γ, la GluProRS quitte le complexe pour s'associer à 4 autres protéines. Ce nouveau complexe est capable d'interagir avec la partie 3' non traduite de l'ARNm de la ceruloplasmine, impliquée dans la réponse inflammatoire, et d'inhiber ainsi sa traduction.

**Implication dans le cycle viral.** Durant l'assemblage du Virus de l'Imunodéficience Humaine 1 (VIH-1), les $ARNt^{Lys}$ (3 isoaccepteurs parmi lesquels l'$ARNt_3^{Lys}$ sert d'amorce à la rétrotranscription virale) et la LysRS de l'hôte sont sélectivement empaquetés dans les virions (Jiang et al., 1993). La LysRS aurait un rôle dans l'encapsidation des $ARNt^{Lys}$ et dans l'efficacité d'infection du virus. L'assemblage des virions est entre autres dépendant du clivage par une protéase des précurseurs des protéines virales Gag et Gag-Pol. Il est donc essentiel de prévenir la maturation prématurée de Gag-Pol. La LysRS aurait un rôle dans l'inhibition de la protéase en empêchant l'association dimérique de Gag-Pol qui est responsable de l'activité protéolytique (Guo et al., 2005).

## 2.8. Paralogues d'aminoacyl-ARNt synthétases

Les informations obtenues à partir des structures cristallographiques et des alignements de séquences ont permis de mettre en évidence un grand nombre de protéines possédant des domaines similaires à ceux constituant les aaRS et dont la fonction est autre que l'aminoacylation. Ces protéines nommées « paralogues » seraient apparues par duplication du gène codant pour une aaRS, puis par divergence d'une des copies. Ceci témoigne également de l'émergence des aaRS par association de modules, qui auraient par la même occasion été dispersés plus largement dans le génome pour donner lieu aux protéines paralogues (Schimmel & Ribas de Pouplana, 2000).

Il existe des protéines homologues aux domaines d'édition de la ProRS et de l'AlaRS qui sont capables de corriger le transfert incorrect d'un acide aminé sur l'ARNt, mais de façon autonome (Ahel et al., 2003 ; An & Musier-Forsyth, 2004). D'autres protéines, impliquées dans la biogenèse des acides aminés possèdent des empreintes d'aaRS. Citons par exemple l'AsnA ou de l'HisZ qui sont similaires au cœur catalytique de l'AspRS et de l'AsnRS (Nakatsu et al., 1998 ; Roy et al., 2003) ou de l'HisRS (Sissler et al., 1999), respectivement. Ainsi, les mécanismes ancestraux de la biosynthèse des acides aminés auraient été dépendants

des ARNt. L'évolution des ancêtres des aaRS aurait donné lieu aux enzymes contemporaines et à des voies de biosynthèse des acides aminés ARNt-indépendantes impliquant les paralogues d'aaRS. Pour dernier exemple, citons la protéine YadB, renommée glutamyl-Q $ARNt^{Asp}$ synthétase d'après sa fonction, qui est fortement similaire au cœur catalytique de la GluRS mais ne présente pas de domaine de liaison à l'anticodon. Cette protéine est responsable de la modification de l'$ARNt^{Asp}$ par l'ajout d'un résidu Gln en position 34 de la boucle anticodon (Blaise et al., 2004 ; Dubois et al., 2004 ; Salazar et al., 2004).

## 2.9. Les aminoacyl-ARNt synthétases comme cibles thérapeutiques

Face à l'émergence de bactéries pathogènes multi-résistantes aux drogues (qui à ce jour ciblent principalement des mécanismes n'impliquant pas les aaRS), la découverte de nouvelles cibles thérapeutiques est devenue une priorité. Les aaRS sont des cibles potentielles de choix du fait de leur diversité structurale et notamment de leurs particularités espèces-spécifiques, exploitables pour le « design » d'inhibiteurs sélectifs. De façon naturelle, il existe un grand nombre d'inhibiteurs des aaRS. Ces composés peuvent être utilisés directement comme antibiotique ou comme modèle pour la création de nouveaux composés dérivés plus sélectifs et plus actifs.

Un des composés naturels les plus connus et utilisés cliniquement est la mupirocine ou acide pseudomonique. Elle est produite naturellement par *Pseudomonas fluorescens* et inhibe sélectivement l'IleRS de bactéries, notamment pathogènes (Kim et al., 2003). Cette drogue se lie au domaine de Rossman et inhibe ainsi la formation de l'adénylate (Yanagisawa et al., 1994). Depuis 1985, la mupirocine est largement utilisée contre les infections par les staphylocoques, notamment contre la souche de *S. aureus* résistante à la methicilline, et par les streptocoques. Malheureusement, différentes études montrent l'émergence de souches plus ou moins résistantes (Cookson, 1998 ; Schmitz et al., 1998 ; Yun et al., 2003). Alors qu'une résistance faible est liée à l'altération des résidus proches de la séquence signature KMSKS du domaine de Rossman, une plus forte résistance résulte de l'acquisition d'un gène supplémentaire d'une IleRS exogène. Cette dernière ne partage que 30% de similarité de séquence avec l'enzyme endogène (Yun et al., 2003).

Il existe beaucoup d'autres substances naturelles capables d'inhiber l'activité des aaRS telles que par exemple la borrelidine (inhibition de la ThrRS), la furanomycine (IleRS), la granaticine (LeuRS), l'indolmycine (TrpRS), l'ochratoxine A (PheRS), ou la cispentacine (ProRS) (Kim et al., 2003). Même si aucune de ces drogues n'a été exploitée en tant qu'antibiotique, elles illustrent la diversité de molécules naturelles inhibitrices des aaRS.

En plus des composés naturels et de leurs dérivés, de nouveaux inhibiteurs des aaRS sont en voie d'exploration. Leur découverte et leur développement sont et seront largement facilités par l'accumulation des connaissances fonctionnelles et structurales des systèmes d'aminoacylation, notamment de structures de synthétases complexées à des agents inhibiteurs, et à l'utilisation de nouvelles techniques telles que le « design » de drogues assisté par ordinateur et le criblage de chimiothèque à haut débit.

## 3. Le second code génétique

La fidélité de la traduction repose sur deux associations successives : celle de l'acide aminé et de l'ARNt, et celle de l'interaction de l'ARNt avec l'ARNm au niveau du ribosome. La correspondance entre le codon porté par l'ARNm et l'anticodon de l'ARNt, associé à un acide aminé donné, est dirigée par un langage nucléique et défini le code génétique. En revanche, l'association de l'ARNt à l'acide aminé est indirecte et utilise deux langages, nucléique et acide aminé. Par analogie et en raison de l'importance de la nature de l'acide aminé porté par l'ARNt dans la définition du code génétique, l'ensemble des signaux permettant la spécificité d'aminoacylation a été définie comme le second code génétique (de Duve, 1988).

Dans la cellule, une aaRS donnée doit reconnaître l'ARNt ou les isoaccepteurs qui lui sont spécifiques parmi une quarantaine d'ARNt différents, tous structuralement très proches. Il a été montré que les ARNt interagissent aussi bien avec les aaRS spécifiques que non spécifiques (Ebel et al., 1973). La formation de couples ARNt/aaRS est dynamique et l'aminoacylation spécifique est le résultat net entre les interactions productives ARNt/aaRS homologues et les interactions non productives ARNt/aaRS non homologues. La fidélité de l'aminoacylation des ARNt repose principalement sur l'efficacité catalytique des aaRS pour leurs substrats spécifiques et surtout est gouvernée par la constante catalytique $k_{cat}$ de la réaction. On parle ainsi de spécificité cinétique (Ebel et al., 1973 ; Giegé et al., 1998).

L'identité a été largement explorée tant au sein des systèmes en tant que tels (couples aaRS/ARNt naturels) qu'en utilisant des ARNt disséqués voire des systèmes artificiels (aussi appelés systèmes orthologues) capables d'incorporer des acides aminés originaux.

## 3.1. Identité des ARNt

Mieux connus pour l'ARNt que pour les synthétases, les éléments d'identité résultent de la combinaison d'éléments positifs (déterminants), qui permettent la reconnaissance et l'aminoacylation par l'aaRS spécifique, et d'éléments négatifs (antidéterminants) qui empêchent toute aminoacylation par une synthétase non spécifique. Lorsque l'ARNt possède une identité ambiguë ou en partie redondante avec un autre ARNt, il a été montré que l'aminoacylation correcte dépend aussi de la compétition des différentes aaRS dans la cellule (Sherman et al., 1992a; Sherman et al., 1992b ).

Les éléments majeurs définissant l'identité de tous les ARNt d'*E. coli* ont été identifiés et la plupart sont connus pour les ARNt de *S. cerevisiae*. L'identité a aussi été explorée dans un certain nombre d'autres organismes tels que *T. thermophilus*, *C. elegans*, *A. thaliana*, *H. sapiens*, etc. La connaissance de tous ces systèmes a permis de dégager un certain nombre de règles générales (revues par (McClain, 1995 ; Giegé et al., 1998 ; Beuning & Musier-Forsyth, 1999 ; Giegé & Frugier, 2003). Ainsi, l'expression de l'identité d'un ARNt donné se fait par un nombre restreint d'éléments, préférentiellement localisés aux deux extrémités de la structure tridimensionnelle en L de l'ARNt (l'anticodon et l'extrémité acceptrice) (Figure 4). Un élément d'identité isolé peut être un nucléotide (base discriminatrice G73 de l'ARNt$^{Asp}$, *e.g.* (Pütz et al., 1991) et revue par (Giegé et al., 1996)), une paire de bases (G3 :U70 de l'ARNt$^{Ala}$, (Hou & Schimmel, 1988 ; Francklyn & Schimmel, 1989)), un élément structural (la paire de Levitt atypique G15 :G48 de l'ARNt$^{Cys}$ d'*E. coli* (Hou et al., 1993 ; Hamann & Hou, 1997)), la région variable de l'ARNt$^{Ser}$ (Lenhard et al., 1999)) ou une modification post-transcriptionnelle (Lysidine de l'ARNt$^{Ile}$ (Muramatsu et al., 1988 ; Ikeuchi Y, 2005)). Il peut agir soit par contact direct avec des acides aminés précis de l'aaRS ou de manière indirecte en contribuant au bon positionnement de l'ARNt sur l'enzyme (*e.g.* G10-U25 de l'ARNt$^{Asp}$ de levure (Pütz et al., 1991)). Il a été montré que les éléments d'identité communiquent entre eux et qu'ils peuvent interagir de façon coopérative, additive ou anticoopérative (Pütz et al., 1993). L'importance de la structure de l'ARNt pour sa reconnaissance par les aaRS n'est pas

négligeable. En effet, un certain nombre d'études ont montré que l'architecture globale de l'ARNt est nécessaire à la présentation optimale des éléments d'identité et que le squelette sucre-phosphate de l'ARN contribue à la spécificité de l'interaction ARNt/aaRS. Toutefois, il est possible de modifier le cœur structural de l'ARNt sans abolir ses propriétés catalytiques, à condition que les éléments d'identité restent correctement présentés.

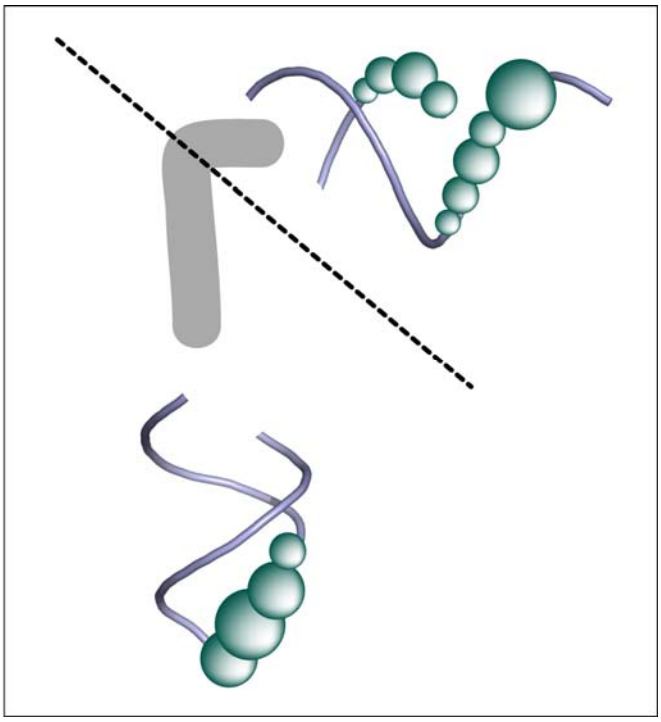


**Figure 4. Localisation des éléments majeurs d'identité aux deux extrémités de la structure 3D de l'ARNt.** Les éléments d'identité sont représentés par des sphères dont la taille est proportionnelle à la fréquence d'utilisation de la position en tant qu'élément d'identité. Le trait en pointillé symbolise l'architecture symétrique de l'ARNt qui se divise en deux domaines structuraux indépendants. Figure adaptée de (Giegé & Frugier, 2003).

### *Approches expérimentales*

De façon générale, l'étude de l'identité d'un ARNt consiste à créer des mutants à des positions d'intérêt et d'observer la réactivité de ces mutants. Ainsi un déterminant d'identité sera un élément qui, suite à sa mutation, entraîne une perte d'activité. Selon l'effet plus ou moins fort de la mutation, un élément sera qualifié de majeur ou mineur. Différentes approches permettant de définir les éléments d'identité, tant *in vivo* que *in vitro*, ont été développées.

Les méthodes *in vivo* consistent à exprimer des mutants de gènes d'ARNt soit suppresseur du codon ambre soit initiateur et dont l'anticodon a été modifié. La capacité de

ces ARNt à lire les codons correspondant à leur anticodon dans un gène rapporteur est déterminée et l'acide aminé mis en place est identifié par séquençage de la protéine rapporteur (Normanly et al., 1986 ; Normanly et al., 1990). Une alternative consiste à construire des souches nulles (« knockout ») pour un ARNt donné. Les gènes d'ARNt essentiels peuvent ainsi être utilisés comme outil génétique pour la sélection de mutants (pour des études chez *E. coli* voir (Gabriel & McClain, 1999 ; McClain & Gabriel, 2001); pour des études chez la levure voir (Geslain et al., 2003). L'avantage des méthodes *in vivo* est que l'effet des mutations de l'ARNt est observé dans des conditions physiologiques. Cependant, un effet délétère peut être le résultat de plusieurs évènements à différents stades de la vie de l'ARNt. Ainsi, les mutations peuvent non seulement affecter l'aminoacylation mais aussi l'expression, la maturation et la modification de l'ARNt, ainsi que l'interaction de l'ARNt avec le reste de la machinerie traductionnelle.

Les méthodes *in vitro* consistent à analyser l'aminoacylation d'un ARNt purifié extrait de cellules ou transcrit *in vitro* par l'ARN polymérase du bactériophage T7 (Sampson & Uhlenbeck, 1988), en présence d'aaRS, d'un extrait brut ou purifié, et de l'acide aminé correspondant. L'avantage de cette méthode est qu'elle permet d'analyser la seule réaction d'aminoacylation et ainsi de déterminer les paramètres cinétiques $k_{cat}$ et $K_M$ et de définir une efficacité catalytique associée à chaque couple enzyme/substrat. Ainsi on peut déterminer l'efficacité de l'aaRS pour un substrat de séquence sauvage ou mutée. La synthèse *in vitro* d'ARN permet en effet de préparer de nombreux variants, voir des fragments d'ARNt. Les conditions ne reflètent pas la compétition intracellulaire, et les ARNt transcrits *in vitro* ne sont pas modifiés. Cependant, les transcrits, sauf exception ($ARNt^{Glu}$ pour lequel la modification du U34 est déterminant d'identité) sont généralement de bons substrats pour leur synthétase, même s'ils ne sont pas modifiés.

### *Transplantation*

Par définition, le jeu d'identité d'un ARNt peut être transplanté dans un autre ARNt et va lui conférer la spécificité de l'ARNt donneur (greffe du jeu d'identité), à condition qu'il n'y ait pas d'éléments négatifs présents dans la séquence ou la structure de l'ARNt hôte. Les expériences de transplantation ont largement été utilisées pour confirmer qu'un jeu d'identité était complet. Ainsi, le jeu d'identité définissant la spécificité aspartate chez la levure a été vérifiée par transplantation de ce jeu dans l'$ARNt^{Phe}$ (Pütz et al., 1991). L'ARNt obtenu devient un substrat efficace de l'AspRS. Cependant la transplantation dans un autre ARNt, l'$ARNt^{Gln}$, a permis de mettre en évidence l'implication de l'élément C38 et ainsi de

compléter le jeu (Frugier et al., 1994b). Les expériences de transplantation ont aussi montré qu'un petit nombre d'éléments suffit pour l'aminoacylation par une aaRS donnée puisque la seule transplantation de la paire de base G3•U70, élément majeur d'identité alanine, confère des capacités d'alanylation à d'autres ARNt et même à des substrats minimalistes minihélices (voir plus loin) (Francklyn & Schimmel, 1989). D'autres expériences ont permis de montrer que des structures fines de l'ARNt sont importantes pour la présentation optimale des éléments d'identité. Ainsi, l'ARNt$^{Asp}$ ne devient un substrat efficace de la PheRS qu'après transplantation des régions $\alpha$ et $\beta$ de la boucle D et de la région variable de l'ARNt$^{Phe}$ donneur (Perret et al., 1992). Ceci étaye la notion selon laquelle la présentation optimale des éléments d'identité à l'aaRS exige une architecture précise. Cependant, la présence de jeu partiel d'identité dans des structures alternatives telles que les structures pseudo-ARNt naturelles (voir §A.1.4.) ou les subtrats minimalistes créés artificiellement (par exemple les minihélices, voir plus loin) montrant qu'une certaine flexibilité structurale est permise. Néanmoins l'optimum réside dans la structure de l'ARNt natif.

### *Localisation des éléments d'identité*

Les éléments d'identité sont surtout présents aux deux extrémités de l'ARNt, au niveau de la boucle anticodon et dans le bras accepteur (Figure 4). Ainsi excepté les ARNt$^{Ser}$, ARNt$^{Ala}$ et ARNt$^{Leu}$ d'*E. coli*, au moins un des nucléotides de l'anticodon joue un rôle de déterminant de l'identité (Giegé & Frugier, 2003). L'absence d'information dans l'anticodon peut s'expliquer dans le cas des spécificités serine et leucine par la présence de plusieurs isoaccepteurs, et est compensée dans le cas des ARNt$^{Ser}$ par la présence d'une grande région variable impliquée dans l'identité. En ce qui concerne le bras accepteur, les premières paires de bases sont impliquées dans plus de la moitié des identités d'*E. coli* et, excepté les identités pour les ARNt$^{Thr}$ et ARNt$^{Glu}$, la base discriminatrice N73 est un déterminant. Notons le cas particulier de l'ARNt$^{His}$ qui possède un nucléotide en position -1 (G-1 codé chez la bactérie mais ajouté post-transcriptionnellement chez les eucaryotes (Gu et al., 2003)) qui est un déterminant d'identité (Himeno et al., 1989). Les éléments d'identité peuvent aussi être localisés au cœur de l'ARNt (citons par exemple la paire de Levitt 515•G48 de l'ARNt$^{Cys}$ d'*E. coli* (Hou et al., 1993 ; Hamann & Hou, 1997), la région variable de l'ARNt$^{Ser}$ (Lenhard et al., 1999) ou le résidu A20 de la boucle D de l'ARNt$^{Arg}$ d'*E. coli* (McClain & Foss, 1988 ; Tamura et al., 1992).

### *Identité ambiguë et antidéterminants*

Les éléments constituant un jeu d'identité peuvent se retrouver dans d'autres ARNt voir peuvent participer à une autre identité, laissant supposer l'existence de charge croisée entre les différents systèmes ou de la présence d'antidéterminants. Ces ambiguïtés ont notamment été révélées grâce à l'utilisation de transcrits d'ARNt. Ainsi, l'ARNt$^{Ile}$ d'*E. coli* non modifié est un substrat non seulement de l'IleRS mais aussi de la MetRS. Dans la cellule, cette erreur est empêchée par la présence de la modification post-transcriptionnelle lysidine en position 34 de l'ARNt$^{Ile}$ (Muramatsu et al., 1988). Un autre exemple concerne l'arginylation efficace de l'ARNt$^{Asp}$ de levure, transcrit *in vitro* (Perret et al., 1990). De façon originale, l'ARNt$^{Asp}$ de levure possède un jeu d'identité arginine propre, mais masqué par une modification post-transcriptionnelle dans la cellule (Pütz et al., 1994 ; Sissler et al., 1996). De plus, les ambigüités d'aminoacylation peuvent être révélées dans un contexte de compétition (similaire à ceui de la cellule). Par exemple, les ARNt$^{Glu}$ et ARNt$^{Asn}$, qui possèdent des éléments de l'anticodon communs à l'ARNt$^{Asp}$, sont aspartylés dans un extrait d'ARNt total de levure à des taux importants de 70 et 35 % respectivement, en présence d'un excès d'AspRS (Ryckelynck et al., 2003). Ces ambigüités naturelles ont suggéré des expériences originales d'ingénierie d'ARNt synthétiques possédant une spécificité d'aminoacylation multiple (*e.g.* l'ARNt$^{Asp}$ de levure peut être transformé en une molécule chimérique capable d'être aminoacylé par les 3 enzymes AlaRS, PheRS et ValRS (Frugier et al., 1993)).

Ainsi, la spécificité d'aminoacylation d'ARNt à identité ambigües doit être maintenue grâce à des éléments négatifs (antidéterminants) qui empêchent toute reconnaissance par une aaRS non homologue. Bien qu'il y ait encore peu d'antidéterminants connus à ce jour, il est probable que tous les ARNt en possèdent pour se protéger d'une ou plusieurs aaRS non spécifiques. Un antidéterminant peut agir directement en provoquant des contraintes stériques ou en masquant un déterminant, et peut être un nucléotide isolé (*e.g.* A73 de l'ARNt$^{Leu}$ humain contre la SerRS (Breitschopf et al., 1995)), une paire de bases (*e.g.* G3•U70 de l'ARNt$^{Ala}$ de levure contre la ThrRS (Nameki, 1995)), une modification post-transcriptionnelle *(e.g.* Lysidine de l'ARNt$^{Ile}$ d'*E. coli* contre la MetRS (Muramatsu et al., 1988)), ou encore un groupement chimique (*e.g.* méthyl en position N1 du résidu G37 de l'ARNt$^{Asp}$ de levure contre l'ArgRS (Pütz et al., 1994)). La structure de l'ARNt peut aussi restreindre la spécificité d'un ARNt. En effet, l'architecture d'un variant de l'ARNt$^{Asp}$ de levure, possédant les éléments d'identité Phe, empêche la reconnaisance optimale par la PheRS (Perret et al., 1992).

### *Subtilité de l'identité*

L'identité ne résulte pas uniquement de la combinaison de déterminants et d'antidéterminants. En effet, plusieurs études ont déjà montré que l'architecture de l'ARNt joue un rôle crucial dans la présentation optimale de ces éléments. Ainsi, la séquence du bras accepteur de l'ARNt$^{Asp}$ d'*E. coli* est telle que la conformation est idéale pour l'interaction au niveau du site catalytique de la synthétase (Choi et al., 2003). D'autres combinaisons de séquence sont tolérées tant que le squelette initial sucre-phosphate de l'ARNt est maintenu, montrant que ce n'est pas la séquence en tant que telle mais bien la conformation résultant de la combinaison qui est importante. La transplantation des éléments d'identité Phe dans l'ARNt$^{Asp}$ ont montré que l'ARNt$^{Asp}$ hôte ne devient un substrat efficace de la PheRS qu'après transplantation des régions $\alpha$ et $\beta$ de la boucle D et de la région variable de l'ARNt$^{Phe}$ donneur (Perret et al., 1992). Etayant le fait que l'identité requiert un environnement favorable, créé par la structure et/ou la combinaison de nucléotides, des expériences de transplantation de l'identité Phe de levure dans l'ARNt$^{Ala}$ d'*E. coli* a permis de définir un nouveau type d'éléments, les éléments permissifs (Frugier et al., 1998). En effet, l'aminoacylation de l'ARNt hôte par la PheRS est inhibée par la présence de la paire de bases G2-C71. Cet effet négatif peut être compensé par l'insertion de la paire de bases G3•U70, qui n'a pas d'effet en soi sur la phénylalanylation. Il en découle directement que l'identité résulte d'une combinaison précise de nucléotides non due au hasard. Enfin, des éléments d'identité peuvent êtres alternatifs ou cryptiques (Saks & Sampson, 1996). L'exemple de l'arginylation efficace de l'ARNt$^{Asp}$ transcrit *in vitro* grâce à un jeu alternatif d'identité de celui démontré dans l'ARNt$^{Arg}$ (C36 et G37 au lieu de C35 et G/U36, (Sissler et al., 1996)) montre que l'identité sélectionnée par la nature n'est pas forcément unique.

## 3.2. Identité des aminoacyl-ARNt synthétases

L'identité des aaRS concerne tant la reconnaissance de l'acide aminé à incorporer que celle de l'ARNt. Les mécanismes assurant la reconnaissance de l'acide aminé sont qualifiés de « double passoire ». Le premier est passif et exclu les acides aminés plus grands, alors que le deuxième est actif et retrouvé dans un nombre limité d'aaRS. Il s'agit de l'édition (ou élimination) des acides aminés activés non spécifiques ou activés et transférés à l'ARNt par erreur (voir §2.2.6.). En ce qui concerne les éléments permettant la reconnaissance spécifique

des ARNt, leur identification est compliquée dû à la taille importante des séquences des aaRS et l'hétérogénéité de leurs structures. Si les acides aminés qui interagissent avec les éléments d'identité des ARNt peuvent être localisés grâce aux données des structures cristallographiques des complexes aaRS/ARNt, ils agissent dans un contexte non négligeable et encore plus complexe que celui de l'ARNt. Néanmoins, une approche ingénieuse a permis de cartographier les résidus essentiels du site actif de l'AspRS (Ador, 1999) et de révéler deux domaines fonctionnels dans l'ArgRS (Geslain et al., 2000) de *S. cerevisiae*. La technique se base sur la sélection *in vivo* de mutations létales de l'aaRS, à partir d'une banque de gènes aléatoirement mutés et exprimés dans la levure. Par ailleurs, l'analyse *in vitro* de l'efficacité d'aminoacylation de mutants de l'AspRS de levure ont permis de montrer que les acides aminés en contact avec les éléments d'identité de l'anticodon de l'ARNt contribuent de façon prédominante à la stabilité du complexe (Eriani & Gangloff, 1999). Par analogie aux nucléotides d'identité de l'ARNt, on peut parler d'acides aminés d'identité.

D'autres expériences de mutagénèse ont montré qu'il est possible de changer la spécificité de reconnaissance de l'ARNt d'une aaRS donnée. Par exemple, il est possible de changer l'identité de la TyrRS d'*E. coli* (qui reconnaît la paire de bases G1-C72 de l'ARNt$^{Tyr}$ homologue) en celle de la TyrRS humaine (qui reconnaît la même paire de bases mais inversée dans son substrat) grâce à la transplantation d'un court peptide correspondant au domaine CP1, l'inverse étant également faisable (Wakasugi et al., 1998).

## 3.3. Interaction des ARNt/aaRS

Les complexes ARNt/aaRS des systèmes glutamine d'*E. coli* (enzyme de classe I) (Rould et al., 1989 ; Perona et al., 1993) et aspartate de levure (enzyme de classe II) (Ruff et al., 1991 ; Cavarelli et al., 1993) ont été les premiers dont la structure cristallographique a été résolue, et sont ainsi devenus des modèles quant à la reconnaissance de l'ARNt selon la classe d'enzyme (Figure 5). Pour les deux complexes, l'interaction de l'ARNt avec la synthétase conduit à des changements conformationnels importants au niveau de la boucle anticodon avec les nucléotides du triplet pointant vers la synthétase et formant des interactions très spécifiques. Dans le cas de l'ARNt$^{Gln}$, deux plateaux de bases supplémentaires se forment (32-38, 33-37) dans la boucle. La structure globale de l'ARNt$^{Asp}$ complexé subit de grands

changements avec un angle entre les deux branches de la molécule considérablement réduit (110° à 95°). Au niveau du bras accepteur, alors que l'interaction de l'ARNt$^{Asp}$ avec l'AspRS ne produit aucune modification drastique de la structure, celle de l'ARNt$^{Gln}$ avec la GlnRS crée une torsion du squelette responsable du recourbement de l'extrémité CCA et de l'ouverture des premières paires de bases. Ces différences structurales au niveau du bras accepteur reposent sur les modes distincts d'interaction des sites catalytiques des enzymes de classe I et II : au niveau du sillon mineur de l'ARNt$^{Gln}$ (classe I) et du sillon majeur de l'ARNt$^{Asp}$ (classe II). Cependant, les changements de conformation des ARNt, d'abord supposés conservés en fonction de la classe d'aaRS, apparaissent plutôt propres à chaque système. Ainsi, la première paire de bases de l'ARNt$^{Val}$ interagissant avec la ValRS (Liu et al., 1997) n'est pas défaite, ou la boucle anticodon de l'ARNt$^{Phe}$ ne subit pas de changement de conformation (Goldgur et al., 1997).

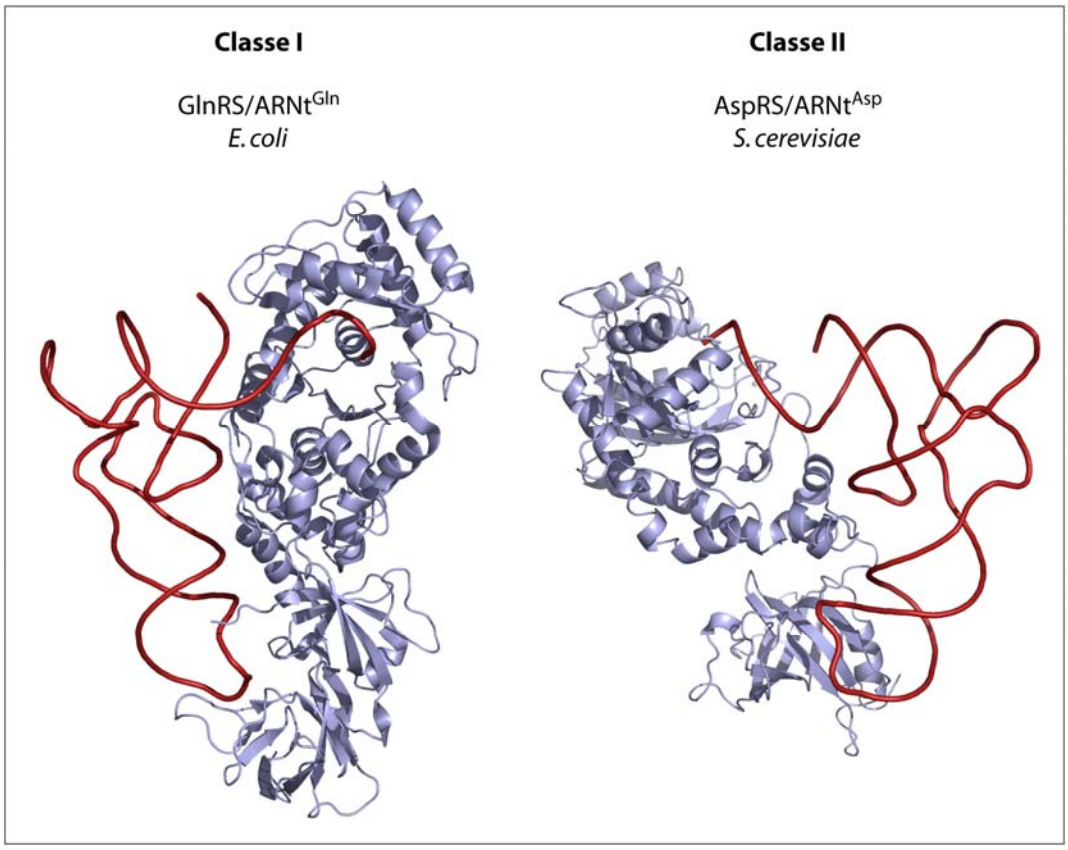



**Figure 5. Structures de complexes ARNt/aaRS de classes I et II.** Les coordonnées du complexe GlnRS/ARNt$^{Gln}$ d'*E. coli* sont de (Bullock et al., 2003) et celles du complexe AspRS/ARNt$^{Asp}$ de *S. cerevisiae* de (Cavarelli et al., 1994). Les aaRS (en bleu clair) sont sous forme schématique où hélices α et brins β sont visibles, les ARNt (en rouge) sont représentés en fil de fer. La figure a été générée en utilisant le programme PyMol (DeLano, 2002).

## 3.4. Aminoacylation de fragments d'ARN et systèmes d'aminoacylation artificiels

La nécessité de comprendre les mécanismes de reconnaissance entre aaRS et ARNt a motivé des expériences tout à fait originales. En effet, certains groupes ont poussé l'analyse jusqu'à la dissection proprement dite de l'ARNt, voire à la création de systèmes artificiels d'aminoacylation.

Le « design » de pseudo-ARNt artificiels peut se faire soit par mutagénèse classique et approche rationnelle, soit par SELEX (« Systematic Evolution of Ligands by Exponential Enrichment ») qui permet d'explorer toute une population de molécules à diversité combinatoire (revue par (Vörtler et al., 2001)). La Figure 6 illustre des exemples de structures obtenues à partir de l'ARNt$^{Asp}$ de levure (haut) ou du bras accepteur d'ARNt (minihélices) de spécificités variées (*e.g.* Ile, Met, Val, Leu, Asp, Ala, Gly, His et Ser, revue par (Schimmel et al., 1993)). Postulant le fait que les éléments majeurs d'identité sont essentiellement situés aux deux extrémités de l'ARNt, il devrait être possible de réorganiser le cœur de l'ARNt sans abolir ses propriétés d'aminoacylation, à condition que ces éléments soient encore correctement positionnés. Cette accomodation est au moins possible pour 3 dérivés de l'ARNt$^{Asp}$ de levure qui restent des substrats de l'AspRS (Wolfson et al., 1999) (Figure 6.B.). La première structure ressemble à l'ARNt$^{Ser}$ mitochondrial de mammifère dont le bras D est manquant ; la deuxième structure est délétée des bras D et T, avec les hélices acceptrice et anticodon jointes par deux connecteurs ; et la troisième consiste en la partie 3' du bras accepteur et le bras anticodon reliés par un connecteur. Un deuxième postulat consiste à admettre que l'ARNt primordial était formé du bras accepteur. Des substrats minimalistes dérivés du bras accepteur (Figure 6.C.), appelés minihélices (empilement des bras accepteur et T) ou microhélices (bras accepteur et boucle T), ont été testées et conservent des propriétés importantes d'aminoacylation. La méthode par approche combinatoire consiste à produire une banque d'ARN et à sélectionner de façon itérative les molécules d'intérêt. Ainsi, à partir d'une librairie de molécules dégénérées, dérivées de la séquence de l'ARNt$^{Phe}$ d'*E. coli*, plusieurs ARN interagissant fortement avec la PheRS ont pu être sélectionnés par des expériences de rétention sur filtre (Tinkle Peterson et al., 1994). De façon intéressante, toutes ces molécules dévient de la structure classique des ARNt, et présentent un bras anticodon de séquence atypique, mais conservant le triplet anticodon Phe. Cependant, aucune de ces molécules n'est aminoacylée par la PheRS.

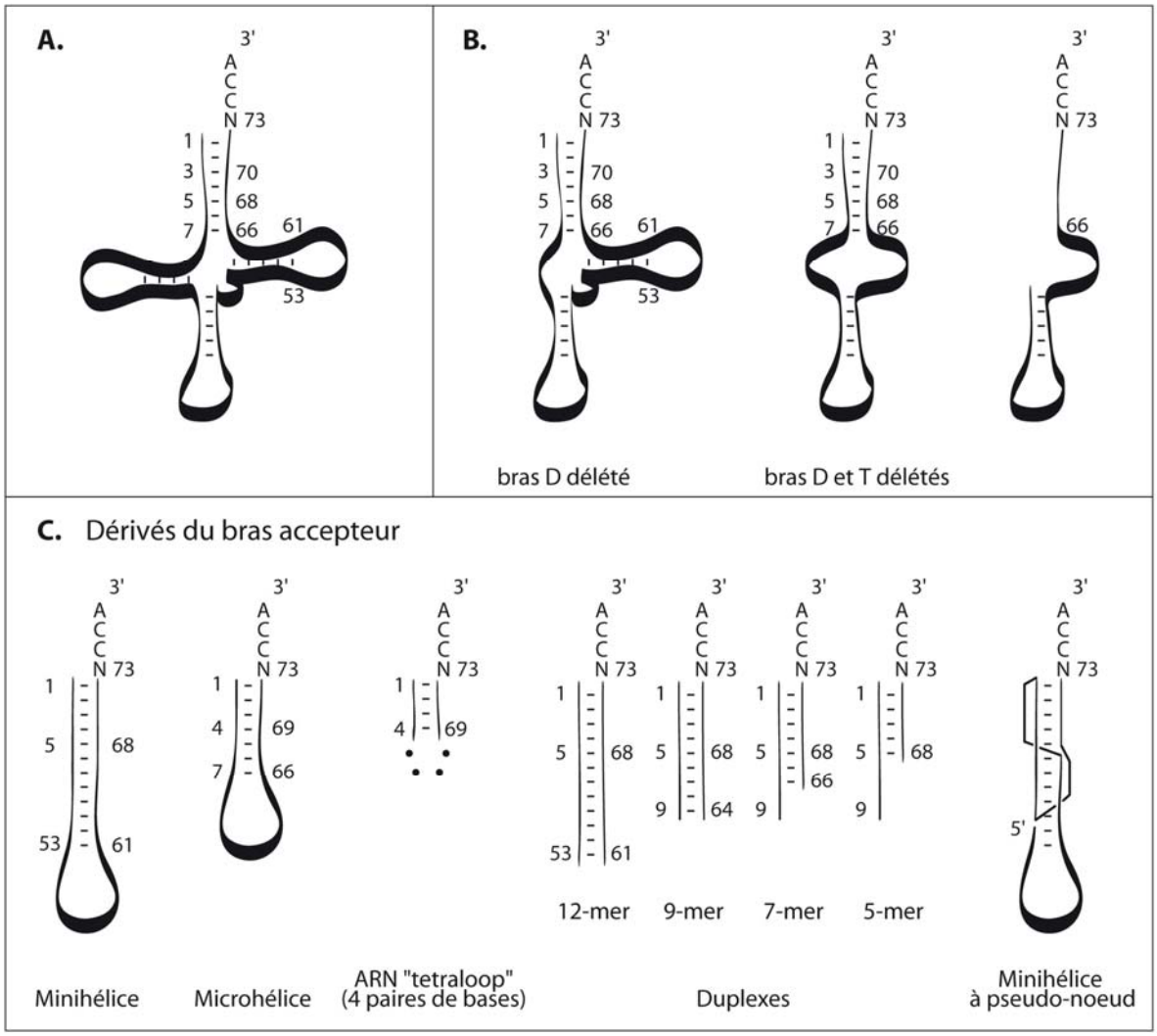



**Figure 6**. **Représentation schématique de différents dérivés d'ARNt aminoacylables. A.** ARNt classique en feuille de trèfle. **B.** Différents dérivés de l'ARNt$^{Asp}$ de levure aminoacylables par l'AspRS (Wolfson et al., 1999). **C.** ARNt minimalistes dérivés du bras accepteur d'ARNt de spécificités variées (revue par (Schimmel et al., 1993)).

Bien qu'il existe plus de 120 acides aminés retrouvés dans les protéines (Uy & Wold, 1977), seulement 20 sont codés par le génome, avec les exceptions de la sélénocystéine (Leinfelder et al., 1988) et de la pyrrolysine (Atkins & Gesteland, 2002). La synthèse de protéines repose sur 61 codons sens et 3 codons stop. Aujourd'hui, grâce aux progrès de la chimie et de la biologie moléculaire, il est possible d'étendre le code génétique et d'inclure des acides aminés artificiels de façon co-traductionnelle dans *E. coli*, la levure et des cellules de mammifères (Xie & Schultz, 2005). Pour ce faire, sont requis un codon qui ne désigne que l'acide aminé artificiel (souvent un codon stop ainsi réassigné), un nouvel ARNt spécifique de cet anticodon et une aaRS correspondante. De façon à incorporer l'acide aminé artificiel

uniquement au site spécifique, il faut que l'ARNt ne soit pas aminoacylé par les aaRS endogènes, et que l'aaRS ne reconnaisse que l'ARNt et l'acide aminé spécifique. On parle alors de système orthogonal, c'est à dire qui ne réagit pas avec les molécules analogues de l'hôte. Pour exemples, citons le cas des systèmes orthogonaux dérivés des spécificités tyrosine, le premier utilisé (Wang et al., 2001), ou glutamine (Köhrer C, 2004). Le système tyrosine orthogonal dérive du système de *Methanococcus jannaschii* et est exprimé chez *E. coli* pour incorporer un analogue de la tyrosine en réponse à un codon stop. Le système glutamine orthogonal dérive du système d'*E. coli* et est exprimé dans des cellules de mamifères afin d'incorporer dans la chaîne polypeptidique un analogue de la glutamine en réponse aux 3 codons stop possibles. D'autres laboratoires sont allés encore plus loin dans la démarche en proposant l'incorporation d'analogues d'acides aminés en réponse à un codon à 4 lettres, s'affranchissant ainsi du réassignement de codons préexistants (Moore et al., 2000 ; Hohsaka & Sisido, 2002 ; Sisido et al., 2005). L'incorporation de nouveaux acides aminés non seulement pousse la compréhension des systèmes d'aminoacylation et notamment des règles de spécificité au plus loin mais aussi permettra probablement de créer de nouvelles protéines avec des propriétés et structures inédites.

## 4. Aspects évolutifs

Ce qui suit a pour objectif de décrire les différentes hypothèses discutées actuellement sur l'origine et l'évolution du vivant et en particulier des systèmes d'aminoacylation.

L'évolution au niveau moléculaire ne laisse pas de traces. Il n'y a pas de fossiles de la plus ancienne forme de vie ou des molécules qui la constituaient. Cependant les molécules biologiques contemporaines fournissent des indices précieux. Ainsi, le **fossile moléculaire** a été défini comme étant une entité dont la structure et/ou la fonction contemporaine renseignent sur son histoire évolutive (Maizels & Weiner, 1994). Le « challenge » pour comprendre l'évolution au niveau moléculaire est de déduire des molécules contemporaines le chemin inverse jusqu'à leur forme ancienne.

## 4.1. Le paradigme actuel : de la soupe primordiale aux trois règnes du vivant

Parmi les différentes théories proposées pour l'apparition et l'évolution de la Vie, celle qui fait figure de paradigme suggère le schéma de la Figure 7. La vie serait apparue sur une terre primitive dans des conditions extrêmes, sans doute localement propices à des réactions prébiotiques. Ce type de réactions aurait généré par condensation les briques élémentaires du vivant (sucres, bases, acides aminés…), constituant de la soupe primordiale. De ce monde prébiotique serait né des polynucléotides aléatoires puis un monde entièrement composé et dirigé par des ARN (Gilbert, 1986). La théorie du monde des ARN a été imaginée suite à la découverte au début des années 80 des premiers ARN catalytiques (intron de groupe I d'un ARNr de *Tetrahymena* capable d'auto-épissage (Kruger et al., 1982), activité catalytique portée par la sous-unité ARN de la RNAse P (Guerrier-Takada et al., 1983), activité peptidyl-transferase de l'ARNr 23S de la sous-unité 50S du ribosome de *Thermophilus aquaticus* (Noller et al., 1992)), faisant des ARN des biomolécules fondamentales à la fois support de l'information génétique et catalyseurs. Depuis, d'autres exemples de ribozymes ont été décrits (*e.g.* revue par (Fedor & Williamson, 2005)). La création de ribozymes artificiels, par sélection *in vitro* ou SELEX (défini §A.3.4.), catalysant des réactions biochimiques fondamentales telles que la formation de liaisons peptidiques (Zhang & Cech, 1997) et phosphodiesters (Bartel & Szostak, 1993) ou telle que l'aminoacylation (Lee et al., 2000), pourrait démontrer la faisabilité d'un monde dirigé par les ARN.

Dans un second temps, ce monde des ARN aurait évolué vers un monde ribonucléo-protéique (RNP). L'association d'acides aminés, puis de courts polypeptides, aurait amélioré les capacités catalytiques des ARN. L'émergence de la synthèse de protéines et notamment d'enzymes protéiques empiétant sur les fonctions de l'ARN, aurait induit progressivement la perte de la capacité catalytique des ARN. Puis l'ADN, en tant que forme catalytiquement inactive de l'ARN, est devenu le support stable de l'information génétique. Une alternative est que l'ADN aurait existé plus tôt, peut-être même dans le monde à ARN. (Doolittle & Brown, 1994 ; Maizels & Weiner, 1994).

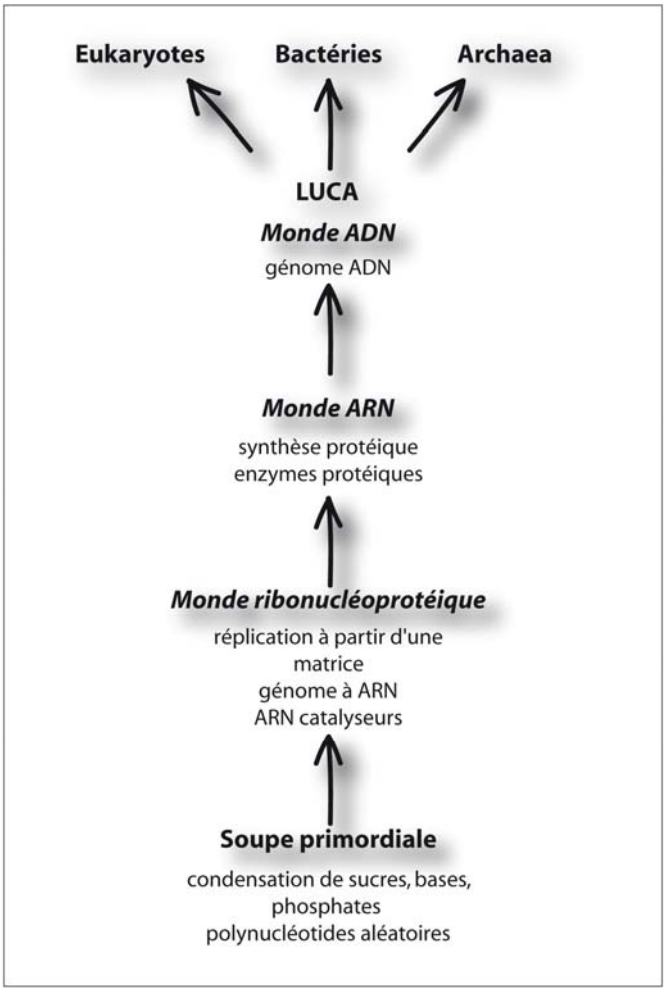



**Figure 7. Schéma illustrant les différentes étapes de l'origine de la Vie aux trois règnes du vivant.** LUCA signifie « Last Universal Common Ancestor ». Figure adaptée de (Maizels & Weiner, 1994).

L'étonnante similarité retrouvée dans la biochimie de tous les êtres vivants, a largement suggéré un ancêtre commun aux trois règnes de la Vie (Bactérie, Archaea et Eucaryote), LUCA (« Last Universal Common Ancestor »), nommé également progénote pour entité génétique (Woese, 1990). Mais qui était LUCA ? Etait-il simple ou complexe, unique ou pluriel ? La seule approche pour déterminer les vagues contours de cet ancêtre reste d'analyser les fossiles moléculaires conservés dans les 3 règnes du vivant. Devant la complexité à établir des liens de « parenté » évidents entre les différents organismes à partir d'arbres phylogénétiques (Forterre & Philippe, 1999), différentes propositions ont émergé. Si l'on considère que LUCA devait posséder tous les gènes conservés parmi les organismes contemporains, il devient une entité quasi totipotente et plus complexe que les cellules actuelles (Doolittle, 2000). Une alternative donc serait que ces gènes n'auraient pas forcément

été présents dans LUCA mais auraient été transférés latéralement d'organismes à organismes (*e.g.* (Martin, 1999)), expliquant la génération d'arbres phylogénétiques différents selon la molécule biologique utilisée comme critère de comparaison. Finalement, C. Woese propose que LUCA n'aurait pas été une entité à part mais une communauté complexe et hétérogène de progénotes simples capables d'évoluer rapidement (Woese, 1998). C'est la population dans son ensemble qui aurait été le réservoir de tous les gènes conservés actuellement. Se développant grâce au transfert de gènes, progressivement cette communauté se serait divisée pour former les 3 règnes du Vivant.

## 4.2. Émergence des systèmes d'aminoacylation du monde des ARN

Dans la biologie contemporaine, la connexion entre le monde ARN et les protéines se fait grâce aux systèmes d'aminoacylation. D'après le paradigme de l'évolution de la Vie, il est probable que l'émergence des protéines dans le monde des ARN n'aurait été qu'un accident, permis par des ribozymes liant les acides aminés comme co-facteurs à leur fonction, et qui peu à peu auraient donné lieu à la synthèse de nouveaux acides aminés puis de peptides (non codés dans un premier temps) (*e.g.* (Schimmel & Kelley, 2000)).

Dans le monde des ARN, les ribozymes catalysant l'aminoacylation auraient évolué par étapes successives : la fixation d'acides aminés puis de peptides leur auraient conférés de meilleures propriétés (réactivité, reconnaissance du substrat, stabilité) avant de laisser place aux protéines (aaRS primordiales) plus aptes à répondre à la nécessité d'une spécificité accrue de reconnaissance des substrats. Par exemple, il a été possible de créer par SELEX un ribozyme capable de reconnaître un acide aminé activé spécifique et de le transférer à un ARNt donné (Lee et al., 2000). La spécificité de reconnaissance de l'ARNt est directement médiée par des interactions Watson-Crick entre une séquence guide du ribozyme et l'extrémité acceptrice de l'ARNt. Le même laboratoire a réussi à créer un ribozyme précurseur d'ARNt capable de s'auto-aminoacyler (Saito et al., 2001). La catalyse en *trans* par le ribozyme artificiel décrit par (Lee et al., 2000) est particulièrement intéressante puisqu'elle se place à l'extrémité acceptrice de l'ARNt proposé comme site de localisation des signaux primordiaux reconnus par les aaRS. En effet, des expériences d'aminoacylation spécifiques de minihélices et microhélices dérivées d'ARNt (définies §A.3.4.) ont conduit à la

proposition d'un « code opérationnel ARN » dans le domaine accepteur qui aurait précédé le code génétique actuel. Ce code aurait ensuite co-évolué avec un système de synthèse de protéines (codées par des gènes) (Schimmel et al., 1993 ; Schimmel, 1995 ; Ribas de Pouplana & Schimmel, 2001b). Progressivement, les systèmes d'aminoacylation ont été mis en place. Dans un premier temps, ils auraient été dédiés à la synthèse d'acide aminés ARNt dépendante catalysée par des aaRS primordiales (voir §A.2.8.), puis se seraient spécialisés dans la synthèse de peptides.

## 4.3. Évolution des ARNt

L'ARNt est une molécule ubiquitaire qui a certainement été présente dans LUCA. (Maizels & Weiner, 1994) ont d'ailleurs placé son origine et sa fonction dans les mécanismes de réplication des génomes à ARN, avant même l'invention de la synthèse de protéines par traduction du code génétique, c'est-à-dire avant l'apparition du code et avant les aaRS (*e.g.* (Ribas de Pouplana et al., 1998)). Maizels & Weiner, 1994 se basent sur l'existence de structures en pseudo-ARNt impliquées dans la réplication des génomes viraux à ARN, et qui seraient donc les fossiles moléculaires de l'ARNt primordial. Bien sûr, l'ARNt primordial ne présentait sans doute pas les caractéristiques structurales d'aujourd'hui.

A l'origine, l'ARNt aurait été une molécule simple composée d'une tige boucle (Schimmel et al., 1993 ; Maizels & Weiner, 1994 ; Ribas de Pouplana & Schimmel, 2001b ; Giegé & Frugier, 2003 ; Widmann et al., 2005). C'est la nature même de l'ARNt contemporain, sa structure symétrique en deux branches (de l'anticodon et acceptrice) et leurs fonctions indépendantes (lecture du codon et fixation de l'acide aminé), qui a suggéré cette hypothèse (Figure 8A.). Notons qu'il existe d'ailleurs des modules distincts dédiés à la reconnaissance de chacune de ces branches au niveau des aaRS. De plus, la possibilité d'aminoacylation d'ARNt minimalistes (minihélices, *e.g.* (Francklyn & Schimmel, 1989 ; Frugier et al., 1994a)) et la présence d'éléments d'identité forts au niveau du bras accepteur (Giegé et al., 1998 ; Giegé & Frugier, 2003) définissant le code opérationnel ARN soutiennent fortement cette hypothèse. La structure en L de l'ARNt contemporain aurait été générée par l'association de deux tiges boucles, soit différentes soit jumelles. Cette hypothèse est particulièrement séduisante puisque l'association de deux molécules pour former un ARNt a été observée dans un organisme contemporain (Randau et al., 2005) et que beaucoup

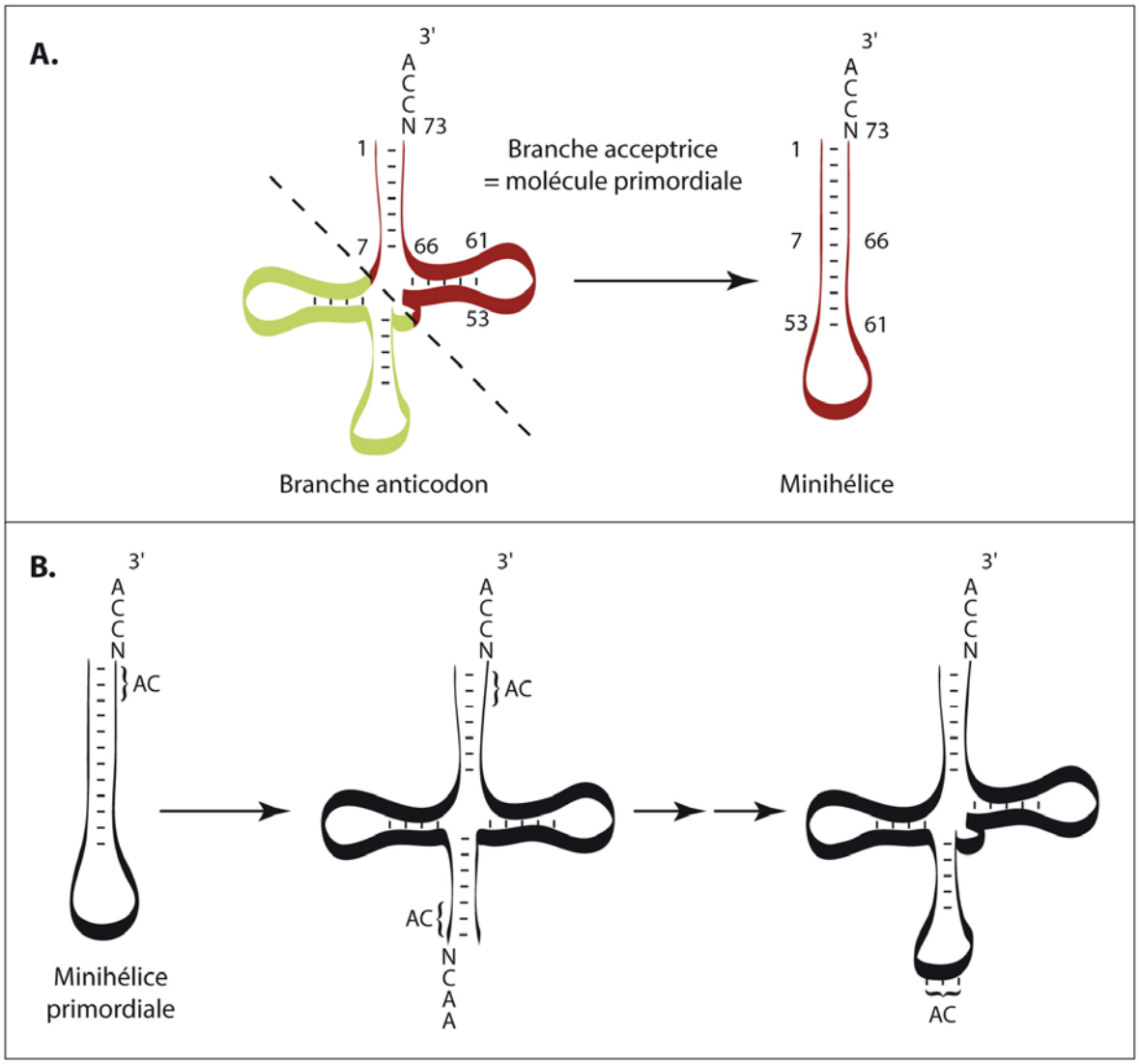



**Figure 8. Evolution de l'ARNt à partir d'une minihélice primordiale. A.** Mise en évidence des deux domaines structuraux (branches acceptrice et anticodon) au sein de l'ARNt contemporain, dont la symétrie est symbolisée par un trait en pointillé. La molécule primordiale (à droite) aurait été une minihélice formée uniquement de la branche acceptrice. **B.** Schéma évolutif de l'ARNt proposant la duplication, l'association et la ligation de deux minihélices portant à la fois l'anticodon (AC) et le code opérationnel. Les deux molécules primordiales auraient évolué vers la spécialisation et la séparation des fonctions de lecture du codon et fixation de l'acide aminé.

d'ARNt possèdent des introns au niveau de l'anticodon (Sprinzl & Vassilenko, 2005). Ceci suggère la création de l'ARNt contemporain par l'association de deux mini-gènes grâce à un mécanisme d'« exon shuffling » (*e.g.* (Gilbert et al., 1997)). Toutefois, l'ARNt primordial n'était pas forcément la branche acceptrice telle que nous la connaissons. (Di Giulio, 2004) propose une hypothèse quant à la forme originale de l'ARNt qui permet de concilier l'apparition du code génétique et sa conservation dans les 3 règnes du vivant alors que la présence de l'ARNt sous sa structure contemporaine dans LUCA est discutée (Di Giulio,

1999). L'ARN en tige boucle aurait porté les signaux à la fois de l'anticodon (définissant plus tard le code génétique) et du code opérationnel ARN. La duplication, l'association et la ligation des deux tiges boucles (au niveau de ce qui constitue aujourd'hui la boucle de l'anticodon) auraient généré la feuille de trèfle actuelle (Figure 8.B.). Puis, progressivement, les séquences des deux molécules primordiales auraient évolué vers leurs spécialisation et séparation au sein de l'ARNt contemporain en deux fonctions : lecture des codons et fixation de l'acide aminé.

La théorie de co-évolution du code génétique suggère l'existence d'un code primitif qui se serait étendu parallèlement à l'enrichissement en ARNt (Xue et al., 2003). La diversité en ARNt aurait été générée grâce à la duplication de gènes d'ARNt primitifs à spécificité relâchée (Saks et al., 1998 ; Lynch, 2000 ; Xue et al., 2003 ; Santos et al., 2004). Les nouveaux ARNt auraient été capturés soit par un acide aminé préexistant déjà codé (donnant lieu à un isoaccepteur) soit par un nouvel acide aminé (donnant lieu à un nouvel accepteur ou alloaccepteur) (Xue et al., 2003). Certains liens évolutifs entre les ARNt paralogues (issus d'un même ancêtre) sont encore détectables (fossile moléculaire), illustré par l'aminocylation croisée de certaines espèces d'ARNt (*e.g.* l'ARNt$^{Asn}$ qui est aspartylé, (Becker & Kern, 1998)) ou la forte similarité de séquences entre deux gènes (*e.g.* les ARNt$^{Tyr}$ et ARNt$^{Phe}$ dont les séquences partagent jusqu'à 94 % d'identité de séquence dans certaines archaea (Xue et al., 2003)).

## 4.4. Évolution des aminoacyl-ARNt synthétases

Les aaRS auraient émergé à partir du monde des ARN, primitivement sous la forme de ribozymes (voir précédemment), puis de ribonucléo-protéines et enfin de protéines.

Les aaRS contemporaines sont équivalemment réparties en 2 classes distinctes, sur la base de leurs sites catalytiques. Cette remarquable symétrie, l'analyse des séquences et la possibilité d'un mode de reconnaissance des deux faces opposées du bras accepteur de l'ARNt par des paires d'aaRS des deux classes (Ribas de Pouplana & Schimmel, 2001c ; 2001a), a conduit à l'hypothèse suivante : les 2 domaines catalytiques (aaRS primordiales) auraient évolué à partir de gènes dont le brin sens et anti-sens auraient simultanément codé pour les ancêtres des deux classes (Rodin & Ohno, 1995; Ribas de Pouplana & Schimmel, 2001c ; Carter Jr & Duax, 2002 ). Une autre théorie suggère que les deux groupes d'enzymes

seraient apparues de façon indépendante et aurait fonctionnellement convergé (Nagel & Doolittle, 1991). De façon alternative, les enzymes de classes II seraient les plus anciennes (Klipcan & Safro, 2004). Les arguments en faveur de cette dernière hypothèse sont entre autres que : (i) les acides aminés reconnus par les aaRS de classe II sont plus simples et plus petits (Carter Jr, 1993 ; Hartman, 1995), (ii) beaucoup d'entres eux ont des rôles cruciaux dans la structure et/ou activité des protéines (Gly, Pro, Ala, Asp, His, Thr et Ser), et de façon remarquable (iii) les substrats minimalistes (minihélices acceptrices) récapitulant le code opérationnel ARN (ou second code génétique primitif) sont préférentiellement reconnus par les aaRS de classe II (*e.g.* (Schimmel et al., 1993)).

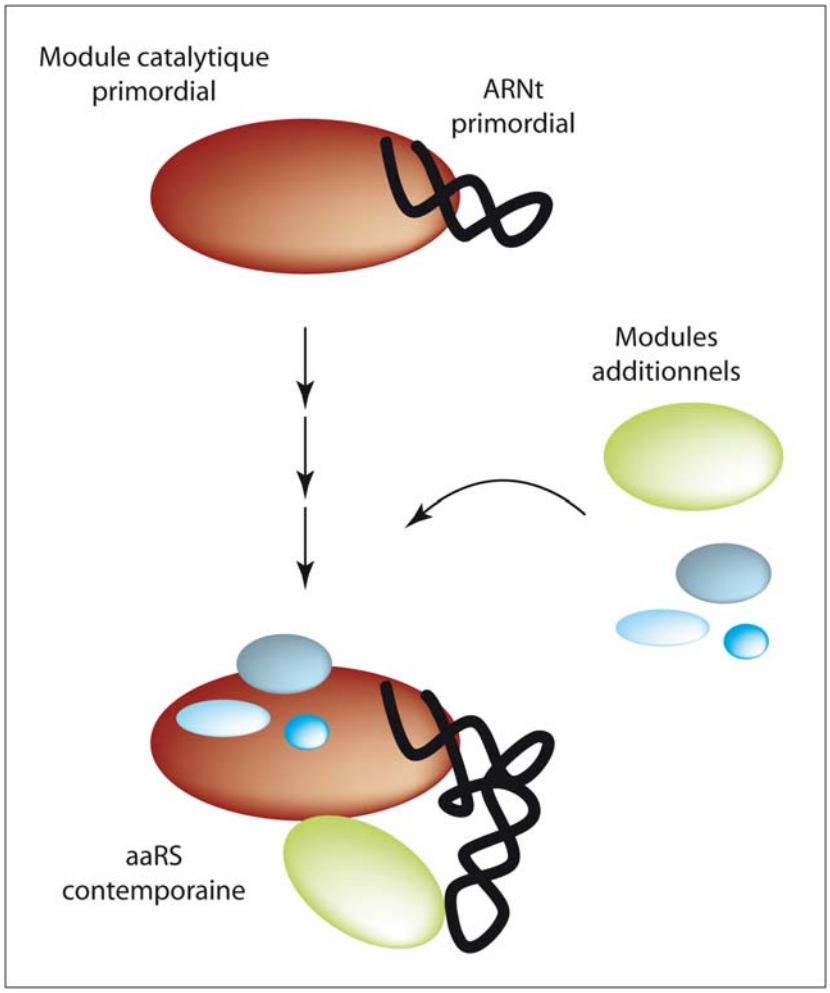



**Figure 9. Evolution des aminoacyl-ARNt synthétases.** La forme primitive des aaRS aurait été réduite au cœur catalytique (en rouge) auquel se serait ajouté le domaine de liaison à l'anticodon (en vert) et d'autres domaines additionnels, de façon concomitante à l'évolution de l'ARNt.

Quoiqu'il en soit, tous les auteurs s'accordent pour dire que la forme primitive des aaRS aurait été réduite au cœur catalytique, auquel se seraient ajoutés des domaines additionnels de façon concomitante à l'évolution de l'ARNt et à la nécessité d'une spécificité

d'aminoacylation accrue (*e.g.* domaine de liaison à l'anticodon, d'édition,…), ou à l'apparition de nouveaux mécanismes cellulaires (*e.g.* domaine EMAPII de la TyrRS impliquées dans l'angiogenèse, voir §A.2.7) (Figure 9). Ceci est largement illustré par l'organisation modulaire des aaRS contemporaines (résultat probable d'association de domaines par « exon shuffling ») et l'existence de protéines paralogues (voir §A.2.8).

La comparaison des séquences a permis d'établir des liens entre les aaRS, suggérant des ancêtres communs. Par exemple, la duplication et l'évolution du gène d'une AspRS primitive non discriminante aurait généré l'AspRS spécifique et l'AsnRS. Les empreintes de cette histoire évolutive sont clairement retrouvées dans certains organismes pour lesquels l'AsnRS est manquante et qui conservent alors une AspRS non discriminante capable d'aminoacyler l'ARNt$^{Asp}$ mais aussi l'ARNt$^{Asn}$ (Becker & Kern, 1998). Notons la conservation de l'AspRS non discriminante dans *T. thermophilus* alors que l'AsnRS est présente (Becker et al., 1997). Ceci pourrait refléter la conservation d'un mécanisme ancestral de biosynthèse d'acide aminé ARNt dépendante (Francklyn, 2003; Roy et al., 2003 ).

## 4.5. Conservation des éléments d'identité ou barrières d'espèces

Selon la théorie du code opérationnel ARN, les premiers éléments d'identité, se retrouvent dans l'extrémité acceptrice de l'ARNt contemporain (base discriminatrice N73 et généralement les 3 premières paires de bases). Pour certains systèmes, les éléments majeurs d'identité sont conservés (Giegé et al., 1996), et de façon étonnante, ils sont en partie retrouvés dans l'extrémité acceptrice de l'ARNt (les autres éléments majeurs constituant généralement l'anticodon). Par exemple, la base discriminatrice G73, élément d'identité majeur aspartate, est globalement conservée dans les ARNt$^{Asp}$ et suffit à l'aminoacylation de minihélices (Frugier et al., 1994a). Il en va de même pour la paire de bases G3-U70 qui spécifie l'alanylation des ARNt$^{Ala}$ et de substrats minimalistes (Francklyn & Schimmel, 1989). Cependant, c'est également au niveau du bras accepteur que peuvent exister des variations d'identité entre les systèmes hétérologues (de même spécificité mais d'organismes différents). Dans certains cas, ces variations suffisent à empêcher toute aminoacylation hétérologue, définissant une barrière d'espèces entre les systèmes d'aminoacylation. De telles délimitations existent pour plusieurs systèmes d'aminoacylation différents, et se placent

généralement entre des systèmes de type bactérien et archaea/eucaryote. Notons que de façon étonnante, la variation de la nature des éléments d'identité des ARNt selon les règnes est également corrélée à l'existence d'aaRS bien distinctes, suggérant l'adaptation étroite des enzymes. Par exemple, la glycylation des $ARNt^{Gly}$ de type bactérien ou eucaryote/archaea repose sur des nucléotides différents en position 73 (U et A respectivement) empêchant ainsi toute reconnaissance croisée. L'échange expérimental de cette base inverse la spécificité des différents types d'ARNt et permet l'aminoacylation hétérologue (Shiba et al., 1994a ; Hipps et al., 1995). Notons qu'en accord avec la barrière d'espèces, on observe deux groupes structuralement distincts de GlyRS. En particulier, les enzymes bactériennes possèdent une organisation oligomérique de type $\alpha 2\beta 2$ alors que les GlyRS eucaryotiques et d'archaea sont de type $\alpha 2$ (Shiba et al., 1997). De la même façon, la nature de la paire de bases 1-72 (G1-C72 chez les bactéries et C1-G72) empêche l'aminoacylation hétérologue. Ceci est corrélé à l'existence deux groupes distincts de TyrRS. Notamment un peptide de 39 acides aminés dans la partie N-terminale du domaine CP1 est responsable de la différence de reconnaissance (*e.g.* revue par (Bonnefond et al., 2005b)).

# B. Le système d'aspartylation dans divers organismes

## 1. Identité des $ARNt^{Asp}$

Il n'existe qu'un seul accepteur $ARNt^{Asp}$ dont l'anticodon est GUC. En effet, le code génétique ne comporte que deux codons aspartate GAU et GAC. Grâce à l'interaction de type wobble entre le premier nucléotide du codon et le dernier de l'anticodon, l'$ARNt^{Asp}$GUC est capable de lire les deux codons.

### Eléments d'identité

A ce jour, les éléments d'identité aspartate ont été principalement étudiés par des approches *in vitro* au sein des systèmes bactériens d'*E. coli* (Hasegawa et al., 1989 ; Nameki et al., 1992), de *T. thermophilus* (Becker et al., 1996) et eucaryotique de *S. cerevisiae* (Pütz et

al., 1991 ; Frugier et al., 1994b ; Ryckelynck et al., 2003). Ces éléments sont globalement conservés (Giegé et al., 1996) et sont la base discriminatrice G73 et l'anticodon G34U35C36 pour les majeurs (aux deux extrémités de l'ARNt) et la paire de bases G10-U/C25 et le nucléotide C38 pour les mineurs (Figure 10). A noter la contribution mineure de la paire de bases G2-C71 dans le système d'*E. coli*. La force des éléments d'identité est différente en fonction du système étudié. En effet, la substitution de l'élément majeur U35 par un A induit une perte d'efficacité de 85 fois chez *S. cerevisiae* (Ryckelynck et al., 2003), >5000 fois (Nameki et al., 1992) chez *E. coli* et 3 $10^5$ fois chez *T. thermophilus* (Becker et al., 1996). De telles différences indiquent des réseaux d'interaction hydrogènes différents entre les nucléotides d'identité et les acides aminés de l'AspRS ou plus simplement renseigne sur l'importance fonctionnelle des groupements chimiques des résidus de l'identité (Giegé et al., 1996).

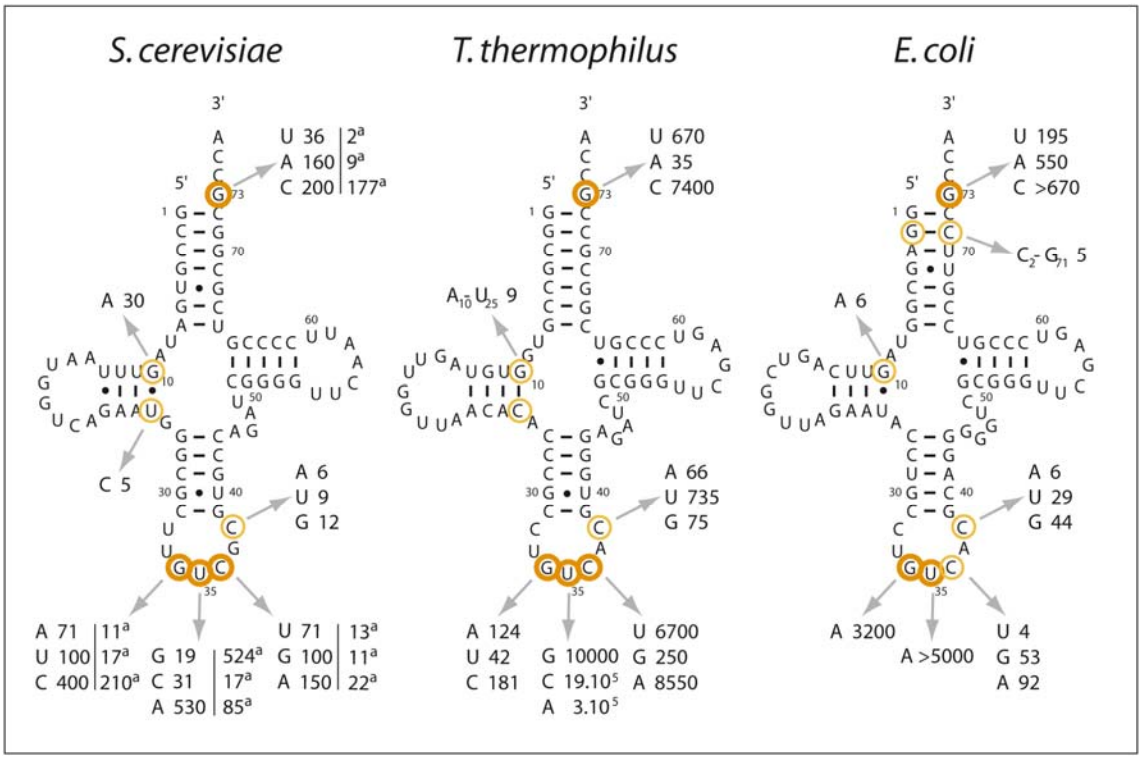



**Figure 10. Identité des ARNt$^{Asp}$ de *S. cerevisiae, T. thermophilus* et *E. coli*.** Les éléments d'identité sont mis en évidence par des cercles oranges épais (éléments majeurs) ou fins (éléments mineurs). Les pertes d'efficacité induites par les mutations sont indiquées (pour références voir texte). Dans le cas de la levure, deux valeurs sont données : elles correspondent soit à l'utilisation de l'enzyme extraite de la levure (Pütz et al., 1991), soit clonée et surproduite dans *E. coli* (valeurs indiquées par [a] , (Ryckelynck et al., 2003)).

Les structures cristallographiques des complexes (voir plus loin) ont révélé la présence d'interactions directes, via des liaisons hydrogènes, entre éléments d'identité de l'ARNt et acides aminés de l'AspRS. C'est le cas pour les éléments majeurs G73, G34U35C36 et mineur G10 et C38 (visible dans les structures des complexes d'*E. coli* (Eiler et al., 1999) et de levure (Cavarelli et al., 1993)). L'élément G10 interagit directement avec l'AspRS d'*E. coli* et non avec celle de levure. Dans ce dernier cas, sa contribution est indirecte, il participe à la présentation optimale des autres éléments. L'interaction des éléments majeurs d'identité a aussi été montré par des expériences d'empreintes en solution en ce qui concerne le système de levure (Garcia & Giegé, 1992; Rudinger et al., 1992 ).

Ces jeux d'identité ont été vérifiés par leur transplantation dans des ARNt d'autres spécificités. Transplantés, l'ARNt$^{Phe}$ de levure et l'ARNt$^{Gln}$ d'*E. coli* deviennent des substrats de l'AspRS de levure (Pütz et al., 1991 ; Frugier et al., 1994b), l'ARNt$^{Phe}$ de levure substrat de l'enzyme de *T. thermophilus* (Becker et al., 1996), et les ARNt$^{Val}$ et ARNt$^{Asn}$ d'*E. coli* des substrats de l'AspRS d'*E. coli* (Nameki et al., 1992).

Les éléments majeurs d'identité aspartate sont probablement conservés au cours de l'évolution, puisqu'ils sont retrouvés dans toutes les séquences d'ARNt$^{Asp}$ connues (Sprinzl & Vassilenko, 2005), avec pour seule exception les ARNt mt de mammifères (voir chapitre 2, C.).

### Communications entre les éléments

L'étude cinétique de variants de l'ARNt$^{Asp}$ de levure transcrits *in vitro*, mutés aux positions d'identité G34, U35, C36, G73 et G10-U25 ont permis d'établir un schéma de communication entre les différents éléments d'identité (Pütz et al., 1993). Ainsi trois types de relations ont été démontrés. Alors que les nucléotides situés loin l'un de l'autre dans la structure de l'ARNt agissent de façon coopérative (l'effet d'une mutation est amplifié par les autres positions mutées) alors que les éléments de l'anticodon agissent soit de façon additive (l'effet d'une mutation n'est pas altéré par une autre mutation) ou anti-coopérative (la mutation d'une position compense l'effet des autres mutations).

### Aspartylation hétérologue

Alors que l'AspRS de levure aminoacyle efficacement l'ARNt$^{Asp}$ d'*E. coli*, l'AspRS d'*E. coli* ne reconnaît pas l'ARNt de levure. L'étude cristallographique du complexe inactif hétérologue ARNt$^{Asp}$ de levure/AspRS d'*E. coli* a révélé une structure asymétrique, avec un seul des deux ARNt fixés entrant dans le site actif de la sous unité enzymatique

correspondante (Moulinier et al., 2001). La substitution des deux premières paires de bases de l'ARNt$^{Asp}$ de levure par celles retrouvées dans l'ARNt$^{Asp}$ bactérien permettent une aspartylation hétérologue efficace. Notons que ces mutations introduisent l'élément d'identité mineur spécifique du système d'*E. coli*. En ce qui concerne le système de *T. thermophilus*, l'AspRS est capable d'aminoacyler les ARNt$^{Asp}$ de levure et d'*E. coli* (Becker et al., 1996). Ceci suggère qu'il n'y a pas d'antidéterminants au sein de ces deux ARNt et que l'enzyme est suffisamment flexible pour s'adapter.

### Aspartylation de minihélices et importance de la base discriminatrice G73

Les ARNt peuvent être disséqués en deux à la jonction entre la branche acceptrice et de l'anticodon. Il a été montré que différentes structures, minihélices acceptrice complète ou dérivées de l'ARNt$^{Asp}$ de levure, restent des substrats de l'AspRS de levure, pouvant atteindre un plateau de charge jusqu'à 30% (Frugier et al., 1994a). Notons que des minihélices dérivant des ARNt$^{Asp}$, ARNt$^{Gln}$, ARNt$^{Sec}$ d'*E. coli* et ARNt$^{Arg}$ de levure sont aussi substrats de l'AspRS de levure, (Frugier et al., 1994a ; Rudinger et al., 1994 ; Martinis & Schimmel, 1995 ; Rudinger et al., 1996). Dans tous les cas, l'aspartylation de ces structures est fortement dépendante et spécifique de la base discriminatrice G73. Des expériences plus anciennes avaient déjà montré la quasi-conservation du résidu G73 dans la majorité des séquences d'ARNt mis-aspartylés (Ebel et al., 1973).

## 2. Structure/fonction de l'AspRS

Actuellement, il existe plus de 500 séquences d'AspRS répertoriées (http://www.ebi.ac.uk/interpro), et de nombreuses informations tant fonctionnelles que structurales existent pour les trois règnes du vivant, archae, eucaryotes et bactéries. Les AspRS, comme la plupart des enzymes de classe II, sont homodimériques et les deux sous-unités sont reliées par un axe de symétrie d'ordre 2. La dimérisation des deux sous-unités est médiée par une grande hélice α (dite de dimérisation) et implique des résidus conservés du motif 1 du site catalytique. Les différentes informations décrites dans ce qui suit sont issues en grande partie de la revue (Giegé & Rees, 2005) et sont illustrées dans la Figure 11.

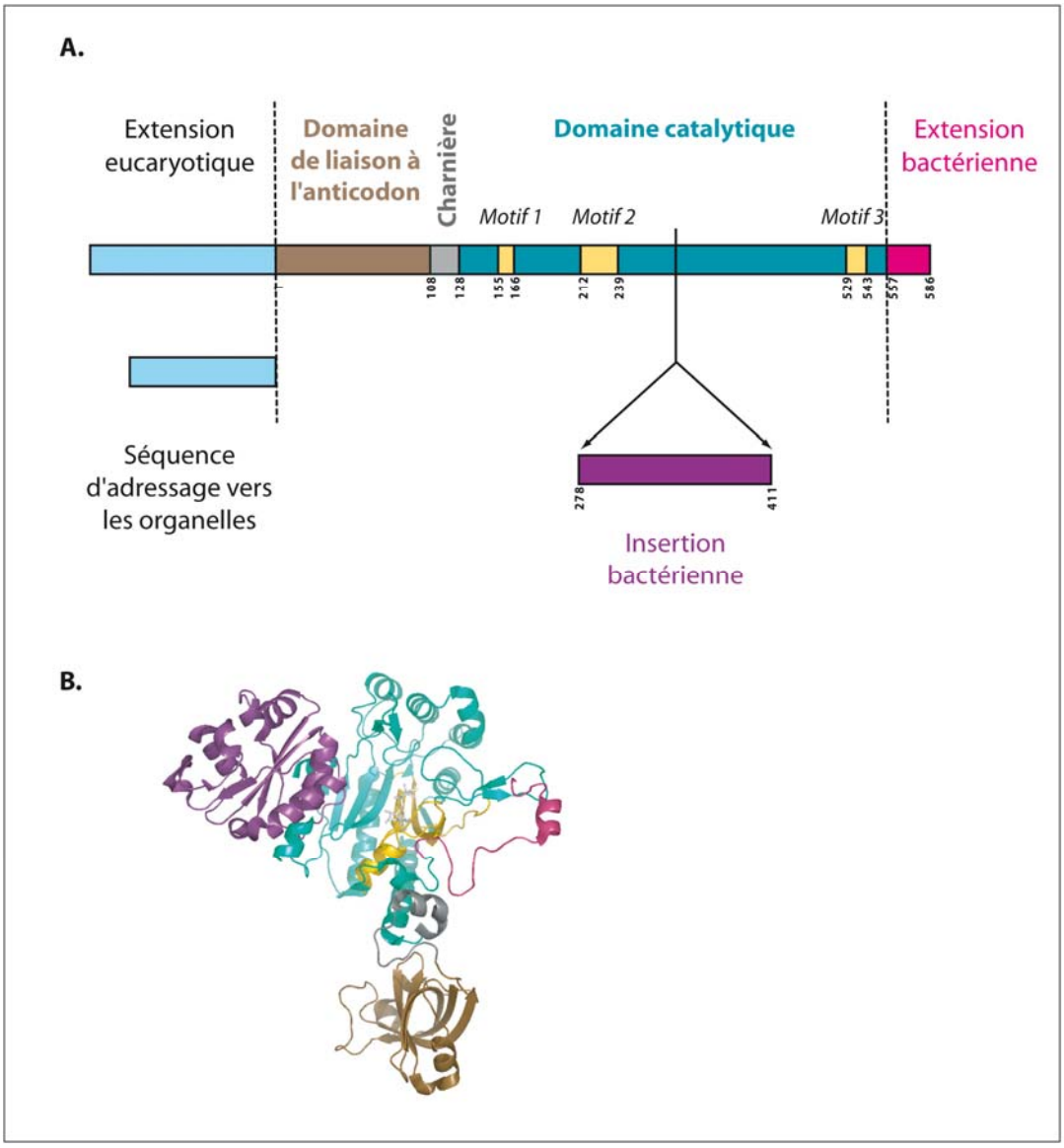



**Figure 11. Organisation modulaire des AspRS. A.** Représentation schématique des différents domaines. Ceux conservés dans l'ensemble des règnes du vivant sont indiqués en gras, et ceux spécifiques des eucaryotes, bactéries ou organelles sont donnés. La numérotation délimite les régions de l'AspRS d'*E. coli*. **B.** Structure cristallographique d'un monomère de l'AspRS d'*E. coli* (Rees et al., 2000) où les différentes régions sont mises en évidence. Le code couleur est le même qu'en **A**. La figure **B.** a été générée en utilisant le programme PyMol (DeLano, 2002).

Bien que les AspRS possèdent peu de similarités de séquences entre elles, la comparaison de leurs structures tertiaires montre une organisation commune composée de trois modules, mais aussi de modules additionnels spécifiques de chaque règne du vivant. Les modules communs sont :

(i) Le domaine catalytique, le plus large, autour duquel s'articule tous les autres modules. Son coeur est composé d'un feuillet β de 6 brins antiparallèles, entouré d'hélices α. Ce repliement est commun à toutes les aaRS de classes II. Le site actif présente les motifs

signatures typiques des enzymes de classe II. Des résidus conservés du motif 1 sont impliqués dans la dimérisation des deux monomères.

(ii) Le domaine de liaison à l'anticodon situé dans la partie N-terminale. Son repliement est un OB-fold. Il est composé d'un tonneau β de 5 brins avec une hélice α entre les 3$^{ème}$ et 4$^{ème}$ brins. Ce repliement est aussi retrouvé dans les autres aaRS de sous-classe IIb (AsnRS et LysRS). Les trois enzymes reconnaissent des anticodons possédant tous un U en position centrale.

(iii) La région charnière reliant les deux domaines. Petit module, sa taille varie selon les règnes : il est plus long dans les AspRS eucaryotiques (37 à 46 résidus) que dans les AspRS bactériennes ou d'archaea (21 à 30 résidus). Il est composé de courtes hélices α. Cette région de la protéine interagit avec le bras D de l'ARNt.

En plus de ces trois domaines conservés, des modules sont spécifiques à chaque règne. Les AspRS bactériennes possèdent un domaine d'insertion entre les motifs 2 et 3 du site actif et une extension dans sa partie C-teminale. Le domaine d'insertion est constitué d'un feuillet β de 5 brins antiparallèles, flanqué de deux hélices α antiparallèles sur la face dirigée vers le cœur catalytique de la protéine et d'une hélice α sur la face opposée. Ce repliement, nommé « ferredoxin fold », est retrouvée dans d'autres protéines telles que la glutamine synthétase (Almassy et al., 1986), la nucléoside diphosphate kinase (Janin et al., 2000) ou la sous unité B de la Glu-tRNA$^{Gln}$ amidotransférase (GatB) (Wolf et al., 1999). Les deux hélices antiparallèles interagissent avec le bras accepteur de l'ARNt. Le domaine serait impliqué dans la stabilisation de l'interaction entre l'AspRS bactérienne et l'ARNt spécifique (Delarue et al., 1994). L'extension C-terminale, de 20 à 30 résidus, fait partie de l'interface entre les deux sous-unités. Notamment, les deux sous unités interagissent via un petit feuillet β à deux brins présent dans cette extension. Les AspRS eucaryotiques possèdent quant à elles une extension N-terminale. L'extension de l'AspRS de *S. cerevisiae*, composée des 70 premiers acides aminés et prédite pour se structurer en trois hélices (Agou et al., 1995), n'est pas essentiel pour l'aminoacylation mais contient un motif de liaison à l'ARN riche en lysines et conservés dans les AspRS eucaryotiques et les autres aaRS de la sous-classe IIb (AsnRS et LysRS). Ce motif interagit de façon aspécifique avec le sillon mineur du bras anticodon de l'ARNt et augmente la stabilité du complexe ARNt/aaRS et l'efficacité catalytique d'aminoacylation (Frugier et al., 2000). Par l'intermédiaire de cette extension, l'AspRS est également capable de se lier en 5' de son propre ARNm et de réguler ainsi son expression (Frugier et al., 2005). Dans le cas de l'AspRS de rat, l'extension est nécessaire à l'association de l'enzyme au

complexe multisynthétasique *in vivo* (Mirande et al., 1992) et dans le cas de l'enzyme cytoplasmique humaine, (Reed et al., 1994) ont suggéré un rôle important dans la mise à disposition de l'Asp-ARNt$^{Asp}$ pour le facteur d'élongation 1α (EF1α). Par une étude structurale par spectroscopie (RMN, Résonance Magnétique Nucléaire) de l'extension N-terminale de l'AspRS humaine, son interaction avec l'ARNt est confirmée et un repliement en tournant β suivi d'une hélice amphiphile est proposé (Cheong et al., 2003).

Actuellement, il existe 11 structures cristallographiques d'AspRS référencées sur le site de la « Protein Data Bank » (PDB). Ces structures ont été obtenues pour les enzymes, libres ou en complexes, des trois règnes du vivant : bactéries (*E. coli* et *T. thermophilus*), archaea (*Pyrococcus kodakaraensis*) et eucaryotiques (*S. cerevisiae*). Il n'existe pas encore de structure d'AspRS d'organelles. Huit de ces structures sont montrées (Figure 12), à savoir l'AspRS libre de chacun des organismes, et l'AspRS en complexe avec l'ARNt et l'ATP ou l'adénylate des organismes pour lesquels la structure existe. Les données correspondantes sont indiquées dans le tableau suivant.

| Organisme | Ligand | Code PDB | Référence |
|---|---|---|---|
| **Bactéries** | | | |
| *E. coli* | | | |
| libre | | **1EQR** | (Rees et al., 2000) |
| complexée | $ARNt^{Asp}$ + adénylate | **1C0A** | (Eiler et al., 1999) |
| | $ARNt^{Asp}$ de levure + adénylate | 1IL2 | (Moulinier et al., 2001) |
| *T. thermophilus* | | | |
| AspRS-1 | | | |
| libre | | **1L0W** | (Ng et al., 2002) |
| complexée | adénylate | 1G51 | (Poterszman et al., 1994) |
| | $ARNt^{Asp}$ d'*E. coli* | **1EFW** | (Briand et al., 2000) |
| AspRS-2 | | | |
| libre | | **1N9W** | (Charron et al., 2003) |
| **Archaea** | | | |
| *P. kodakaraensis* | | | |
| libre | | | |
| complexée | ATP<br>adénylate<br>aspartate | **1B8A** | (Schmitt et al., 1998) |
| **Eucaryotes** | | | |
| *S. cerevisiae* | | | |
| libre | | **1E0V** | (Sauter et al., 2000) |
| complexée | $ARNt^{Asp}$ | 1ASY | (Ruff et al., 1991) |
| | $ARNt^{Asp}$ + ATP | **1ASZ** | (Cavarelli et al., 1994) |

**Tableau 2. Ensemble des données concernant les structures cristallographiques d'AspRS..** Sont indiquées en gras les codes PDB des structures représentées Figure 12. Les données ont été extraites et complétées à partir de (Giegé & Rees, 2005).

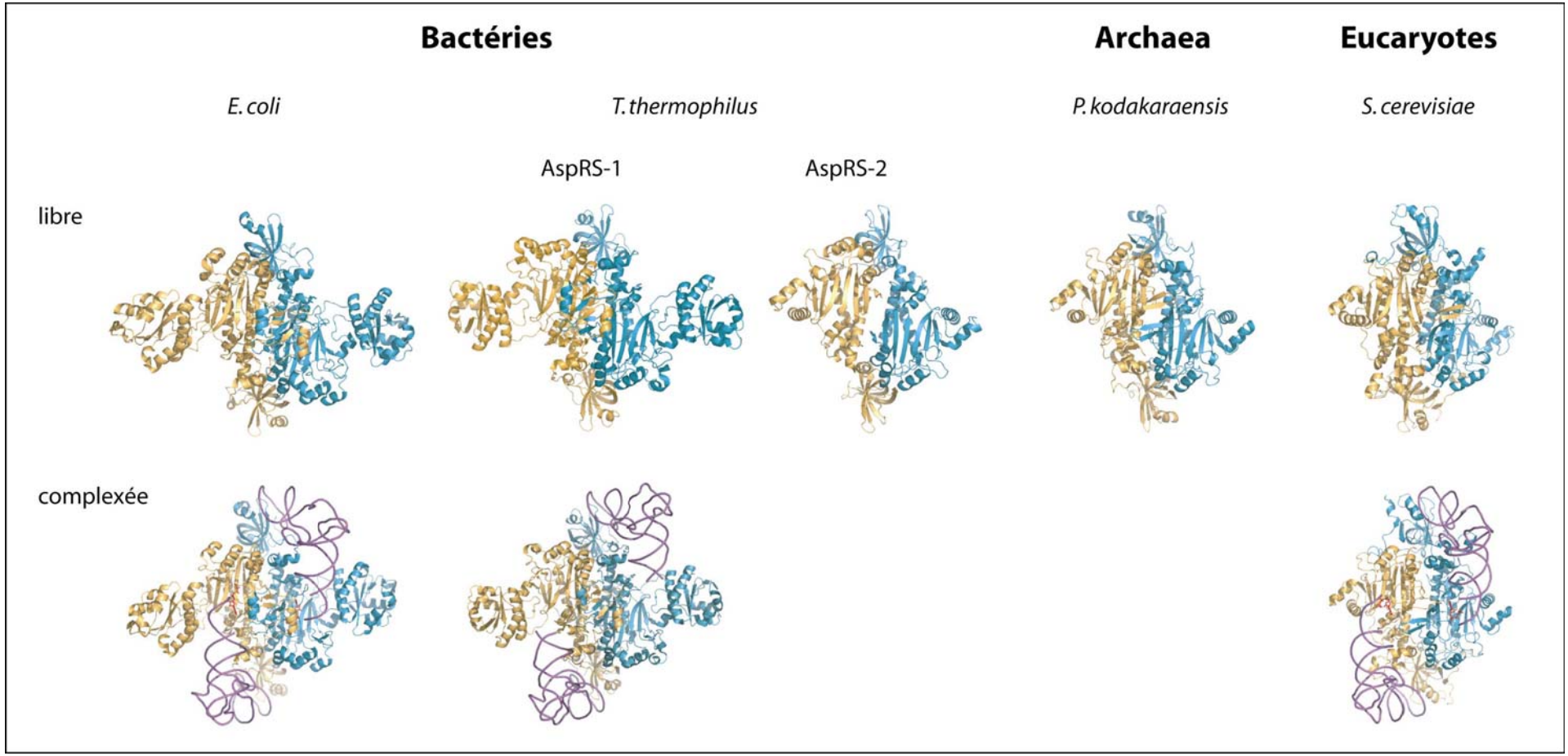



**Figure 12. Structures cristallographiques d'AspRS des trois règnes du vivant.** Les enzymes sont libres ou complexées à l'ARNt. Les dimères d'AspRS (orange et vert) sont sous forme schématique où hélices α et brins β sont visibles, et les ARNt (violet) sont représentés en fil de fer. L'adénylate ou l'ATP sont en rouge. La figure a été générée en utilisant le programme PyMol (DeLano, 2002). Pour références et compléments voir Tableau 2.

## 3. Interaction ARNt$^{Asp}$ / AspRS

Chaque monomère de l'AspRS interagit avec une molécule d'ARNt. L'interaction la plus importante, et sans doute la première a être mise en place, est celle des nucléotides de l'anticodon avec le domaine correspondant de l'enzyme (partie N-terminale). Les contacts établis entre l'enzyme et l'ARNt sont assez limités (pour les complexes d'*E. coli* ou de levure, ils ne représentent que 5-6% des acides aminés de l'AspRS (Cavarelli et al., 1993; Eiler et al., 1999 ). La plupart des contacts se font via des liaisons hydrogènes ARN/protéine. Certains contacts sont médiés par des molécules d'eau (surtout au niveau du site actif de l'enzyme qui interagit avec le bras accepteur). Un seul contact par empilement est retrouvé au niveau de l'interaction avec le résidu U35, qui est spécifique des classes IIb d'aaRS (Cavarelli et al., 1993).

Au sein du complexe AspRS/ARNt, on observe des changements de conformation tant au niveau de la protéine (changement de l'orientation relative des domaines) qu'au niveau de l'ARNt (réduction de l'angle entre les branches acceptrice et de l'anticodon de 110° pour la forme libre à 95° dans les complexes d'*E. coli* et de levure). En particulier, l'interaction au niveau du domaine anticodon induit des changements importants dans la boucle anticodon de l'ARNt$^{Asp}$. Alors que dans la forme libre de l'ARNt les bases de l'anticodon sont empilées, dans la forme complexée elles sont complétement évasées vers la synthétase.

L'analyse des propriétés catalytiques de variants de l'AspRS (Eriani & Gangloff, 1999) et la comparaison de différentes structures d'AspRS libres et en complexe avec l'ARNt$^{Asp}$ homologue ou non homologue (voir Tableau 2) suggère une interaction en plusieurs étapes successives. La première interaction aurait lieu au niveau de l'anticodon. En effet, celui-ci se retrouve correctement positionné dans tous les complexes, homologues ou non homologues, alors que le bras accepteur, ou une partie, n'est pas toujours visible. Un argument en faveur de ceci est que les acides aminés en contact avec les résidus de l'anticodon contribuent le plus fortement à la stabilité du complexe, les mutations de ces acides aminés dans l'AspRS conduisant à la plus forte perte d'énergie libre (Eriani & Gangloff, 1999). La deuxième étape consiste au positionnement du bras D de l'ARNt au niveau de la région charnière qui subit un réarrangement structural (Sauter et al., 2000). Puis, l'interaction de l'extrémité acceptrice avec le domaine catalytique (et le domaine d'insertion chez les bactéries) induit la réduction de l'angle entre les deux branches de l'ARNt. Ceci permet à l'élément d'identité G73 d'être reconnu par des résidus du motif 2 du site

catalytique. Ce motif subit un changement conformationnel en entrant dans le sillon majeur de l'hélice acceptrice. Finalement, une boucle située entre les motifs 1 et 2 du site catalytique, nommée « flipping loop », s'ouvre de façon à permettre l'ancrage correcte du nucléotide terminal A76. Une illustration claire de cette dernière étape vient du complexe non productif de l'AspRS et de l'ARNt$^{Asp}$ de levure (Moulinier et al., 2001). Dans ce cas, la « flipping loop » reste fermée sur la partie acide aminé de l'adénylate.

# C. Les systèmes d'aminoacylation mitochondriaux de mammifères

Un intérêt grandissant pour les mitochondries ces dernières décennies vient du fait qu'elles sont le siège de voies métaboliques d'importance vitale (*e.g.* synthèse d'ATP) mais aussi de la corrélation de mutations du génome mitochondrial (mt) à des pathologies humaines graves, maladies neurodégénératives pour les plus courantes. Bien que la mitochondrie ait été mise en évidence pour la première fois à la fin du 19$^{ème}$ siècle par les cytologistes, beaucoup d'aspects restent encore mal connues. En effet, la mitochondrie est une entité à part, ayant son propre génome, sa propre machinerie traductionnelle et son propre environnement. Les systèmes d'aminoacylation mt sont encore peu connus, à l'inverse de ceux cytoplasmiques, bactériens ou d'archaea. Il est pourtant nécessaire d'accumuler des connaissances fondamentales sur ces systèmes afin d'en comprendre les particularités et d'en contrecarrer les disfonctionnements.

Avant de développer plus largement la description des systèmes d'aminoacylation et essentiellement celle des ARNt bizarres (ils possèdent des séquences et structures dégénérées par rapport aux ARNt canoniques), les deux premières parties de ce chapitre seront consacrées d'une part au rappel de l'histoire, de la structure et de la fonction de la mitochondrie, et d'autre part à la description des mécanismes fondamentaux permettant l'expression du génome mt. Une dernière partie sera consacrée aux pathologies mt.

L'ensemble des connaissances fondamentales biologiques, biochimiques et moléculaires concernant la mitochondrie, peuvent être trouvées dans l'ouvrage « Mitochondria » (Scheffler, 1999).

# 1. Généralités

## 1.1. Origine endosymbiotique

Largement controversée au début du 20ème siècle, la théorie de l'endosymbiose fait aujourd'hui autorité. Elle propose que la mitochondrie résulte de l'endosymbiose d'une alpha-protéobactérie à l'intérieur de cellules eucaryotes primitives. Au cours de l'évolution, la redondance de beaucoup de fonctions avec celles de l'organisme hôte aurait généré la perte d'un certain nombre de gènes de l'endosymbiote. Un certain nombre d'autres gènes auraient été transférés dans le noyau (Adams & Palmer, 2003). L'existence de nombreux pseudogènes homologues à l'ADN mt humain retrouvés dans le génome nucléaire témoignerait de ce transfert (*e.g.* revue par (D'Errico et al., 2004)). Ainsi, le génome mt est devenu de plus en plus compact. En particulier chez l'Homme, il ne code que pour 37 gènes dont seulement 13 protéines. Le maintien de ces gènes serait lié au caractère très hydrophobe des protéines codées, difficilement importables du cytoplasme. Cependant, la description d'une organelle particulière, l'hydrogénosome (Palmer, 1997), très proche de la mitochondrie et ne possédant pas de génome suggère que le degré de compaction maximal du génome mt n'est peut-être pas encore atteint.

Bien qu'une origine monophylétique ait été confirmée (Burger et al., 1996), à savoir qu'une seule protéobactérie serait à l'origine de toutes les mitochondries connues à ce jour, les génomes mitochondriaux varient en taille et en nombre de gènes selon les organismes dans lesquels ils évoluent. Ainsi, le génome mt du zooflagellé *Reclinomonas americana* est resté comme oublié par le temps (Lang et al., 1997; Palmer, 1997 ) avec une taille de 69 034 paires de bases (comparé à 16 569 pb pour la mitochondrie humaine) et une quantité surprenante de gènes (97 au total). De façon intéressante, ce génome mt code non seulement pour de nombreuses protéines, dont les gènes ont été transférés dans le noyau dans la plupart des autres organismes, mais aussi pour les sous unités protéiques d'une ARN-polymérase de type bactérienne qui apparaît perdue par toutes les autres mitochondries et remplacée dans ce cas par une enzyme codée par le noyau et similaire à celle des bactériophages T3 et T7. Ainsi, la connaissance des génomes mitochondriaux et le séquençage de beaucoup d'autres permettront peut-être l'ébauche d'un schéma définissant le statut initial et l'histoire évolutive de l'endosymbiote.

## 1.2. Structure et fonction de la mitochondrie

La mitochondrie (« mito » pour filaments et « chondrie » pour grains), organelle d'approximativement 0,3-0,5 µm, est connue pour être la centrale énergétique de la cellule aérobie, puisqu'elle est la principale productrice d'ATP. Le nombre de mitochondries par cellules peut varier de quelques centaines à plusieurs milliers en fonction du type cellulaire. Dans les cellules spécialisées des vertébrés, ces variations peuvent être considérables. Alors que l'ovocyte (ovule non fécondé) peut contenir environ 100 000 mitochondries, le spermatozoïde n'en contient que quelques-unes. La mitochondrie est délimitée par deux membranes : la membrane externe perméable aux petites molécules et la membrane interne très repliée sur elle-même formant les crêtes mitochondriales (Figure 13). Cependant, la notion de mitochondire isolée est aujourd'hui révisée avec la mise en évidence de mitochondries liées par une dynamique membranaire se scindant et se réassociant activement, conduisant à des fusions/fissions (Nunnari et al., 1997 ; Rizzuto et al., 1998).

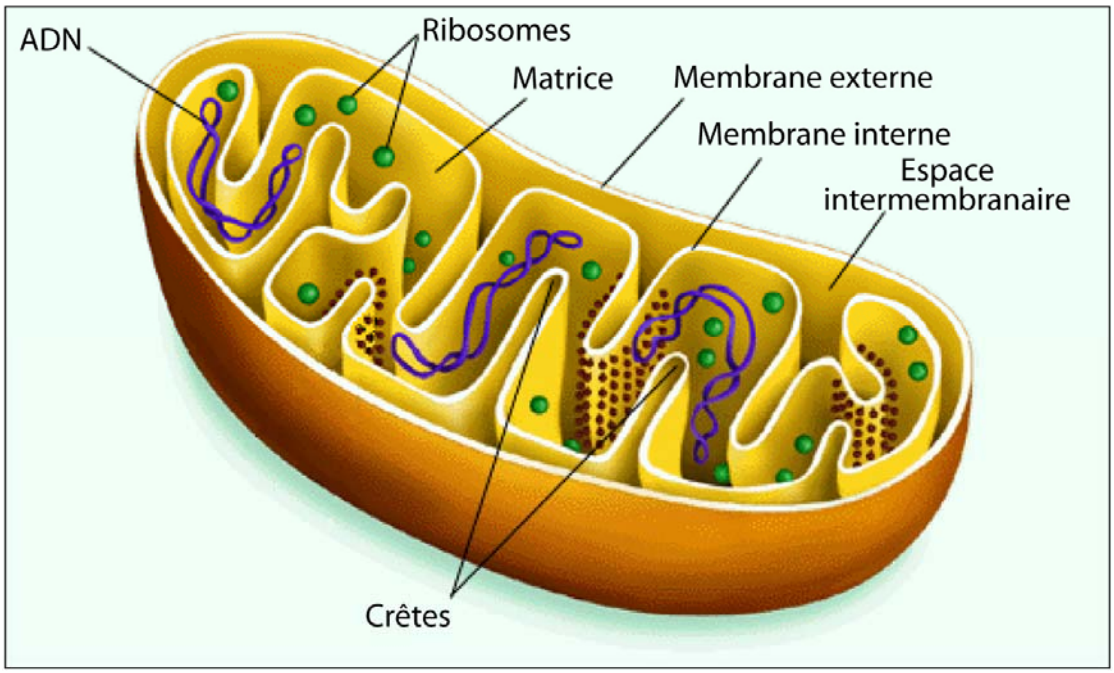



**Figure 13. Représentation schématique de la mitochondrie.** Illustration extraite de Encyclopedia Britanica.

La mitochondrie possède son propre génome et, de façon unique par rapport aux autres organelles, elle peut contenir deux à dix copies d'ADN (ADNmt). L'ADNmt est uniquement transmis par la mère. D'une part, la faible quantité de mitochondries que possède le

spermatozoïde est largement diluée dans l'énorme population mt de l'ovocyte. D'autre part, des mécanismes d'ubiquitinylation seraient impliqués dans la dégradation des mitochondries mâles (Sutovsky et al., 1999).

Le taux de mutations de l'ADNmt est en moyenne 10-20 fois supérieur à celui de l'ADN nucléaire (Allen & Raven, 1996 ; Pesole et al., 1999 ; Saccone et al., 1999). Ceci s'expliquerait par différents facteurs. L'ADNmt n'est pas protégé par des protéines de type histones. Il évolue dans un contexte riche en anions superoxydes mutagènes, rejetés par la chaîne respiratoire. Finalement, les systèmes de réparation et de recombinaison mitochondriaux seraient moins efficaces.

La mitochondrie occupe plusieurs fonctions. En plus de produire de l'ATP, elle est impliquée dans de nombreuses réactions métaboliques essentielles pour la vie de la cellule (par exemple néoglucogenèse, cycle de l'urée, dégradation des acides gras, etc…). Elle joue également un rôle dans les mécanismes de mort cellulaire programmée ou apoptose (Scheffler, 1999).

## 1.3. Expression du génome

L'ADN mt humain est un ADN circulaire bicaténaire de 16 569 pb, dont la séquence complète a été déterminée dès 1981 (Anderson et al., 1981) et révisée récemment (Andrews et al., 1999). Les deux brins n'ont pas la même densité, ce qui permet de distinguer un brin lourd plus riche en guanine, d'un brin léger plus riche en cytosine. L'organisation du génome (Figure 14) témoigne de son origine endosymbiotique et d'une évolution vers un très haut niveau de compaction. En effet, l'ADN mt ne comporte quasiment pas de régions intergéniques et ne possède pas d'introns. Les gènes codant pour des protéines sont généralement ponctués par des gènes d'ARNt et dans un cas extrême, deux gènes se chevauchent. La seule région non codante, appelée boucle D (« displacement loop ») est une région indispensable, puisque nécessaire à la réplication et la transcription du génome.

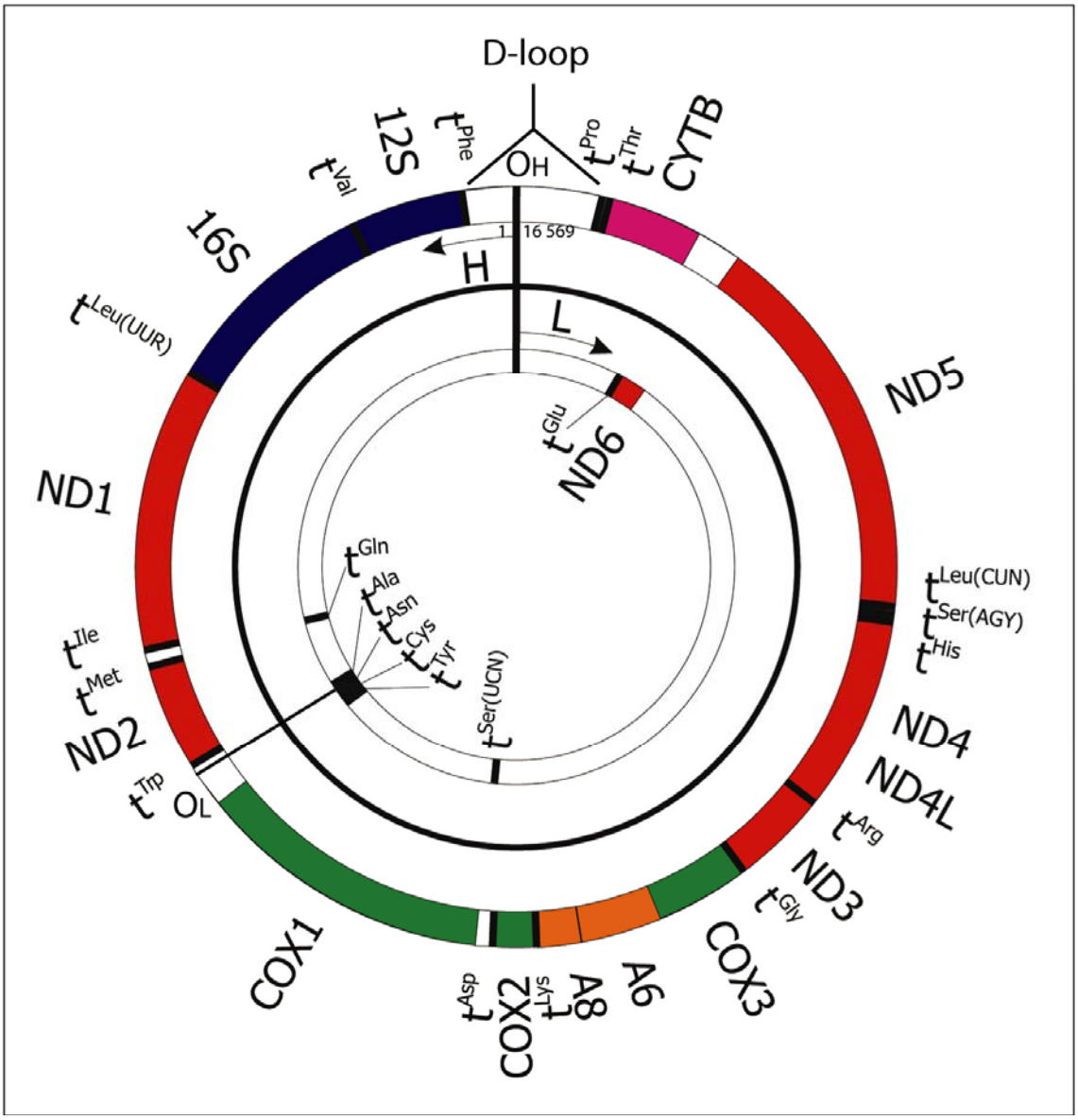



**Figure 14. Organisation du génome mitochondrial humain.** Le brin lourd (H) est représenté par le cercle extérieur riche en couleur, le brin léger (L) par le cercle intérieur blanc. $O_H$ et $O_L$ indiquent les origines de réplication du brin lourd et léger, respectivement. Les gènes d'ARNt sont représentés par des traits épais noirs et indiqués par « $t^{aa}$ ». Les gènes d'ARN ribosomiques 12S et 16S sont indiqués en bleu marine, et les gènes des sous-unités de la chaîne respiratoire sont en rouge pour le complexe I (ND1-6), en rose pour le complexe III (cytB), en vert pour le complexe IV (COX1 et COX3), et en orange pour le complexe V (ATPase A6 et A8). La région non codante « D-loop » est en blanc. Le génome est numéroté (de 1 à 16 569) à partir de l'origine de réplication $O_H$, dans le sens inverse des aiguilles d'une montre.

Le génome mt humain ne contient que 37 gènes codant pour 2 ARN ribosomiques (ARNr), 22 ARNt et pour 13 protéines toutes composantes de la chaîne respiratoire. Les autres sous unités de cette chaîne (environ 80 protéines) et toutes les autres protéines (estimées à 1000-2000) sont importées du cytoplasme. Ainsi, une mitochondrie fonctionnelle

requiert une coordination étroite entre les génomes mt et nucléaire. L'ensemble des informations est schématisé Figure 15.

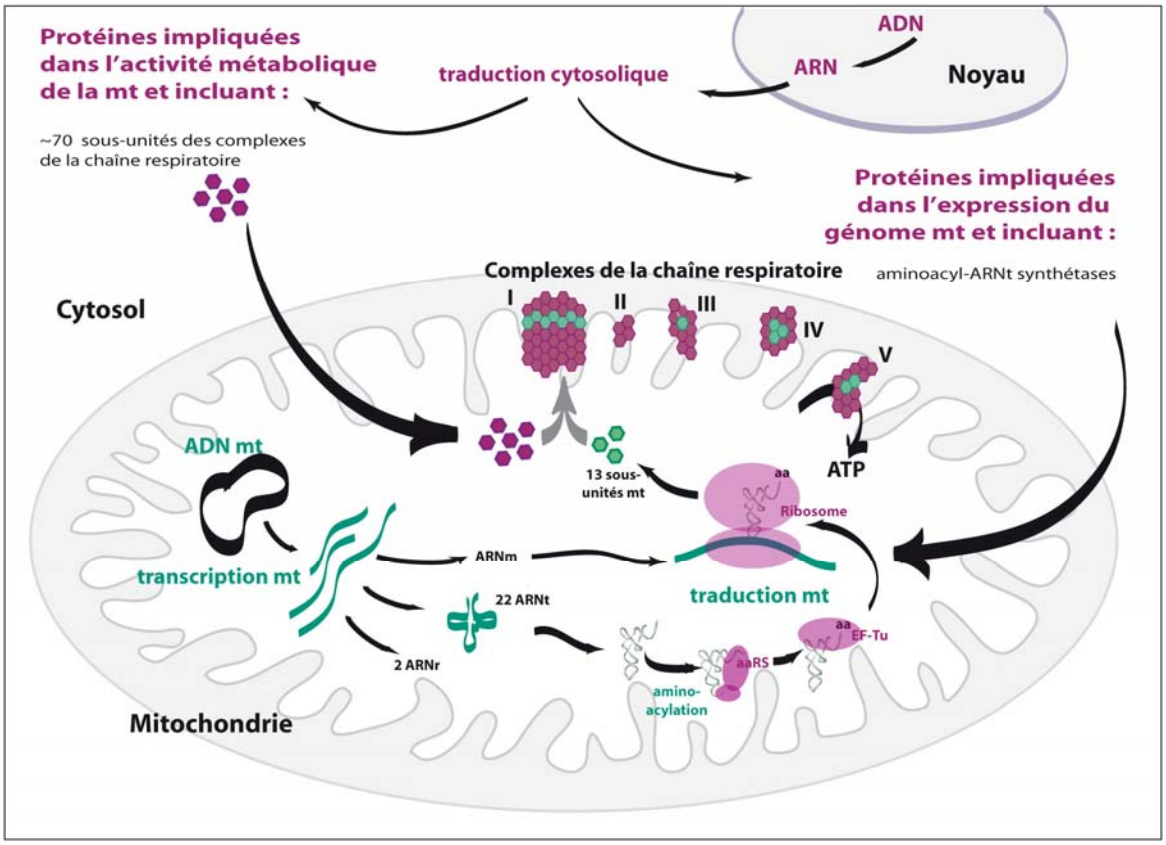



**Figure 15. Origine double de la machinerie traductionnelle mitochondriale.** Le génome mt code pour 22 ARNt et 2 ARNr (indiqués en vert) alors que le génome nucléaire code pour toutes les autres macromolécules impliquées dans la machinerie traductionnelle mt (en violet), en particulier les aaRS. La traduction des 13 gènes mt de protéines conduit à des sous-unités (hexagone vert) des complexes de la chaîne respiratoire. Ceux-ci sont complémentés par ~70 sous-unités additionnelles (hexagones violets) traduites des gènes nucléaires et importés du cytosol.

Chaque molécule d'ADN mt contient deux origines de réplication distinctes, l'une pour le brin lourd (OH) et l'autre pour le brin léger (OL). Deux modes de réplication sont à ce jour discutés : les modes asymétrique et symétrique. Dans le premier mode (Clayton, 1982), le cycle de réplication débute avec la synthèse du brin lourd jusqu'à l'origine de réplication du brin léger localisée au deux tiers de la molécule. Puis la synthèse des deux brins se fait en parallèle et de façon bidirectionnelle. Dans le mode symétrique, la réplication se fait par un mécanisme couplant les deux brins (Holt et al., 2000). Il n'est pas exclu que les deux mécanismes existent et que selon les conditions physiologiques l'un prédomine sur l'autre.

La transcription de l'ADNmt, bidirectionnelle et indépendante de chaque brin, est effectuée par une ARN-polymérase de type T3/T7 et donne lieu à trois ARN polycistroniques.

Deux ARN sont longs et synthétisés à partir des promoteurs HSP2 du brin lourd et LSP du brin léger. Un troisième ARN court est synthétisé à partir du promoteur HSP1 situé juste en amont de HSP2. Les transcrits polycistroniques sont maturés par clivage au niveau des structures en ARNt par la RNase P (en amont), et par une 3'-ARNt endonucléase ou tRNase Z (en aval) (Levinger et al., 2004). Les ARNt ainsi libérés sont modifiés et la séquence CCA, nécessaire à leur aminoacylation, est ajoutée (Levinger et al., 2004). Pour les ARN non ponctués par des ARNt, d'autres enzymes sont impliquées (Masucci & Schon, 1996). Les ARNm sont polyadénylés par une polyA-polymérase, permettant dans certains cas de reconstituer le codon stop partiellement codé. Ils ne possèdent pas de coiffe en 5' (Attardi & Montoya, 1983), ni de séquence particulière, de type Shine-Dalgarno chez les bactéries, en amont du codon d'initiation. Les ARNt et les ARNr sont modifiés post-transcriptionnellement (pour les ARNt, voir plus loin). Les ARNr possèdent globalement moins de nucléotides modifiés (Baer & Dubin, 1981 ; Ofengand & Bakin, 1997).

La traduction des ARNm a été montrée associée à la membrane interne de la mitochondrie (Liu & Spremulli, 2000). Bien que possédant des traits typiquement bactériens (elle débute par un N-formyl-méthionine-ARNt et est inhibée par le chloramphénicol) la traduction mt présente des caractéristiques qui lui sont propres. Le code génétique mt possède quelques variations par rapport au code génétique universel, variations qui peuvent être différentes selon les organismes. Dans le génome mt de mammifères, le codon stop UGA est décodé par l'ARNt$^{Trp}$, les codons Ile AUA et AUU par l'ARNt$^{Met}$ et les codons Arg AGA et AGG sont des codons stop (il y a donc quatre codons stop au lieu de trois, et il n'y a plus que quatre codons Arg au lieu de six). La lecture des codons requiert aussi une plus grande flexibilité d'interaction entre les ARNt et l'ARNm puisqu'il n'y a que 22 ARNt pour lire 60 codons. Cette flexibilité serait générée soit par une interaction codon/anticodon limitée à deux paires de bases ou à de plus nombreuses possibilités d'interaction du nucléotide en troisième position de l'anticodon de l'ARNt (revue par (Florentz et al., 2003)). Une autre particularité de la traduction mt vient de son ribosome. En effet, le ribosome mt de mammifère est composé d'une petite sous unité 28S et d'une grande sous unité 39S. Il est caractérisé par une composition plus faible en ARN (seulement 33% contre 66% dans le ribosome bactérien) et notamment par la perte de certaines régions structurales composées d'ARN qui pourrait être compensée par une plus forte teneur en protéines (revue par (O'Brien, 2003)). L'initiation de la traduction ne repose ni sur une séquence de type Shine-Dalgarno (bactéries), ni sur la coiffe en 5' (eucaryotes).

# 2. Les systèmes d'aminoacylation mitochondriaux

La machinerie traductionnelle dans la mitochondrie est de deux origines (Figure 15), avec les ARNt et les ARNr codés par le génome mt, les autres partenaires codés par le génome nucléaire et importés dans la mitochondrie. En particulier les aaRS mt sont toutes codées par le génome nucléaire et importées dans l'organelle. Ceci rend le problème de la spécificité d'aminoacylation tout à fait intéressant. De plus, le génome mt possède un taux de mutation environ 10 à 20 fois supérieur à celui du génome nucléaire (Allen & Raven, 1996 ; Pesole et al., 1999 ; Saccone et al., 1999). Ce facteur est singulièrement augmenté pour les gènes d'ARNt qui sont mutés jusqu'à 100 fois plus que leurs homologues cytoplasmiques (Saccone et al., 1999). En conséquence, alors que les gènes d'aaRS évoluent peu et ressemblent aux gènes d'aaRS connues (revue par (Sissler et al., 2005)), les ARNt mt de mammifères apparaissent dégénérés par rapport aux ARNt classiques (revue par (Florentz et al., 2003)). Ceci soulève de nombreuses questions concernant les mécanismes de reconnaissance entre aaRS et ARNt, et notamment en ce qui concerne les signaux d'identité, mais aussi concernant la co-évolution entre deux molécules codées par deux génomes différents.

## 2.1. Les ARNt mitochondriaux

Le génome mt de mammifère code pour 22 ARNt, à savoir un ARNt par acide aminé à l'exception de la leucine et de la sérine pour lesquels il y en a deux (Anderson et al., 1981). Alors que l'import d'ARNt cytoplasmiques est courant dans les mitochondries de plantes, de la levure et des protozoaires, il n'existe à ce jour qu'un seul exemple d'importation d'ARNt dans les mitochondries de mammifères. En effet, pour au moins 11 marsupiaux différents, le gène de l'ARNt$^{Lys}$ mt est en fait un pseudogène et l'ARNt$^{Lys}$ fonctionnel est importé du cytoplasme (Dörner et al., 2001).

Les gènes d'ARNt sont répartis tout au long du génome mt, avec 14 gènes transcrits à partir du brin lourd (riche en G) et conduisant à la production d'ARNt « légers », et les huit autres transcrits à partir du brin léger et conduisant à des ARNt « lourds ». De ce fait, la composition nucléotidique des ARNt mt est biaisée avec des ARNt « légers » pauvres en G et riches en C, et des ARNt « lourds » riches en G et pauvres en C. Un nombre plus élevé de

dinucléotides CpA et UpA (séquences très sensibles à la dégradation (Dock-Bregeon et al., 1989) se retrouvent dans les séquences des ARNt « légers », alors que les ARNt « lourds » se caractérisent par une plus grande proportion de paires de bases G•U (Florentz & Sissler, 2003).

### *Maturation*

Tout comme les ARNt classiques, les ARNt mt subissent des étapes de maturation. Leurs extrémités 5' et 3' sont maturés par la RNase P et la RNAseZ (Levinger et al., 2004). Les ARNt ainsi libérés sont modifiés et la séquence CCA, nécessaire à leur aminoacylation, est ajoutée (Levinger et al., 2004). Actuellement, environ 40 ARNt mt ont été séquencés au niveau ARN, mettant en évidence 14 modifications différentes (Sprinzl & Vassilenko, 2005). En ce qui concerne les ARNt mt humains, les modifications ont été déterminées pour 6 ARNt (Lys, Ile, Leu(UUR), Ser(UCN), Ser(AGY) et Pro). La plupart de ces modifications sont celles retrouvées dans d'autres ARNt (*e.g.* $m^1G$, T, ψ, D, etc…). Toutefois, il ne sera pas étonnant d'en découvrir de toutes nouvelles, à l'exemple du dérivé de la taurineméthyl spécifique du résidu U34 des $ARNt^{Lys}$ (Yasukawa et al., 2000b) et $ARNt^{Leu}$ (Yasukawa et al., 2000c) mt humains. Par ailleurs, certaines modifications classiques sont présentes à de nouvelles positions comme la méthylation de l'Adénine en position 9 (connecteur entre les bras accepteur et D) qui n'est retrouvée qu'au sein de l'$ARNt^{Lys}$ mt, et qui est nécessaire au repliement de la molécule en feuille de trèfle (Helm et al., 1999). Globalement, les ARNt mt possèdent moins de modifications post-transcriptionnelles que leurs homologues cytoplasmiques, suggérant que celles présentes ont un rôle important. Par exemple, un défaut de modification en position 34 de l'$ARNt^{Leu}$ est corrélé à une pathologie mt humaine (Kirino et al., 2005). Les ARNt mt de mammifères peuvent aussi être édités. C'est le cas de l'édition de l'$ARNt^{Gly}$ mt de marsupiaux dont le C en position 34 de l'anticodon est partiellement édité en U pour générer l'$ARNt^{Asp}$ qui n'est pas codé par le génome mt (Mörl et al., 1995).

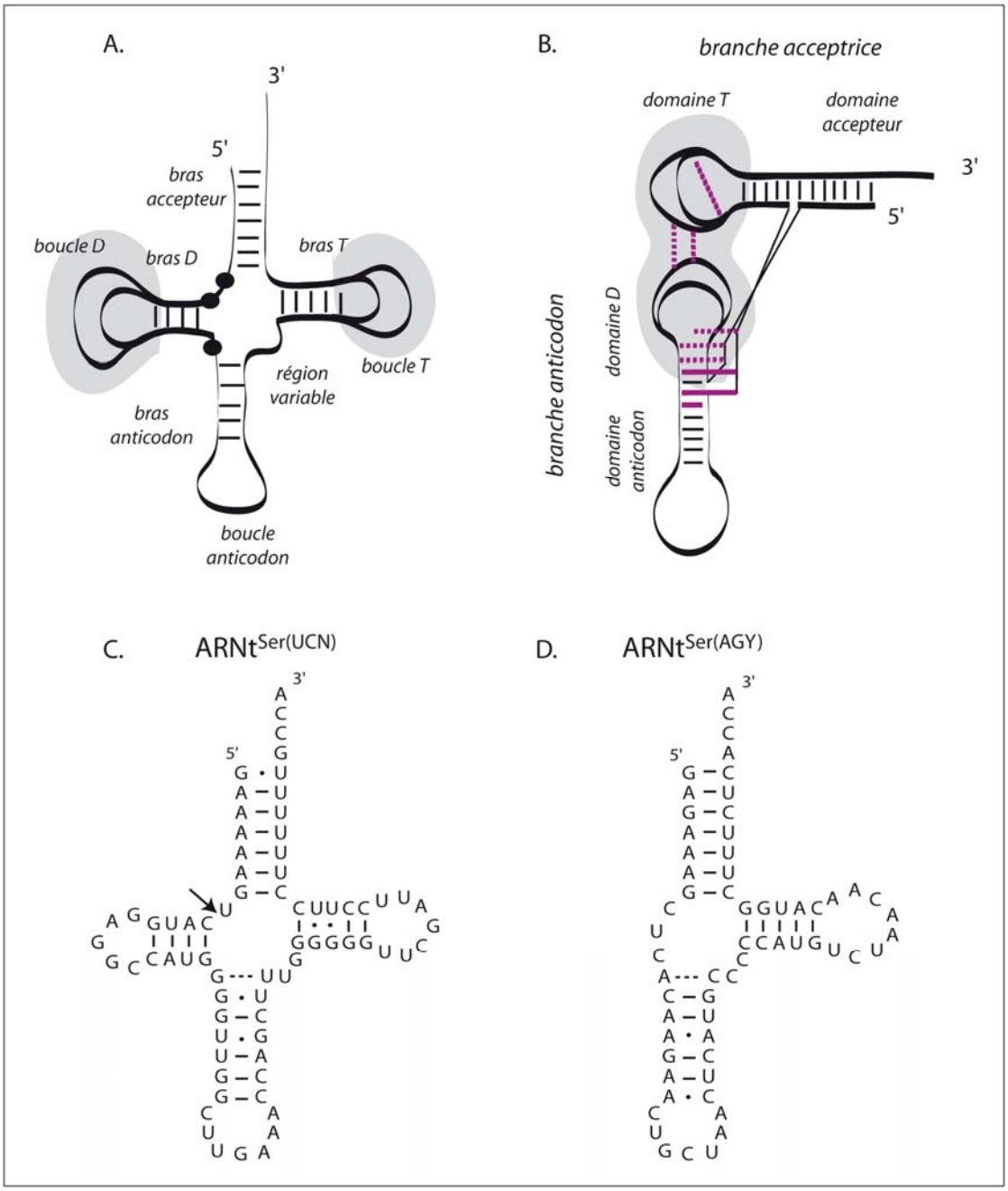



**Figure 16. Propriétés structurales des ARNt mitochondriaux humains.** Représentation schématique des structures secondaire en feuille de trèfle (**A.**) ou tertiaire (**B.**) de 20 des 22 ARNt mt humains. Les différents domaines sont indiqués. Les régions grisées mettent en évidence la grande variation des tailles des boucles D et T. En **B.**, les interactions à longue distance sont indiquées par des traits pleins, ou en pointillé pour celles non strictement conservées. **C.** Cas de l'ARNt$^{Ser(UCN)}$ pour lequel le connecteur 1 est réduit à un nucléotide (flèche) et le bras anticodon est étendu. **D.** Cas de l'ARNt$^{Ser(AGY)}$ pour lequel le bras D est absent.

### *Structures secondaires et tertiaires*

En plus d'être plus petits que les ARNt classiques (de 62 à 78 contre 75 à 95 pour les ARNt classiques), les ARNt mt présentent des caractéristiques architecturales qui les distinguent et participent largement à leur qualification d'ARNt « bizarres ». Les exemples les plus marquants sont l'ARNt$^{Ser(UCN)}$ mt qui comporte un bras anticodon étendu (6 paires de bases au lieu de 5) et un connecteur 1 plus court, et l'ARNt$^{Ser(AGY)}$ qui ne possède pas de bras D (Figure 16.C&D.). Les 20 autres se replient, du moins théoriquement, en une feuille de trèfle. Cependant, ils possèdent tous des variations importantes au niveau des tailles des boucles D et T (Figure 16.A&B). Alors que la longueur d'une boucle D d'ARNt classique peut être de 5 à 7 nucléotides, celle des ARNt mt peut être réduite à 1 ou étendue à 10 nucléotides. La taille de la boucle T, qui est invariable dans les ARNt classiques (7 nucléotides), peut comprendre de 5 à 9 résidus. En revanche, la région variable des ARNt mt varie très peu (4-5 nucléotides) par rapport à celle des ARNt classiques qui peut comprendre jusqu'à 24 résidus et former un bras supplémentaire (*e.g.* cas de l'ARNt$^{Ser}$ d'*E. coli*). Une autre particularité vient directement de la séquence et de la composition biaisée en nucléotides des ARNt mt qui comportent plus de paires de bases G•U et A-U, voir des mésappariements, et très peu de paires de bases G-C. Ceci conduit à des ARNt globalement plus flexibles.

Le repliement tridimensionnel des ARNt mt est mal compris. En effet la recherche de résidus stratégiques aux positions impliquées dans le réseau d'interaction tertiaire dans les ARNt classiques a révélé une grande souplesse. Notamment les résidus G18G19 et T54ψ55C56 impliqués dans l'interaction des boucles D et T des ARNt classiques, ne sont retrouvés que dans trois des 22 ARNt mt (ARNt$^{Leu(UUR)}$, ARNt$^{Ser(UCN)}$ et ARNt$^{Gln}$). Récemment, (Ishitani et al., 2003) ont mis en évidence un changement structural drastique dans un ARNt classique lors de la reconnaissance par une enzyme de modification. La forme initiale en L de l'ARNt se replie en une forme alternative en λ (Figure 17). Il est tentant de penser que les ARNt mt pourraient adopter ce type de structure. Malheureusement, le faible nombre d'études structurales des ARNt mt ne permet pas encore d'ébauche de cette ou ces structures.

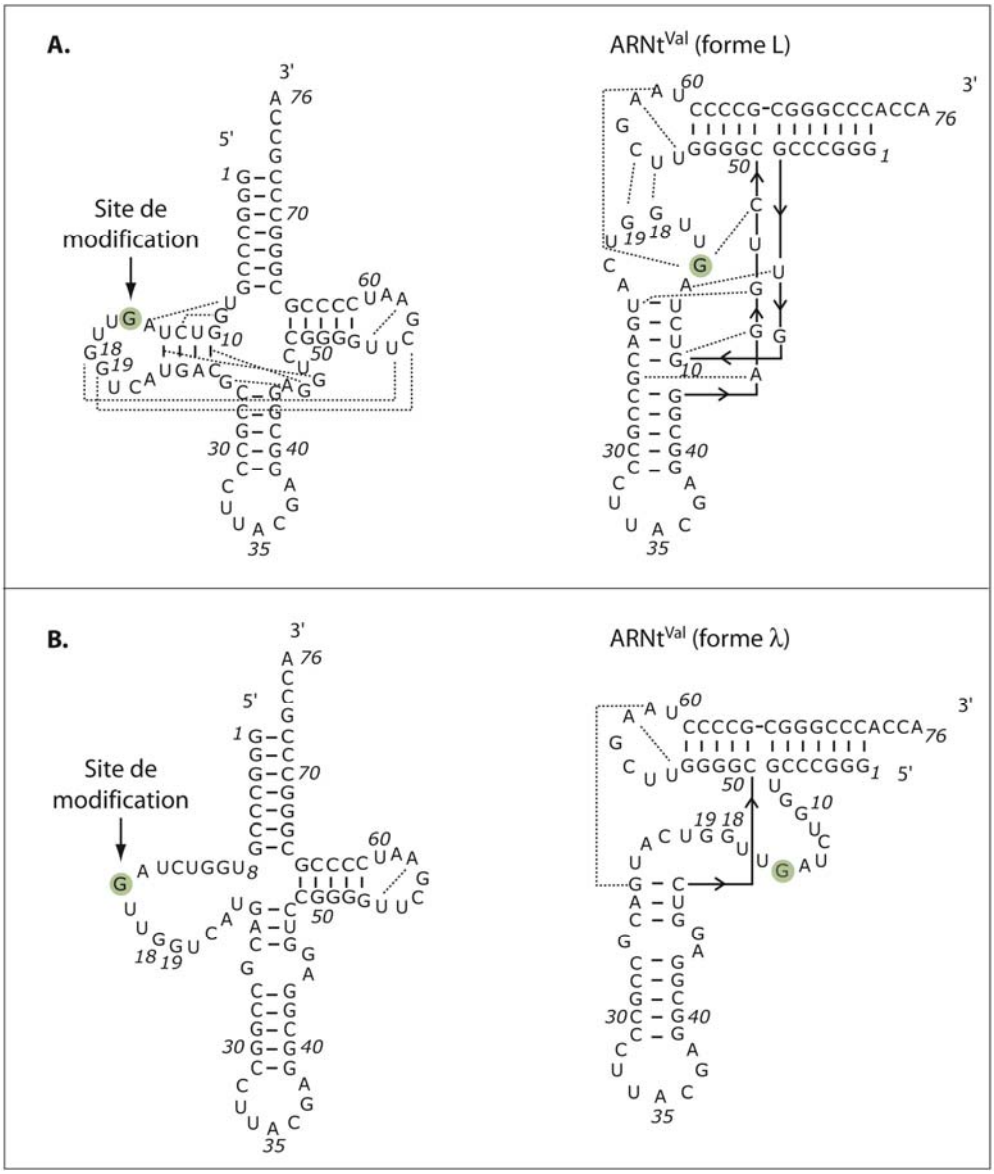



**Figure 17. Repliement alternatif de l'ARNt en forme λ. A.** Structures canoniques, secondaire en feuille de trèfle et tertiaire en L de l'ARNt$^{Val}$ de *Pyrococcus horikoshii*. **B.** Structures alternatives, secondaire et tertiaire en λ induite par l'interaction avec l'enzyme qui modifie le nucléotide G15 (en vert) (Ishitani et al., 2003). Figure extraite de (Schimmel & Tamura, 2003).

### *Rôle alternatif*

Dans le génome mt, les gènes d'ARNt ponctuent les gènes codant pour des protéines. Après transcription des polycistrons, les ARNm sont libérés grâce à la maturation des ARNt. Tout récemment, l'ARNt$^{Met}$ mt humain a été détecté associé à une des protéines du complexe ribonucléoprotéique RISC (« RNA Induced Silencing Complex), suggèrant un rôle potentiel de l'ARNt dans la voie de dégradation des ARNm par ARN interférence (Maniataki & Mourelatos, 2005).

### *Mutations des ARNt : polymorphisme* versus *pathologie*

Le génome mt évolue rapidement : il subit un taux de mutations 10 à 20 fois supérieur à celui du génome nucléaire (Allen & Raven, 1996 ; Pesole et al., 1999 ; Saccone et al., 1999). Alors que certaines mutations sont neutres, les mutations polymorphiques, et contribuent à la diversité des populations humaines (*e.g.* (Krings et al., 1997 ; Ingman et al., 2000), beaucoup sont corrélées à des maladies (DiMauro & Moraes, 1993 ; Larsson & Clayton, 1995 ; Schon et al., 1997 ; Dimauro & Schon, 2001). Plus de 115 pathologies corrélées à des mutations de l'ADN mt ont été recensées dans la banque MITOMAP (Kogelnik et al., 1998), dont 38 % sont liées à des mutations dans les gènes de protéines, 4 %dans les gènes des ARNr et 58 % ARNt.

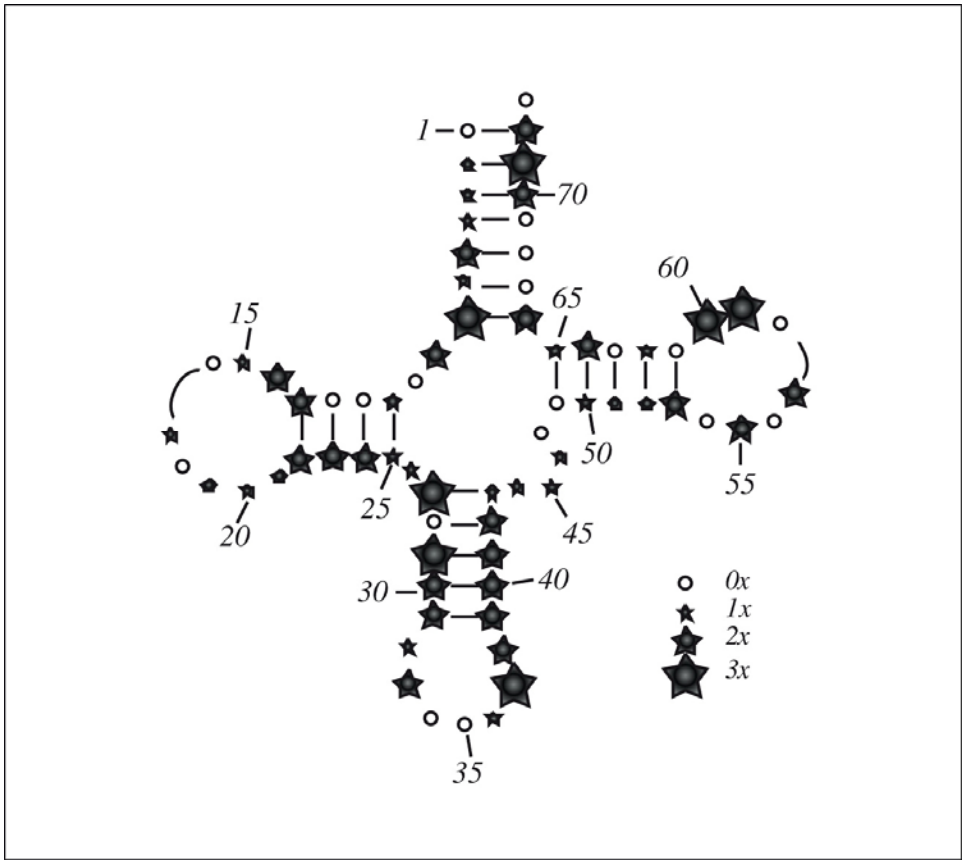



**Figure 18. Distribution des mutations pathologiques sur la structure en feuille de trèfle de l'ARNt.** La taille des étoiles est proportionnelle à l'occupation des sites. Les informations proviennent de la base de données MITOMAP (http://www.mitomap.org/).

Parmi les 22 ARNt mt humains, 21 présentent des mutations corrélées à des pathologies (l'exception étant l'ARNt$^{Arg}$), les ARNt$^{Leu(UUR)}$, ARNt$^{Lys}$ et ARNt$^{Ile}$ constituant des points chauds revue par (Florentz et al., 2003). La superposition de l'ensemble des mutations « pathogéniques » sur une même structure en feuille de trèfle, montre qu'elles

touchent des positions très variées (Figure 18). Leurs effets sur le fonctionnement des ARNt dans la synthèse protéique sont potentiellement nombreux, puisque ces effets peuvent interférer soit avec l'une des nombreuses étapes de la biogenèse de l'ARNt, soit avec l'un ou l'autre de sa fonction (interaction avec l'aaRS, facteurs d'initiation ou d'élongation, protéines ribosomiques, ARNt, interaction anticodon/codon avec l'ARNm, etc…). Certaines mutations affectent la mise en place des modifications post-transcriptionnelles (Brulé et al., 1998 ; Yasukawa et al., 2000c ; Kirino et al., 2004 ; Kirino et al., 2005), la stabilité (Hao & Moraes, 1997 ; Yasukawa et al., 2000a), ou encore la maturation en 3' (Levinger et al., 2001 ; Levinger et al., 2004). Enfin, plusieurs mutations affectent directement les capacités d'aminoacylation des ARNt *in vivo*. Par exemple dans le cas de la mutation A8344G (mutation dans le gène de l'ARNt$^{Lys}$ et corrélée au symptôme MERRF, pour « Myoclinic Epilepsy with Ragged Red Fibers »), seuls 40 % de lysyl-ARNt$^{Lys}$ restent disponibles pour la traduction dans les mitochondries malades (Enriquez et al., 1995 ; Enriquez & Attardi, 1996). Une situation comparable se présente dans le cas de la mutation MELAS (« Mitochondrial Epilepsy with Lactic Acidosis and Stroke-like episodes ») dans l'ARNt$^{Leu(UUR)}$ avec une diminution du taux de leucyl-ARNt$^{Leu(UUR)}$ de 50 % (Chomyn et al., 2000). Des études de capacité et d'efficacité d'aminoacylation menées au laboratoire sur des couples d'ARNt transcrits *in vitro* aboutissent à des conclusions similaires, certaines mutations affectant l'aminoacylation (Degoul et al., 1998 ; Kelley et al., 2000 ; Sissler et al., 2004). La question des effets moléculaires des diverses mutations est décrite en détail dans une revue (Levinger et al., 2004).

## 2.2. Les aminoacyl-ARNt synthétases mitochondriales

### *Séquences et structures*

Les connaissances accumulées sur les aaRS mt de mamifères restent encore limitées (revue par (Sissler et al., 2005)). Les premières études biochimiques, initiées dès les années 60 sur des extraits bruts enzymatiques à partir de foies de rat et bœuf, de placenta humain ou de cellules HeLa, ont permis de mettre en évidence les activités de la plupart des activités synthétasiques (Buck & Nass, 1969 ; Lynch & Attardi, 1976 ; Kumazawa et al., 1991). Ce n'est qu'en 1994 que les premiers gènes d'aaRS mt humaines (GlyRS et IleRS) ont été décrits (Shiba et al., 1994a ; Shiba et al., 1994b). Ces gènes ont été recherchés par comparaison de

séquences dans les banques de données sur la base d'homologies structurales avec les enzymes d'autres organismes. Depuis, 6 autres gènes ont été identifiés par la même approche. Finalement, on dispose d'informations au niveau du gène pour les GlyRS et HisRS (O'Hanlon et al., 1995 ; O'Hanlon & Miller, 2002), et les gènes des IleRS (Kelley et al., 2000), PheRS (Bullard et al., 1999), LysRS (Tolkunova et al., 2000), LeuRS (Bullard et al., 2000), TrpRS (Jørgensen et al., 2000), SerRS (Yokogawa et al., 2000) et MetRS (Spencer et al., 2004) ont été clonés et les protéines correspondantes caractérisées. Le Tableau 3, extrait de (Sissler et al., 2005), indique leurs principales caratéristiques. Les aaRS mt sont toutes codées par le génome nucléaire, synthétisées dans le cytoplasme et importées dans la mitochondrie. Leur import est dirigé par un signal d'adressage généralement présent en N-terminal, riche en acides aminés chargés positivement et formant une hélice α amphiphile (Hammen & Weiner, 1998 ; Pfanner, 2000). Après import, la séquence d'adressage est clivée. La plupart de ces enzymes mt sont codées par des gènes distincts de ceux des enzymes cytoplasmiques (cyt) correspondantes (HisRS, IleRS, LeuRS, MetRS, PheRS et TrpRS). Dans le cas des GlyRS et LysRS, un même gène code pour les enzymes mt et cyt. Les GlyRS mt et cyt sont traduites à partir de deux sites d'initiation, générant ainsi une enzyme possédant une séquence d'adressage (mt) et l'autre non (cyt) (Shiba et al., 1994a ; Mudge et al., 1998). En ce qui concerne les LysRS, un mécanisme d'épissage alternatif permet l'insertion ou non d'un signal d'adressage dans la protéine mature (Tolkunova et al., 2000). Ainsi les LysRS et GlyRS mt matures sont quasiment identiques à leurs homologues cyt. Les autres aaRS mt ressemblent plutôt aux enzymes de type procaryotique, avec l'exception de l'HisRS mt qui partage plus de 70% d'identité de séquence avec l'enzyme cyt (Sissler et al., 2005).

Les caractéristiques structurales des aaRS mt humaines ont été comparées à celles des autres aaRS connues (Sissler et al., 2005). Les tailles et l'oligomérie des enzymes mt sont globalement les mêmes que ceux de leurs homologues dans d'autres organismes. Notons les exceptions de la GlyRS qui est dimérique ($\alpha 2$, cas également d'autres GlyRS mt et des GlyRS cytoplasmiques) alors que les enzymes procaryotiques sont de type $\alpha 2\beta 2$, et de la PheRS qui est monomérique ($\alpha$, cas également de la PheRS mt de levure (Sanni et al., 1991)) alors que les autres PheRS connues sont de type $\alpha 2\beta 2$. Toutes les aaRS mt humaines possèdent les motifs caractéristiques permettant de les répartir dans les 2 classes d'aaRS (définies dans la partie A.2.).

| | Taille de la séquence d'adressage (aa) | Taille de l'aaRS mature (aa) | Oligomérie | Classe d'aaRS | Référence |
|---|---|---|---|---|---|
| **GlyRS** | 54 | 685 | α2 | IIa | (Shiba et al., 1994a) ; Mudge, 1998 #5968} |
| **HisRS** | Probablement en N-terminale | 506 | α2 ? | IIa | (O'Hanlon et al., 1995) |
| **IleRS** | ? | 993 | α ? | Ia | {Shiba, 1994 #2091 |
| **LeuRS** | 39 | 903 | α2 | Ia | (Bullard et al., 2000) |
| **LysRS** | 16 | 625 | α2 | IIb | (Tolkunova et al., 2000) |
| **MetRS** | 18 | 575 | α | Ia | (Spencer et al., 2004) |
| **PheRS** | 37 | 451 | α | IIc | (Bullard et al., 1999) |
| **SerRS** | 20 | 498 | α2 | IIa | (Yokogawa et al., 2000) |
| **TrpRS** | 18 | 360 | α2 | Ib | (Jørgensen et al., 2000) |

**Tableau 3. Caractéristiques des aminoacyl-ARNt synthétases mt humaines.** L'oligomérie des HisRS et IleRS est prédictive et indiquée à partir de celle des enzymes connues (Cusack et al., 1990; Eriani et al., 1990 ).

Notons que la structure cristallographique de la SerRS mt bovine en complexe avec le séryl-adénylate a été résolue tout récemment (Chimnaronk et al., 2005). C'est à ce jour, la seule structure d'une synthétase mt dont nous disposons. La structure met en évidence deux régions uniquement retrouvées dans la SerRS mt dans les parties N- et C-terminales. Des expériences complémentaires de mutagénèse de ces régions et d'aminoacylation *in vitro*, ont montré qu'elles sont impliquées dans la reconnaissance de l'ARNt. De plus, la SerRS pourrait reconnaître les deux $ARNt^{Ser}$ mt (structuralement différents, voir Figure 16) par l'utilisation de combinaisons distinctes d'acides aminés au niveau de ces régions.

### *Propriétés d'aminoacylation*

De façon globale, les activités des aaRS mt humaines sont moins bonnes que celles des enzymes bactériennes correspondantes. Par exemple, les PheRS, TrpRS et LeuRS mt humaines possèdent des efficacités catalytiques réduites de ~100 fois pour l'acide aminé et ~250 fois pour l'ATP (Bullard et al., 1999 ; Bullard et al., 2000; Jørgensen et al., 2000 ). De plus, les capacités d'aminoacylation estimées en présence d'ARNt totaux d'*E. coli*, montrent que l'activité spécifique de la LeuRS est ~250 à 400 fois inférieure à celle observée pour les

enzymes d'origine bactérienne et cytoplasmique (Bullard et al., 1999), et que celle de la PheRS est ~20 à 30 fois moins bonne (Bullard et al., 1999). En conclusion, les aaRS mt apparaissent comme des enzymes moins efficaces que leurs homologues cyt ou bactériens.

Des test d'aminoacylation entre les systèmes mt et cyt de mammifères, réalisés à partir d'extraits de foie de rat (Buck & Nass, 1969), de cellules HeLa (Lynch & Attardi, 1976) ou de mitochondries de bœuf (Kumazawa et al., 1991) ont montré que les aaRS mt de mammifères sont capables d'aminoacyler les ARNt cytoplasmiques et d'*E. coli*, alors que leurs homologues cyt et bactériens ne sont pas capables de reconnaître les ARNt mt. Cette unilatéralité d'aminoacylation pourrait être due à la grande divergence structurale des ARNt mt par rapport aux ARNt canoniques, qui ainsi empêcherait la reconnaissance par des enzymes d'origines différentes (hétérologues).

### *Fonctions alternatives*

Des études réalisées sur la levure, ont montré l'implication des aaRS mt dans des processus tels que l'import d'ARNt (cas de l'ARNt$^{Lys}$ importé grâce à la LysRS mt, (Tarassov & Martin, 1996)) ou l'épissage des introns (cas de la LeuRS mt, (Labouesse, 1990)). La question de savoir si les aaRS mt de mammifères étaient capables de substituer les enzymes de levure pour ces fonctions a été abordée. Ainsi, par des expériences *in vitro*, (Entelis et al., 2001) ont prouvé qu'il est possible d'importer des ARNt dans la mitochondrie humaine, montrant que la cellule possède tous les facteurs requis pour l'importation des ARNt. Notons qu'il n'y a pas d'importation d'ARNt chez l'Homme, mais que l'ARNt$^{Lys}$ cytoplasmique de marsupiaux est importé dans la mitochondrie (Dörner et al., 2001). Cependant l'implication de la LysRS mt n'est pas encore démontrée.

En ce qui concerne l'épissage, bien que le génome mt humain soit dépourvu d'intron, la LeuRS mt humaine est capable de substituer l'enzyme de levure pour cette fonction (Houman et al., 2000 ; Rho & Martinis, 2000).

# D. Objectifs de la thèse

L'objectif de ce travail sera d'explorer les originalités et subtilités des mécanismes sous-jacents à la spécificité d'aminoacylation des ARNt par les aaRS, tant par l'étude de systèmes « classiques », à savoir du cytoplasme de la levure *S. cerevisiae*, que par celle de systèmes « atypiques », à savoir ceux de la mitochondrie humaine.

Tout d'abord, les règles de l'identité pour les systèmes d'aminoacylation spécifiques de l'arginine et de l'aspartate chez la levure seront explorés plus en avant. L'étude conjointe de ces deux systèmes va nous permettre d'approcher de façon originale l'identité au sein de la « famille nombreuse » isoacceptrice de l'arginine qui compte 4 ARNt. Les 4 $ARNt^{Arg}$ étant de même spécificité, il était jusqu'alors admis qu'ils possédaient les mêmes caractéristiques fonctionnelles. L'objectif sera de vérifier cette hypothèse par l'analyse des activités relatives des $ARNt^{Arg}$ transcrits *in vitro* en présence de leur synthétase homologue de levure (l'ArgRS), mais aussi en présence de l'AspRS de levure, enzyme non homologue. La question est alors de savoir si les $ARNt^{Arg}$ sont fonctionnellement équivalents car tous reconnus par la même enzyme, ou s'ils possèdent des comportements différents dus aux variations de séquences et d'architecture qui distinguent chacune de ces molécules.

De façon originale, les systèmes d'arginylation et d'aspartylation de levure sont fonctionnellement reliés, avec la mise en évidence précédemment au laboratoire de l'arginylation efficace du transcrit *in vitro* de l'$ARNt^{Asp}$. L'existence d'un isoaccepteur de l'arginine ($ARNt_4^{Arg}$) aux ressemblances frappantes avec le seul accepteur $ARNt^{Asp}$ permet d'élargir l'investigation de la relation entre ces deux systèmes. En particulier, les propriétés d'aspartylation de l'$ARNt_4^{Arg}$ sauvage ou muté, transcrit *in vitro* ou exprimé dans la levure, seront explorées et notamment la question d'une relation évolutive entre les gènes $ARNt_4^{Arg}$ et $ARNt^{Asp}$ abordée.

Un autre grand axe de recherche portera sur l'étude des mécanismes sous-jacents à la reconnaissance spécifique des ARNt mt de mammifères par leurs aaRS, et notamment ceux humains. Les systèmes d'aminoacylation mt de mammifères sont originaux car leurs ARNt, codés par le génome mt, possèdent des structures dégénérées alors que les aaRS mt, codées par le génome nucléaire, ressemblent aux synthétases classiques. Ceci laisse présager de nouvelles règles de reconnaissance entre ces deux macromolécules et en particulier de nouveaux signaux d'identité. L'accumulation de connaissances fondamentales permettra de

mieux comprendre ces systèmes atypiques et d'en évaluer les relations évolutives avec les systèmes bactériens dont ils sont issus.

Nous avons choisi d'explorer plus en avant le système d'aspartylation mt humain comme modèle d'étude. Ce système est de choix pour plusieurs raisons : (i) l'ARNt$^{Asp}$ mt possède une séquence/structure bizarre par rapport aux ARNt$^{Asp}$ canoniques, (ii) il semble posséder un jeu d'identité différent et (iii) les systèmes aspartate de levure et bactériens (*E. coli* et *T. thermophilus*) étant bien connus, il sera possible de réaliser une analyse comparative des systèmes d'aspartylation et ainsi d'aborder la question de leur évolution.
Ainsi, les éléments d'identité seront recherchés au sein de l'ARNt$^{Asp}$ et les particularités de l'AspRS permettant l'adaptation à son ARNt bizarre explorées. La disponibilité d'AspRS d'organismes divers (clonées, surproduites et purifiées précédemment au laboratoire) va nous permettre d'étudier les raisons de la non reconnaissance des ARNt mt de mammifères par des aaRS d'autres origines, question posée depuis plus d'une décennie. L'ensemble des travaux réalisés sur le système d'aspartylation mt humain sera mené par des approches *in vitro*, basées sur l'utilisation d'ARNt transcrits *in vitro* et de l'AspRS mt clonée, surproduite et purifiée. Une approche structurale basée sur des essais de cristallisation de l'AspRS mt humaine sera également entreprise.

CHAPITRE 1

# Etude des systèmes d'aminoacylation arginine et aspartate chez la levure

# A. Particularités d'une famille isoacceptrice

Dans une cellule, il existe en général plusieurs ARNt isoaccepteurs pour un même acide aminé. Etant tous de même spécificité, il est aisé de penser qu'ils sont tous reconnus et aminoacylés avec la même efficacité par l'aaRS homologue. Nous avons choisi de vérifier ceci au sein de la famille isoacceptrice de l'arginine de *S. cerevisiae*.

L'arginine est, avec la leucine et la sérine, un des rares acides aminés codés par 6 codons différents. Les codons arginine sont CG(A/C/G/U) et AG(G/A). La lecture de ces codons nécessite plusieurs ARNt isoaccepteurs. Le système arginine de *S. cerevisiae* compte 4 ARNt isoaccepteurs différents capables de lire les 6 codons grâce à l'interaction wobble. Le nombre de copies de leurs gènes est en accord avec leur appellation d'isoaccepteurs « majeurs » ou « mineurs » : 6 et 11 copies respectivement pour les $ARNt_2^{Arg}$ et $ARNt_3^{Arg}$, et 1 seule copie pour les $ARNt_1^{Arg}$ et $ARNt_4^{Arg}$. Les 4 ARNt partagent peu de similarité de séquences entre eux (seulement 33 % d'identité de séquence), et les tailles de leurs régions $\alpha$ et $\beta$ (de part et d'autre des nucléotides G18 et G19 de la boucle D) et des régions variables sont différentes (voir figure 1 de l'article n°1).

Le système arginine de levure bénéficie déjà de larges investigations tant fonctionnelles que structurales. Ainsi, par exemple les 2 ARNt majeurs ont été purifiés et séquencés, le gène de l'ArgRS cloné et l'enzyme surproduite (Sissler et al., 1997), le complexe $ARNt_2^{Arg}$/ArgRS cristallisé et sa structure résolue (Delagoutte et al., 2000). De plus, les éléments d'identité arginine ont été (i) déterminés au sein du transcrit de l'$ARNt_3^{Arg}$ et sont les résidus $C_{35}$ et G ou $U_{36}$ de l'anticodon (Sissler et al., 1996), et (ii) explorés *in vivo* au sein de l'$ARNt_4^{Arg}$ (Geslain et al., 2003). Cependant, aucune analyse comparative de l'ensemble des 4 $ARNt^{Arg}$ n'a été réalisée.

La question était alors de savoir si les différents isoaccepteurs $ARNt^{Arg}$ sont fonctionnellement équivalents étant tous reconnus par la même enzyme, ou s'ils possèdent des comportements différents dus aux variations de séquences et d'architecture qui distinguent chacune de ces molécules. Pour ce faire, nous avons analysé les activités relatives des 4 $ARNt^{Arg}$ tant en présence de l'ArgRS qu'en présence de l'AspRS non homologue. L'ensemble des résultats obtenus est décrit et illustré dans l'article n°1.

Article n°1

# Idiosyncrasies fonctionnelles d'ARNt isoaccepteurs dans les systèmes d'aminoacylation spécifique et non spécifique

**A. Fender, M. Sissler, C. Florentz & R. Giegé**

**2004**

**Biochimie, 86 : 21-29**

**Résumé :**

L'aminoacylation spécifique des ARNt par leur aaRS homologue est une étape cruciale pour la synthèse de protéines fonctionnelles. L'identité d'un ARNt ou d'une famille d'ARNt isoaccepteurs repose sur des éléments positifs (déterminants), permettant la reconnaissance et l'aminoacylation par l'aaRS spécifique, et négatifs (antidéterminants), empêchant toute aminoacylation non spécifique. Un jeu d'identité est complet si sa transplantation dans un ou plusieurs ARNt différents permet l'acquisition de la nouvelle spécificité d'aminoacylation aux ARNt hôtes. De telles expériences ont aussi permis de mettre en évidence l'implication de structures fines particulières dans l'expression de l'identité dans l'ARNt hôte. Ici, nous explorons l'expression du jeu d'identité aspartate de levure, bien caractérisé, au sein d'ARNt d'architectures distinctes mais de même spécificité, les quatre isoaccepteurs arginine de levure. L'objectif est de savoir si l'expression de la nouvelle identité est similaire due à la spécificité unique des ARNt hôtes ou différente dû aux séquences et structures propres à chacun. Les transcrits *in vitro* des $ARNt^{Arg}$ de séquences sauvages ou transplantés des éléments d'identité aspartate ont été préparés et leurs propriétés d'aminoacylation établies. Les quatre $ARNt^{Arg}$ sauvages sont actifs en arginylation avec des efficacités catalytiques séparées d'un facteur 20 et sont inactifs en aspartylation. Alors que les transplants des $ARNt_1^{Arg}$ et $ARNt_4^{Arg}$ deviennent des substrats efficaces de l'AspRS, ceux des $ARNt_2^{Arg}$ et $ARNt_3^{Arg}$ restent inactifs. La recherche d'antidéterminants au sein de ces deux

ARNt a révélé des caractéristiques idiosyncrasiques. En effet, la conversion de la seule paire de base C6-G67 en G6-C67, présente dans l'ARNt$^{Asp}$, permet l'expression complète de l'identité aspartate dans le transplant de l'ARNt$_2^{Arg}$, mais non dans l'ARNt$_3^{Arg}$. En conclusion, les différents ARNt isoaccepteurs se protègent de l'aminoacylation hétérologue par des voies idiosyncrasiques d'antidétermination.

*[signalement bibliographique ajouté par : ULP – SCD – Service des thèses électroniques]*

**Aurélie Fender**, Marie Sissler, Catherine Florentz and Richard Giegé.

**Functional idiosyncrasies of tRNA isoacceptors in cognate and noncognate aminoacylation systems.**

Biochimie 86, 21-29 (2004)

Pages 21-29 :

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur /

http://dx.doi.org/10.1016/j.biochi.2003.11.011

Il est également possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: peb.sciences@scd-ulp.u-strasbg.fr.

## Complément : l'efficacité d'arginylation des ARNt$^{Arg}$ n'est pas équivalente

Bien que les 4 molécules sioent de très bons substrats pour l'ArgRS, leurs efficacités relatives d'arginylation sont variables. La question de la conservation *in vivo* de telles différences d'efficacité se pose. Si tel était le cas, quelle pourrait être sa signification biologique ? Le Tableau 4 montre un parallèle entre l'usage relatif des codons arginine chez *S. cerevisiae*, le nombre de copies de gènes des ARNt$^{Arg}$ et leur efficacité catalytique. Bien que l'ARNt$_1^{Arg}$ soit mineur (une seule copie du gène), l'usage dans la synthèse protéique du codon qu'il déchiffre (AGG) est élevé (21 % des codons arginine). Si la quantité d'ARNt exprimée dans la cellule est corrélée au nombre de gènes, l'aminoacylation de l'ARNt$_1^{Arg}$ serait plus efficace pour satisfaire cet usage important. Notons cependant que l'ARNt$_1^{Arg}$ n'est pas indispensable dans la cellule. En effet, le codon qui lui correspond peut être décodé par l'ARNt$_3^{Arg}$. Ce dernier possède une base modifiée en position 34 qui lui permet une souplesse dans la reconnaissance codon–anticodon. Les ARNt$_1^{Arg}$ et ARNt$_3^{Arg}$ pourraient être en compétition dans la cellule pour la lecture des codons AGG. Tout ceci reste non résolu à ce jour et nécessitera des analyses complémentaires.

| **Codon** (5'-3') | **Taux d'utilisation du codon** (%)[a] | **ARNt** | **Nombre de copies de gènes** | **Efficacité catalytique** ($10^{-3}$ $sec^{-1}$ $nM^{-1}$)[b] |
|---|---|---|---|---|
| CGG | 3,9 | ARNt$_4^{Arg}$ | 1 | 1,1 |
| AGG | 20,9 | ARNt$_1^{Arg}$ | 1 | 13,0 |
| AGA | 47,9 | ARNt$_3^{Arg}$ | 11 | 0,7 |
| CGU<br>CGC<br>CGA | 14,6<br>5,9<br>6,8 | ARNt$_2^{Arg}$ | 6 | 3,7 |

**Tableau 4. Correspondance entre l'utilisation des codons arginine, le nombre de copies de gènes des ARNt$^{Arg}$, et leur efficacité catalytique en arginylation, chez *S. cerevisiae*.** [a]Le taux d'utilisation correspond à l'usage relatif d'un codon arginine, parmi les 6, au sein de la synthèse protéique chez *S. cerevisiae* (d'après le site http://www.kazusa.or.jp/codon/). [b]Les efficacités catalytiques sont extraites du Tableau 1 de l'article n°1.

## Note

Les différences fonctionnelles des isoaccepteurs de l'arginine tant dans le contexte homologue que non homologue soulèvent la question de l'origine des ARNt au sein de cette famille isoacceptrice. Le faible pourcentage d'identité de séquences pourrait révéler des ascendances différentes. A savoir les 4 ARNt$^{Arg}$ découlent-ils d'un ancêtre commun ou

possèdent-ils des histoires évolutives différentes. Ce point sera développé dans la partie suivante en particulier pour l'ARN$t_4^{Arg}$.

## B. Relation fonctionnelle et évolutive entre les systèmes arginine et aspartate de levure

De façon remarquable, une relation fonctionnelle entre les systèmes arginine et aspartate existe. En effet, dès 1973, un complexe hétérologue formé de l'ArgRS de levure et de l'ARNt$^{Asp}$ a été isolé (Gangloff et al., 1973). Ce complexe est cependant peu productif puisque l'ARNt$^{Asp}$ est arginylé 12 600 fois moins efficacement que l'ARNt$^{Arg}$. Plus récemment, il a été montré que l'ARNt$^{Asp}$ obtenu par transcription *in vitro*, et donc dépourvu de ses modifications post-transcriptionnelles, est reconnu et aminoacylé de façon très efficace par l'ArgRS et peut de ce fait être considéré comme un pseudo-isoaccepteur arginine (Perret et al., 1990). C'est un simple groupement chimique, la méthylation de la base $G_{37}$ qui suffit à protéger l'ARNt$^{Asp}$ naturel de la reconnaissance par l'ArgRS (Pütz et al., 1994). L'arginylation du transcrit de l'ARNt$^{Asp}$ est liée à la présence d'un jeu d'identité arginine alternatif de celui mis à jour dans l'ARNt$_3^{Arg}$ (C35 et G/U36), tant par sa séquence que par sa position : il s'agit des résidus $C_{36}$ et $G_{37}$ (Sissler et al., 1996).

La comparaison des séquences des 4 ARNt$^{Arg}$ et de l'ARNt$^{Asp}$ montre que 3 des isoaccepteurs arginine se ressemblent (53 % de résidus identiques), alors que l'ARNt$_4^{Arg}$ se singularise par le fait qu'il a une séquence plus proche de celle de l'ARNt$^{Asp}$ (57 % d'identité) que de celles des 3 autres ARNt$^{Arg}$ (33 % d'identité). Cette ressemblance est étendue à des caractéristiques structurales fines. Elles concernent les tailles de la boucle variable (v=4), et des régions $\alpha$ et $\beta$ situées de part et d'autre des résidus G18 et G19 de la boucle D ($\alpha$=$\beta$=3), conservées uniquement entre ces deux ARNt.

L'existence d'un ARNt$^{Arg}$ aux ressemblances frappantes avec l'ARNt$^{Asp}$ permet d'élargir l'investigation de la relation fonctionnelle entre ces deux systèmes. En particulier, les propriétés d'aspartylation de l'ARNt$_4^{Arg}$ sauvage ou muté, transcrit *in vitro* ou exprimé dans la levure, ont été explorées et notamment la question d'une relation évolutive entre les gènes *ARNt$_4^{Arg}$* et de *ARNt$^{Asp}$* a été abordée. L'ensemble des résultats a fait l'objet de l'article n°2.

Article n°2

# Réminiscence d'une ancienne spécificité dans un ARNt contemporain

**Fender, R. Geslain, G. Eriani, R. Giegé, M. Sissler & C. Florentz**

**2004**

**Nucleic Acids Research, 32 (17) :5076-5086**

**Résumé :**

La haute spécificité d'aminoacylation des ARNt par leur aaRS garantit une traduction fidèle et précise de l'information génétique. Les particularités structurales et fonctionnelles, attenantes à chaque couple aaRS/ARNt, suggèrent une indépendance entre les systèmes d'aminoacylation. Cependant, l'existence de misaminoacylation d'ARNt par des aaRS non spécifiques reflète des relations entre de tels systèmes. Chez *S. cerevisiae*, un lien fonctionnel entre les systèmes d'arginylation et d'aspartylation a été montré. En particulier, il a été établi que l'ARNt$^{Asp}$ transcrit *in vitro* est un substrat efficace de l'ArgRS. Dans cette étude, la relation entre les systèmes arginine et aspartate est davantage explorée, tenant compte de la découverte d'un quatrième ARNt isoaccepteur de l'arginine (ARNt$_4^{Arg}$) dans le génome de la levure. Cet ARNt possède une séquence particulièrement proche de celle de l'ARNt$^{Asp}$ mais divergente des trois autres ARNt$^{Arg}$. Suite à la transplantation du jeu complet d'identité aspartate au sein des quatre ARNt$^{Arg}$, l'ARNt$_4^{Arg}$ gagne la plus grande efficacité d'aspartylation. De plus, il est possible de convertir l'ARNt$_4^{Arg}$ en un accepteur de l'aspartate, aussi efficace que l'ARNt$^{Asp}$, par seulement deux mutations ponctuelles, C38 et G73, et ce malgré l'absence des éléments d'identité aspartate majeurs de l'anticodon. Ainsi, des éléments aspartate cryptiques sont enfouis dans l'ARNt$_4^{Arg}$. La capacité dissimulée d'accepteur aspartate au sein d'un ARNt$^{Arg}$ contemporain conduit à la proposition d'un lien évolutif entre les gènes des ARNt$_4^{Arg}$ et ARNt$^{Asp}$.

*[signalement bibliographique ajouté par : ULP – SCD – Service des thèses électroniques]*

**Aurélie Fender**, Renaud Geslain, Gilbert Eriani, Richard Giegé, Marie Sissler and Catherine Florentz.

**A yeast arginine specific tRNA is a remnant aspartate acceptor.**

Nucleic acids research 32, 5076-5086 (2004)

Pages 5076-5086 :

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur /

http://nar.oxfordjournals.org/cgi/reprint/32/17/5076.pdf

Il est également possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: peb.sciences@scd-ulp.u-strasbg.fr.

## Complément : schéma évolutif

Nous proposons que les $ARNt^{Asp}$ et $ARNt_4^{Arg}$ contemporains descendent d'un ancêtre commun, possédant une identité ambiguë. Cette molécule primordiale auraient été reconnue par les AspRS et ArgRS primitives grâce à la présence simultanée des éléments d'identité aspartate G34U35C36 (anticodon) et G73, et arginine C36 et G37 (jeu alternatif, (Sissler et al., 1996)) (Figure 19). De façon remarquable, la composition en nucléotides de la boucle anticodon de cet ARNt primordial est telle qu'elle aurait permis la lecture à la fois des codons aspartate GA(U/C) et arginine CG(G/A) (dont le codon CGG est spécifique de l'$ARNt_4^{Arg}$) par un mécanisme original. En effet, l'utilisation du triplet anticodon classique en position 34-35-36 permet la lecture des codons asparate, alors que l'utilisation d'un triplet décalé d'un seul nucléotide et en position 35-36-37 permet celle des codons arginine (voir figure X). Le gène de cet ARNt à identité ambiguë aurait été dupliqué en réponse directe à l'émergence du codon CGG, permettant ainsi la génération de l'$ARNt_4^{Arg}$.

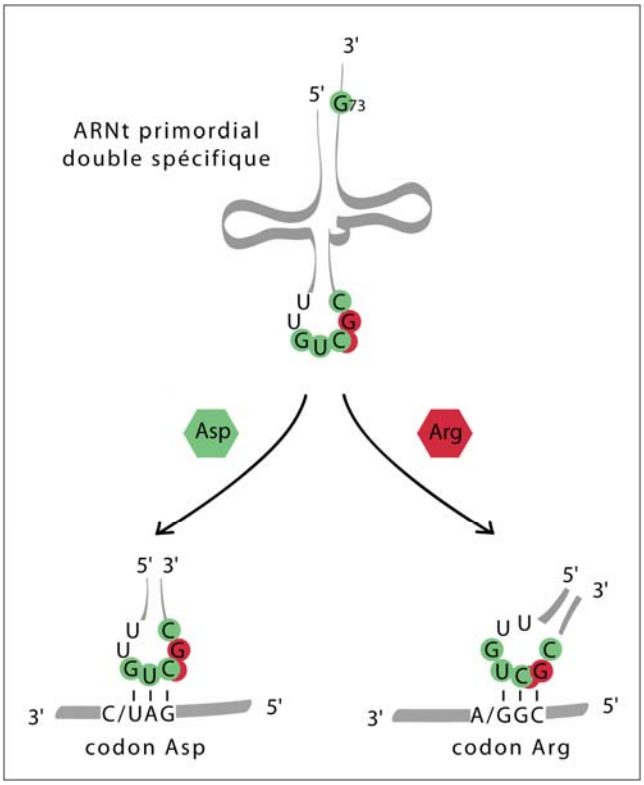



**Figure 19. Possibilité d'un décodage double par un ARNt primordial.** Le jeu d'identité aspartate est représenté en vert, celui alternatif de l'arginine est en rouge. Le décodage des codons arginine serait possible grâce à l'utilisation d'un triplet anticodon alternatif (décalé d'un nucléotide).

CHAPITRE 2

# Etude du système d'aspartylation mitochondrial humain

# A. Vers le jeu complet d'aminoacyl-ARNt synthétases mitochondriales humaines : caractérisation de l'AspRS et de la TyrRS

Les systèmes d'aminoacylation mt de mammifères sont originaux car leurs ARNt, codés par le génome mt, possèdent des structures dégénérées alors que les aaRS mt, codées par le génome nucléaire, ressemblent aux synthétases classiques (Sissler et al., 2005). Ceci laisse présager de nouvelles règles de reconnaissance entre ces deux macromolécules et en particulier de nouveaux signaux d'identité. Nous nous intéressons plus particulièrement aux systèmes mt humains, pour lesquels des mutations des ARNt peuvent être corrélées à des pathologies (Florentz et al., 2003 ; Levinger et al., 2004).

### Vers le jeu complet de gènes d'aaRS mt humaines

Au début de ce travail, les gènes des 22 gènes d'ARNt nécessaires et suffisants à la traduction mt étaient séquencés (Anderson et al., 1981). En revanche, alors que les gènes des 20 aaRS cytoplasmiques et les protéines correspondantes étaient connues, seulement 7 gènes d'aaRS mt avaient été identifiés, clonés et les protéines correspondantes caractérisées (IleRS, PheRS, LyRS, LeuRS, TrpRS, SerRS et MetRS) et deux autres gènes avaient été caractérisés mais les protéines non étudiées (ceux codant pour la GlyRS mt et l'HisRS mt) (Sissler et al., 2005). Au vue des avancées dans le séquençage du génome humain, il devenait possible et essentiel d'avoir une vue d'ensemble des gènes d'aaRS mt humaines. En collaboration avec Marie Sissler et Luc Bonnefond (doctorant sous la direction conjointe de Joëlle Rudinger-Thirion et de Richard Giegé), nous avons effectué une recherche systématique de tous les gènes manquants des synthétases mt humaines. Dix gènes supplémentaires (codant pour les ArgRS, CysRS, GluRS, TyrRS, ValRS, AlaRS, AsnRS, AspRS, ProRS et ThrRS) sur les 11 manquants (l'exception concerne la GlnRS) ont été trouvés et caractérisés. Ceci fait l'objet de l'article n°3.

L'analyse des gènes et de leurs séquences traduites en protéine conduit à deux conclusions majeures : (i) les aaRS mt sont codées par un gène distinct de ceux des enzymes cytoplasmiques (excepté les LysRS et GlyRS pour lesquels un seul gène code pour les enzymes mt et cytoplasmiques) (Figure 20), et (ii) les protéines codées suivent toutes la répartition prévue en deux classes de synthétases selon la nature structurale de leur site

catalytique. Notons toutefois le motif « -MKTR » de l'ArgRS mt au lieu de la séquence signature classique « KMSKS », déviation déjà observée dans d'autres ArgRS (Delagoutte et al., 2000), et l'absence des motifs 1 et 2 dans l'AlaRS mt, également observé dans d'autres AlaRS (Chihade et al., 2000).

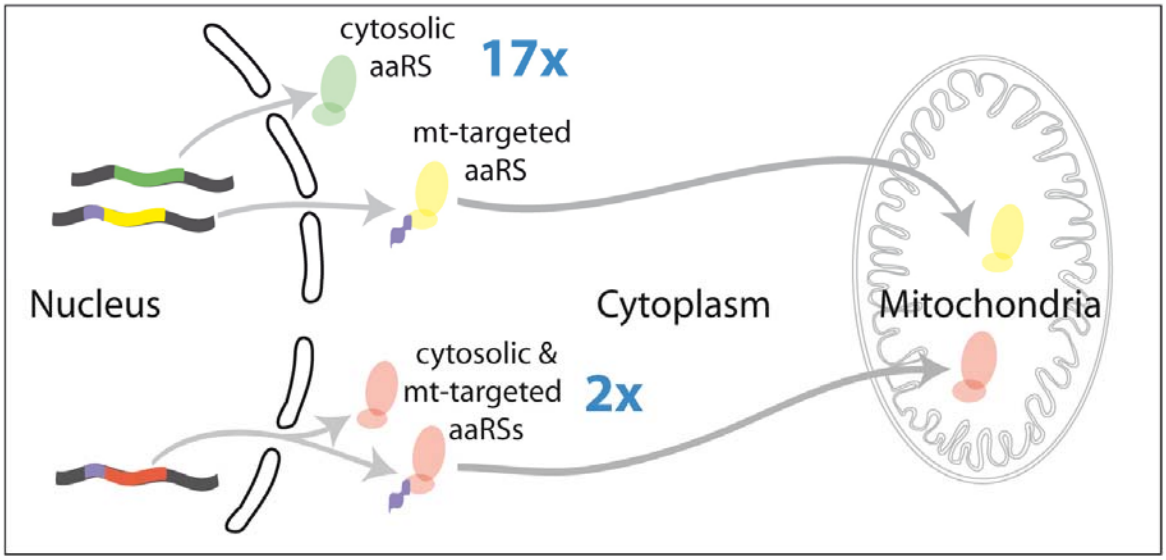



**Figure 20. Diversité des gènes nucléaires pour les aminoacyl-ARNt synthétases cytoplasmiques et mitochondriales.** Deux stratégies existent : soit les deux enzymes sont codées par des gènes différents (17 cas), soit par un gène unique (2 cas). Un signal d'adressage vers la mitochondrie (extension violette) permet l'import vers l'organelle.

Les enzymes mt partagent de 23 à 100% d'identité de séquences avec leurs homologues cytoplasmiques, et présentent quelques différences quant à leur oligomérie. Ainsi, la PheRS est monomérique ($\alpha$) alors que l'enzyme cytoplasmique est tétramérique ($\alpha 2\beta 2$) (situation déjà observée chez la levure, (Sanni et al., 1991)). La GluRS et la ProRS sont des enzymes indépendantes alors que les enzymes cytoplasmiques sont fusionnées (Kaiser et al., 1994).

En ce qui concerne la GlnRS mt, son gène n'a pas était trouvé soit parce que sa séquence varie tellement des autres GlnRS soit qu'il n'existe pas. Plusieurs travaux avaient précédemment suggéré que la synthèse de Gln-ARNt$^{Gln}$ mt se ferait par la voie indirecte de glutamylation de l'ARNt$^{Gln}$ suivie de la transamidation du résidu Glu en Gln sur l'ARNt, déjà décrite chez les bactéries, archaea et chloroplastes (Schön et al., 1988 ; Schön & Söll, 1988; Ibba & Söll, 2000). Récemment, il a été montré que pour pallier au défaut de GlnRS mt chez S. *cervisiae*, le système complet de glutaminylation GlnRS/ARNt$^{Gln}$ cytoplasmique est importé (Rinehart et al., 2005).

## Exploration spécifique des systèmes aspartate et tyrosine

Deux systèmes d'aminoacylation ont été étudiés plus en détail. Il s'agit du système tyrosine, étudié par Luc Bonnefond, et du système aspartate, objet de cette thèse. Les deux enzymes sont de type bactérien (elles présentent chacune des domaines additionnels retrouvés seulement chez les bactéries) et chargent efficacement les ARNt d'*E. coli*. Toutes deux sont capables d'aminoacyler l'ARNt mt correspondant à leur spécificité, produit par transcription *in vitro*, cependant avec une efficacité 10 à 40 fois moins grande que pour les ARNt d'*E. coli*.

Le système d'aspartylation mt humain a été choisi pour plusieurs raisons : (i) l'ARNt$^{Asp}$ mt possède une séquence/structure bizarre par rapport aux ARNt$^{Asp}$ canoniques, (ii) il semble posséder un jeu d'identité différent et (iii) les systèmes aspartate de levure et bactériens (*E. coli* et *T. thermophilus*) étant bien connus, il sera possible de réaliser une analyse comparative des systèmes d'aspartylation et ainsi d'aborder la question de leur évolution.

## Difficultés techniques et précautions

Nous avons choisi d'étudier les systèmes d'aminoacylation mt humains par une approche *in vitro*, mettant en jeu des ARNt produits par transcription *in vitro* et des enzymes clonées, surproduites et purifiées. Cette approche permet de contrecarrer l'obtention de quantités suffisantes de matériel humain qui reste délicate. De plus la stratégie *in vitro* présente l'avantage de pouvoir tester de nombreux mutants aussi bien de l'ARNt que de l'enzyme.

En ce qui concerne les tests fonctionnels, le milieu réactionnel utilisé est celui mis au point précédemment au laboratoire spécifiquement pour les systèmes mt humains (Sohm et al., 2003). Le rapport [$Mg^{2+}$]/[ATP] améliore de façon significative le taux de charge. Ce milieu a été testé et sa validité pour le système d'aspartylation vérifiée. Cependant, nous avons rencontré des problèmes de reproductibilité des résultats et de variation d'activité des transcrits en fonction de la température. Ainsi par exemple, l'ARNt$^{Asp}$ mt transcrit subit une perte d'activité importante au bout de 5 min d'incubation à 37°C. Ceci s'explique par la présence d'un grand nombre de sites CpA et UpA (14 en tout), sites privilégiés d'hydrolyse spontanée des ARN (Dock-Bregeon et al., 1989), mais aussi par une grande flexibilité du transcrit, l'ARNt$^{Asp}$ mt possèdant une très faible proportion de paires de bases G-C (seulement 5 contre 14 retrouvées dans l'ARNt$^{Asp}$ d'*E. coli*). De plus, le transcrit est dépourvu de bases modifiées, qui peuvent avoir un rôle stabilisant pour la structure des ARNt (Dalluge et al., 1996 ; Helm et al., 1998). Nous avons tenté deux stratégies pour pallier à ce problème : (i)

l'introduction de paires de bases G-C aux deux extrémités du bras accepteur de l'ARNt$^{Asp}$, de façon à verrouiller l'hélice acceptrice, et (ii) la baisse de la température réactionnelle, stratégie déjà validée pour un autre système mt (Sissler et al., 2004). Les tests d'aminoacylation montrent que l'introduction de paires de bases G-C n'améliorent pas l'activité de l'ARNt$^{Asp}$, mais que la diminution de la température réactionnelle de 37°C à 15°C constitue un compromis entre l'activité efficace de l'enzyme et la stabilité de l'ARNt. Les transcrits restent actifs en plateaux de charge (au moins jusqu'à 45 min) et les résultats en vitesse sont reproductibles. Ces conditions ont été adoptées pour la suite du travail.

Article n°3

# Vers le jeu complet d'aminoacyl-ARNt synthétases mitochondriales humaines : Caractérisation de l'AspRS et de la TyrRS

**A. Fender, L. Bonnefond, J. Rudinger-Thirion, R. Giegé, C. Florentz & M. Sissler**

**2005**

**Biochemistry, 44 :4805-4816**

**Résumé :**

La mitochondrie humaine possède une machinerie traductionnelle entièrement dévolue à la synthèse de 13 protéines. Alors que les ARNt et les ARNr nécessaires sont produits par transcription du génome mitochondrial, tous les autres facteurs requis pour la synthèse protéique sont synthétisés dans le cytoplasme et importés. Ceci est notamment le cas des aaRS. Alors que les gènes codant pour les enzymes cytoplasmiques sont bien définis, seulement neuf gènes d'aaRS mitochondriales sont connus. Ceux manquants ont été recherchés et identifiés, à l'exception de celui de la GlnRS. Dans cet article, nous décrivons plus particulièrement les gènes codant pour l'AspRS et la TyrRS mitochondriales humaines et la caractérisation initiale des deux enzymes. Les deux synthétases appartiennent à la classe attendue, possèdent une organisation dimérique et aminoacylent aussi bien les ARNt d'*E. coli* que les ARNt mitochondriaux humains transcrits *in vitro*. L'analyse des séquences de l'ensemble des gènes d'aaRS humaines confirme la vue selon laquelle, les synthétases mitochondriales et cytoplasmiques, exceptées la LysRS et la GlyRS, sont codées par deux jeux différents de gènes.

*[signalement bibliographique ajouté par : ULP – SCD – Service des thèses électroniques]*

Luc Bonnefond, **Aurélie Fender**, Joëlle Rudinger-Thirion, Richard Giegé, Catherine Florentz, and Marie Sissler.

**Towards the full set of human mitochondrial aminoacyl-tRNA synthetases: charactérization of AspRS and TyrRS.**

Biochemistry 44, 4805-4816 (2005)

Pages 4805-4816 :

La publication présentée ici dans la thèse est soumise à des droits détenus par un éditeur commercial.

Pour les utilisateurs ULP, il est possible de consulter cette publication sur le site de l'éditeur /

http://pubs.acs.org/cgi-bin/article.cgi/bichaw/2005/44/i12/pdf/bi047527z.pdf

Il est également possible de consulter la thèse sous sa forme papier ou d'en faire une demande via le service de prêt entre bibliothèques (PEB), auprès du Service Commun de Documentation de l'ULP: peb.sciences@scd-ulp.u-strasbg.fr.

## Nouvelles perspectives

La connaissance de l'ensemble des gènes codant pour les aaRS mt humaines ouvre de nombreuses perspectives.

(i) De plus amples analyses des séquences permettront d'apporter des renseignements au niveau de l'évolution des gènes des aaRS et de la mitochondrie en général.

(ii) Ceci permettre la caractérisation individuelle de tous les systèmes d'aminoacylation tant au niveau fonctionnel que structural. Dans ce sens, un projet de génomique structurale a été initié au laboratoire et vise à cloner, purifier et cristalliser les aaRS mt humaines nouvellement découvertes. Ceci apportera des connaissances considérables sur les systèmes d'aminoacylation mt de mammifères, à l'heure où les données structurales restent encore très pauvres puisque seule la structure de la SerRS mt de bœuf est connue et ce depuis peu (Chimnaronk et al., 2004 ; Chimnaronk et al., 2005).

(iii) L'accumulation et la comparaison des données concernant l'ensemble des systèmes d'aminoacylation permettra l'établissement de règles quant aux signaux d'aminoacylation spécifique et aux modes de reconnaissance des ARNt mt « bizarres » par les aaRS mt *a priori* « classiques ».

(iv) D'autres questions pourront être abordées. Ainsi on peut se demander si les aaRS mt forment un complexe multisynthétasique comme les aaRS cytoplasmiques. Requièrent-elles des cofacteurs ou protéines chaperonnes pour une fonctionnalité optimale ? Quelles est la localisation des synthétases au sein de la cellule et de la mitochondrie ? Sont-elles proches des membranes mt comme ceci semble être le cas pour les ribosomes mt ? Pour répondre à ces questions, il sera intéressant de produire des anticorps spécifiques afin de réaliser des expériences de co-immunoprécipitation et de localisation *in situ*.

(v) Par ailleurs, la complexité du rapport génome/protéome humain a été établie par le recensement de seulement 32 000 gènes pour un besoin en protéines fonctionnelles bien plus important. Ceci renforce le rôle majeur de mécanismes d'épissage alternatif des pré-ARNm. La possibilité de fonctions alternatives des aaRS mt, produites par épissage alternatif, pourra être abordée.

Dans ce qui va suivre, divers aspects concernant le système mt aspartate ont été explorés. Ainsi en Partie B, nous avons étudié les éléments d'identité de l'ARNt$^{Asp}$ mt et l'adaptation de l'AspRS à ce substrat particulier. La Partie C décrit la recherche des éléments de l'ARNt$^{Asp}$ mt responsable de sa non-reconnaissance par des AspRS d'autres organismes. Enfin, la Partie D est dédiée à une étude structurale et présente des essais de cristallisation.

Toutefois, tentant d'exploiter le plus largement possible les perspectives proposées, d'autres projets ont été initiés. Ceux-ci sont brièvement décrits ici.

### Production d'anticorps dirigés contre l'AspRS mt humaine

Nous avons produit, chez le lapin, des anticorps dirigés contre l'enzyme mt humaine clonée avec une queue poly-histidines (polyHis) et purifiée. La spécificité des anticorps contenus dans les sérums prélevés a été testée par ELISA (« Enzyme-Linked Immunosorbent Assay ») (voir Matériel et Méthodes) (Figure 21). La présence d'anticorps anti-polyHis a également été testée par ELISA en utilisant la TyrRS-polyHis (clonée et caractérisée par Luc Bonnefond, article n°3) comme antigène. Les résultats montrent que l'activité anti-AspRS mt est importante, mais qu'il existe une population minoritaire d'anticorps dirigés contre la queue polyHis (Figure 21). Nous pensons qu'il sera important de purifier les anticorps anti-AspRS mt avant de les utiliser. Ceci pourra se faire par chromatographie d'affinité, en utilisant une matrice sur laquelle sera fixée covalemment notre protéine. La déplétion des anticorps pourra se faire de la même façon, mais en utilisant une matrice substituée en queue polyHis ou en TyrRS-polyHis.

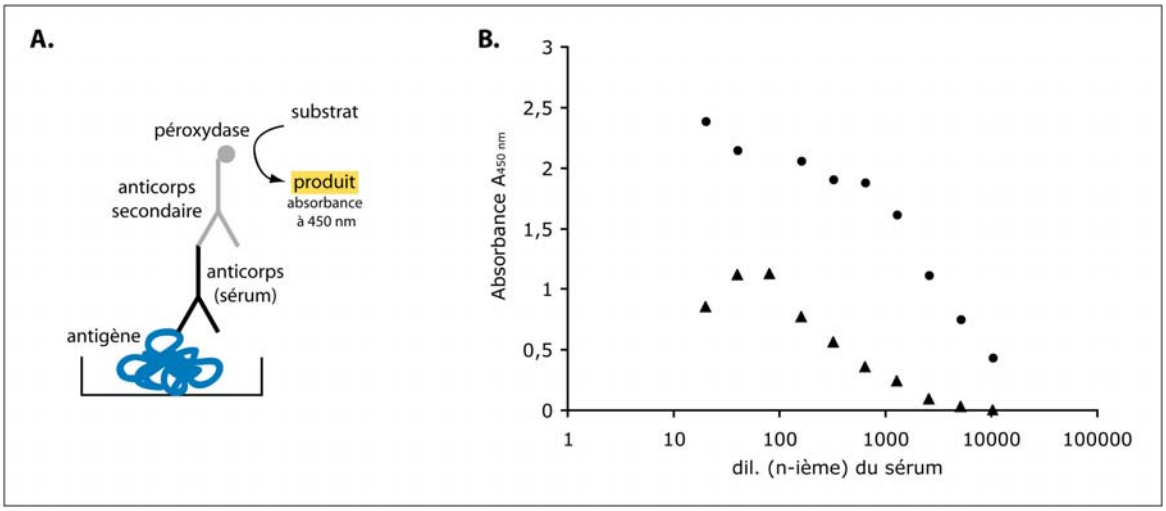



**Figure 21. Activité et spécificité des anticorps anti-AspRS.** L'activité et la spécificité des anticorps ont été vérifiées par test ELISA (« Enzyme-Linked Immunosorbent Assay »). **A.** Schématisation du principe du test ELISA réalisé. L'antigène (AspRS mt humaine : test d'activité, ou TyrRS mt humaine (polyHis) : test de spécificité) est adsorbé sur plaque, puis incubé en présence de différentes dilutions de sérum (contenant les anticorps anti-AspRS). La fixation de l'anticorps est révélée grâce à l'incubation d'un anticorps secondaire couplé à la peroxydase et de son substrat. Le produit jaune libéré est décelable par lecture de l'absorbance à 450 nm. **B.** Courbe d'absorbance à 450 nm (en ordonnée) en fonction de la dilution de sérum appliquée (en abscisse, échelle logarithmique). Les ronds noirs correspondent à l'activité anti-AspRS du sérum, les triangles noirs révèlent la proportion d'activité anti-queue polyHis non spécifiques de l'AspRS.

Une fois purifiés, ces anticorps permettront de répondre à différentes questions. Leur utilisation lors d'expériences d'hybridation *in situ* ou de western blot, réalisées sur différents extraits cellulaires, permettra de localiser l'AspRS mt. Les expériences de western blot permettront également de savoir si l'AspRS existe sous des formes de tailles différentes (voir ci-dessous). Finalement, les anticorps permettront d'entreprendre des expériences de co-immunoprécipitation visant à rechercher des partenaires de l'AspRS (autres aaRS, nouveaux facteurs, ARN autres que l'ARNt$^{Asp}$, etc…).

### AspRS mt et produit d'épissage alternatif ?

Un autre projet intéressant, également initié en parallèle, vise à découvrir des protéines dérivées de l'AspRS mt qui pourraient avoir une fonction nouvelle (voir introduction « Fonctions alternatives des aminoacyl-ARNt synthétases »).

De façon originale, deux produits de tailles différentes ont été obtenus après PCR (ciblée sur le gène de l'AspRS mt) à partir d'une banque d'ADNc humains : l'un d'environ 1800 nucléotides correspondant à la séquence complète de l'AspRS dépourvue du signal d'adressage vers la mitochondrie, et l'autre d'environ 1650 nucléotides (Figure 22.A.). Le clonage et séquençage du fragment plus court, nommé mini-AspRS, a permis d'identifier un exon entier (l'exon 13) comme étant la séquence manquante.

Le gène de l'AspRS mt est partagé en 17 exons distribués sur 32,5 kb (Figure 22.B.). De façon originale, chaque exon correspond à un domaine structural de l'enzyme d'*E. coli*. Un modèle structural de l'AspRS mt humaine a été construit à partir de la structure cristallographique de l'AspRS d'*E. coli* par Claude Sauter (voir article n°4). D'après ce modèle, aucune des jonctions intron/exon ne sectionne une structure en hélice ou en brin β. Une telle conservation exon/domaine structural pourrait permettre des phénomènes d'épissage alternatif de l'ARNm et ainsi par « exon shuffling », la genèse de nouvelles protéines.

La correspondance stricte entre la région manquante et les bornes d'un exon laisse supposer que mini-AspRS résulte de l'épissage alternatif de l'ARNm. La traduction *in silico* de la mini-AspRS donne lieu à une protéine dont une partie du domaine d'insertion typiquement bactérien dans le site catalytique est absente (Figure 22.C.). Dans l'AspRS mt humaine complète, le domaine d'insertion comprend 129 acides aminés repliés en feuillet β à 4 brins anti-parallèles entouré d'hélices α. La fonction de ce domaine n'est pas très bien connue, mais elle pourrait avoir un rôle dans la stabilisation de l'interaction entre l'AspRS

bactérienne et l'ARNt correspondant (Delarue et al., 1994). Dans la mini-AspRS, ce domaine est écourté à 78 acides aminés et serait réduit à un brin β entouré de part et d'autre d'un motif hélice-boucle-hélice. Notons que les AspRS de type eucaryotique ou d'archaea ne possèdent pas d'insertion dans le domaine catalytique mais une courte séquence d'environ 35 acides aminés qui se replie en un court brin β suivi par un motif hélice-boucle-hélice (revue par (Giegé & Rees, 2005)). Cette région est similaire à celle décrite ci-dessus dans la mini-AspRS.

Il s'agira maintenant de vérifier la réalité biologique de la forme courte mini-AspRS au niveau de l'ARN mais aussi au niveau de la protéine, par des expériences de northern blot ou de RT-PCR et par wertern blot. Des expériences préliminaires de northern blot ont été initiées à partir d'ARN total de cellules musculaires humaines. Cependant elles devront être optimisées davantage ou réalisées à partir d'ARNm purifiés. Il serait par ailleurs intéressant de définir si l'expression des différentes formes de l'AspSR est identique en fonction du type cellulaire. Dans le cas où une forme alternative de l'AspRS existerait, il sera important de vérifier sa localisation subcellulaire. Des expériences de western blot réalisées à la fois sur des extraits cytoplasmiques et des extraits mitochondriaux permettront de répondre à cette question. Enfin, la mini-AspRS pourra être exprimée, purifiée et caractérisée biochimiquement afin d'identifier sa(ses) fonction(s) biologique(s). Sa capacité à aminoacyler l'ARNt sera explorée en priorité.

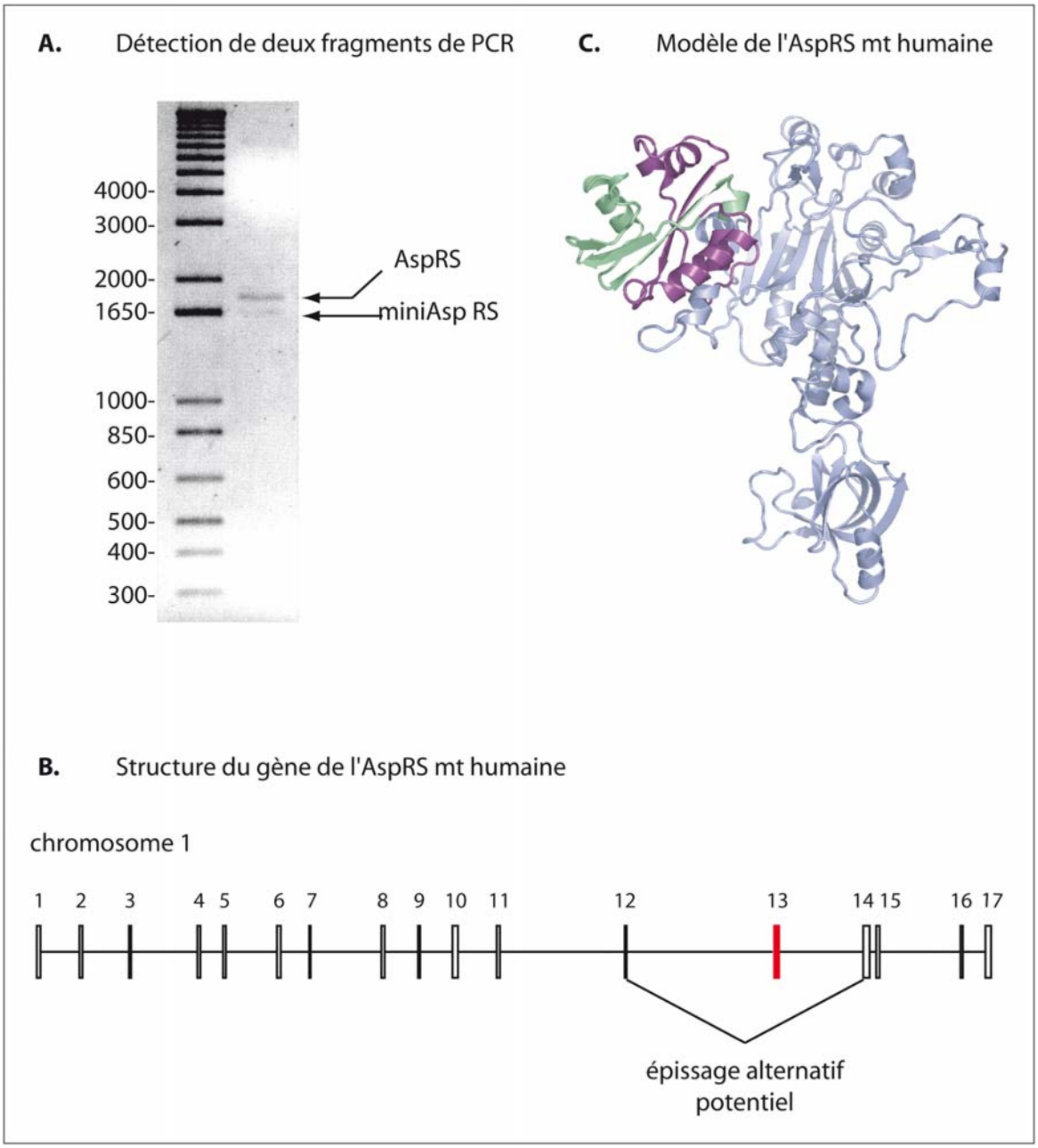



**Figure 22. Epissage alternatif du gène de l'AspRS mitochondriale humaine. A.** Deux fragments de PCR, obtenus à partir d'une banque d'ADNc humain en présence d'oligonucléotides complémentaires aux extrémités de l'ARNm codant pour l'AspRS (dépourvue du signal d'adressage vers la mitochondrie), ont été détectés par électrophorèse sur gel d'agarose 1 %. **B.** Organisation du gène de l'AspRS mt représentant la composition en exons et introns. Les proportions relatives sont conservées. L'exon potentiellement épissé est coloré en rouge. **C.** Mise en évidence de la séquence manquante (en vert) sur la structure d'un monomère de l'AspRS mt modélisée (à partir de la structure de l'AspRS d'*E. coli*, voir article n°4) Le domaine d'insertion spécifique des AspRS bactériennes est en violet. La figure a été générée en utilisant le programme PyMol (DeLano, 2002).

## B. Eléments d'identité dans un ARNt mitochondrial

Bien que les systèmes d'aminoacylation canoniques soient largement étudiés, l'identité et la spécificité d'aminoacylation des systèmes mitochondriaux sont encore peu explorées. Cependant, on peut anticiper que ces systèmes se distinguent des autres en raison (i) de leurs origine génétique (ARNt codés par le génome mitochondrial et les aaRS codées par le noyau et importées dans la mitochondrie), et (ii) du taux de mutations différents dans ces deux gènes. Ainsi, la confrontation des deux partenaires ARNt/aaRS mt pose la question de l'adaptation et de la reconnaissance spécifique au sein des systèmes d'aminoacylation mt de mammifères. Pour répondre à cette question, nous avons choisi d'étudier l'identité du système d'aspartylation mt humain dont la mise en place a été décrite dans les pages précédentes. Cette étude a conduit à l'article n°4 qui montre d'une part que le jeu d'identité aspartate est réduit par rapport aux systèmes aspartate bactériens ou eucaryotiques avec la non importance du résidu 73.

Article n°4

# Evolution reverse d'un système d'aminoacylation mitochondrial

**A. Fender, C. sauter, M. Messmer, J. Pütz, R. Giegé, C. Florentz & M. Sissler**

**2005**

**Soumis à PNAS**

**Résumé :**

La machinerie de traduction mt de mammifères est d'origine double, avec les ARNt codés par le génome mt évoluant rapidement, et les aaRS codées par le noyau et importées. Les ARNt mt sont atypiques : ils possèdent des séquences biaisées, des boucles et bras de taille variable, et ne présentent pas certains résidus requis pour le repliement tertiaire. Les aaRS, elles, semblent classiques. Ceci soulève la question de savoir quels sont les éléments d'identité de tels ARNt et comment les aaRS s'adaptent. Nous avons exploré le système aspartate mt humain, pour lequel l'AspRS de type procaryotique, très similaire à l'enzyme d'*E. coli*, reconnaît un ARNt$^{Asp}$ « bizarre ». L'analyse de transcrits de l'ARNt$^{Asp}$ mt humain montre que les éléments de l'anticodon GUC participe à l'identité comme c'est le cas pour les autres systèmes classiques. En revanche, la position 73 n'est plus impliquée dans l'identité alors que le nucléotide G73 est un élément d'identité aspartate majeur universel. En effet, cette position peut être occupée par n'importe lequel des 4 nucléotides dans l'ARNt$^{Asp}$ mt sans affecter l'aspartylation. La modélisation de l'AspRS mt nous a permis de mettre en évidence le résidu Gly269 à la place de l'Asp220 qui est responsable de l'interaction avec l'élément G73 dans le complexe AspRS/ARNt$^{Asp}$ d'*E. coli*. La substitution de cette glycine par un aspartate suffit à convertir l'AspRS mt en une enzyme discriminant à nouveau l'élément G73. Ainsi, la perte de l'élément primordial d'identité G73 a été compensée par un changement local et subtil dans la séquence de l'AspRS, permettant le relâchement de stringence. La restriction du jeu d'identité, conséquence directe de l'évolution rapide du génome mt, suggère une évolution reverse du système d'aminoacylation mt (d'origine double) par rapport à son ancêtre bactérien.

# Reverse evolution for a mitochondrial tRNA aminoacylation system

**Aurélie Fender, Claude Sauter, Marie Messmer, Joern Pütz, Richard Giegé Catherine Florentz* and Marie Sissler**

Institut de Biologie Moléculaire et Cellulaire du CNRS, UPR 9002, Université Louis Pasteur, Department "Machineries Traductionnelles" 15 rue René Descartes, F-67084 Strasbourg Cedex, France.

**In mammalian mitochondria, the translational machinery is of dual origin with tRNAs encoded by simplified and rapidly evolving mitochondrial (mt) genome and aminoacyl-tRNA synthetases (aaRS) coded by the nuclear genome, and imported. Mt-tRNAs are atypical with biased sequences, size variations in loops and stems and absence of residues forming classical tertiary interactions, while synthetases appear typical. This raises questions about identity elements in mt-tRNAs and adaptation of their cognate mt-aaRSs. We explore here the human mt-aspartate system, where a prokaryotic-type AspRS, highly similar to the *E. coli* enzyme, recognizes a bizarre tRNA$^{Asp}$. Analysis of human mt-tRNA$^{Asp}$ transcripts confirms the identity role of the GUC anticodon as in other aspartylation systems, but reveals the non-involvement of position 73, an otherwise universally conserved major aspartate identity element as G73. This position can be occupied by any of the 4 nts without effects on aspartylation. Modeling human mt-AspRS allowed placing Gly269 at a position occupied by Asp220, the residue contacting G73 in the *Escherichia coli* AspRS/tRNA$^{Asp}$ complex. Replacing this glycine by an aspartate converts human mt-AspRS into a discriminating enzyme for G73. Thus, the loss of a primordial identity element G73 has been faced up by a local and subtle sequence change in the aaRS, leading to relaxation of stringency. Restriction in the identity set, driven by a rapid evolution of the mt-genome, suggests a reverse evolution of the mt-aminoacylation system of dual origin, in regard of its bacterial ancestor**

Specific aminoacylation of transfer RNAs (tRNA) by their cognate aminoacyl-tRNA synthetases (aaRS)[1] is a key step in translation of genetic information. Rules governing recognition between the partner macromolecules have been well studied for a large number of prokaryotic and cytosolic eukaryotic systems (1). They rely on identity nucleotides, mainly located at both extremities of the tRNA L-shape that contact specific amino acids in the aaRS structure. Composed of major and minor elements, identity sets are usually conserved during evolution, as is the architecture of tRNA (1). However, identities can be completed by idiosyncratic subtleties such as cryptic (2) or permissive (3) elements, and structural features including flexibility and the overall tRNA scaffold (4, 5). Despite this knowledge, identity and specificity for aminoacylation reactions remain mainly unknown for organellar tRNAs, in particular for mammalian mitochondrial (mt) tRNAs, a group of atypical macromolecules.

Mammalian mt-tRNAs differ from classical tRNAs both at the structural level and in regard of potential identity sets (6). Coded by the mitochondrial genome, they

[1] aaRS stands for aminoacyl-tRNA synthetase with aa for amino acid (for individual aaRS, aa is given in the 3 letter code, *e.g.* AspRS for aspartyl-tRNA synthetase).

* corresponding author

display numerous non-canonical features and significantly degenerated sequences (6, 7). A large set (14 out of 22) possesses a biased nucleotide composition leading to a high content of weak A–U pairs in their cloverleaf-like structure, variable size of D- and/or T-loops, and misses otherwise highly conserved residues required for 3D folding. Aminoacylation identity elements are only sparsely known. According to the endosymbiotic hypothesis on the origin of mitochondria (8) and the ability of mt-aaRSs to aminoacylate *Escherichia coli* tRNAs (9), identity elements of *E. coli* tRNAs were searched for in the corresponding mammalian mt-tRNA sequences, and a subset has indeed been found (6, 7). Conversely to the large sequence and structural degeneracy of mammalian mt-tRNAs, corresponding mt-aaRSs seem not to depart from the already known synthetases. They are nuclear encoded and possess a mt-targeting signal. Recent completion of the full set of human mt-aaRS genes (10) confirms the view that, except for GlyRS and LysRS, mt and cytosolic synthetases are coded by two different sets of genes. The synthetases of mt location have no apparent particular structural feature (except PheRS which is a monomer instead of a 2 2 tetramer). They belong to the expected aaRS classes since they present either of the two signatures in the catalytic domain (Rossmann-fold for class I; semi-conserved motifs 1, 2 and 3 for class II) (11, 12).

Apposition of "bizarre" tRNAs *versus* standard aaRSs implies modified rules for recognition and specific aminoacylation. To shed light in this question, we have explored the human mt aspartylation system. For the sake of comparison, known residues specifying aspartylation in the *E. coli* (13, 14), *Saccharomyces cerevisiae* (15, 16) and *Thermus thermophilus* (17) model systems were useful. Based on sequence comparison, "classical" (*i.e.* non-mt) $tRNA^{Asp}$ molecules constitute a homogeneous family (18, 19), in particular in regard of the universally conserved G73 discriminator residue (found in all non-mt $tRNA^{Asp}$, 20) and the GUC anticodon, which are major aspartate determinants. Other elements are of minor strength, such as C38 and G10–U/C25, or the *E. coli* specific G2–C71 pair (Fig. 1A). Human mt-$tRNA^{Asp}$ presents characteristic mammalian mt-tRNA features (Fig. 1B). It is shorter (71 nts instead of 76) due to short D- and T-loops, and has a biased nucleotide composition with 29A, 7G, 24U, 11C leading to a high content of weak base-pairs (16 A–U and G–U pairs *versus* only 5 G–C pairs). Typical signals important for tertiary interactions are missing, like G18 and G19 in the D-loop, and the sequence UUC (the precursor of T C) in the T-loop. Only 5 out of the 9 tertiary interactions can be predicted on a theoretical basis. Altogether, this suggests a more flexible and less stable structure for mt-$tRNA^{Asp}$ than for classical tRNAs (6). In a more general perspective, mammalian mt-$tRNA^{Asp}$ from 136 organisms share only a few conserved nucleotides (Fig. 1C), including potential strong identity elements G34U35C36. To be noticed however, the absence of two expected identity elements, namely G73 and C38. These mt-tRNAs miss also the minor *E. coli* aspartate identity element G2–C71. Mammalian mt-$tRNA^{Asp}$ consensus sequence highlights R2–U71 (position of the minor aspartate element in *E. coli*) and A49–U65 as 100% conserved and thus, designate them as potential identity elements. Thus, mt-$tRNA^{Asp}$ either presents a reduced or an alternate identity set as compared to other aspartate tRNAs. In contrast, nuclear encoded human mt-AspRS, recently characterized (10), displays standard features and resembles known AspRSs. Markedly, it shares more than 40% of strictly identical residues with the *E. coli* enzyme and presents the typical bacterial insertion and C-terminal extension domains.

We show here that human mt-$tRNA^{Asp}$ conserves GUC aspartate anticodon as an identity element but not the discriminator position 73. This indicates a simplified set of aspartate identity elements in the human mt system. Sequence comparison and structure-based modeling of human mt-AspRS show

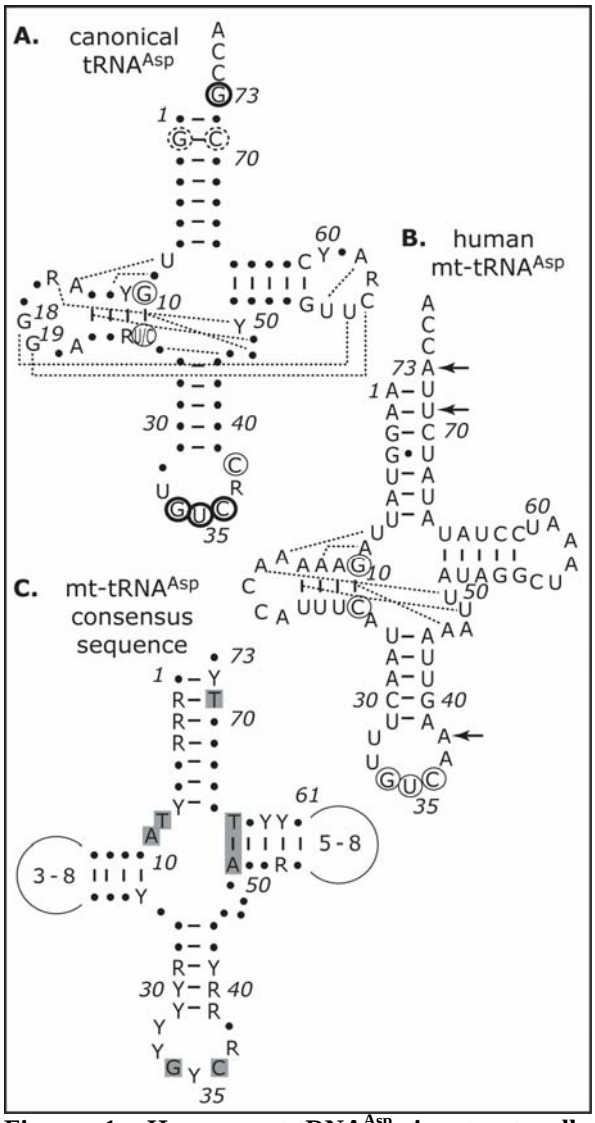



**Figure 1. Human mt-tRNA$^{Asp}$ is structurally bizarre and misses some classical aspartate identity elements.**

**A.** Cloverleaf representation of a canonical tRNA in which strictly conserved residues and tertiary interactions (dotted lines) are emphasized. Y and R stand respectively for pyrimidine and purine. Major (thick circles) and minor (thin circles) aspartate identity elements in yeast, *T. thermophilus*, and *E. coli* (18) are highlighted. An additional minor element specific to *E. coli* is in dotted circles (G2–C71). Numbering is according to ref. (20). **B.** Cloverleaf representation of human mt-tRNA$^{Asp}$ (51). Tentative tertiary interactions (dotted lines) are given. Notice the particular D- and T- loops. Expected identity elements are circled and missing elements pinpointed by arrows. **C.** Consensus sequence of mt-tRNA$^{Asp}$ genes from 136 mammals. Notice variations in size of D- and T-loops. Black dots stand for any nucleotide at a given position. The non-encoded CCA end is omitted.

local structural deviation in the mt-enzyme next to the discriminator position in the tRNA. In agreement with this view, a new mt-AspRS could be engineered that becomes discriminating for G73 after a single point mutation. This is in favor of a local and focused adaptation of the slowly evolving nuclear-encoded enzyme, which allows for non-importance of residue 73 rather than to recognition of an alternative nucleotide at this position. These data will be discussed in the light of the molecular evolution of tRNA aminoacylation systems in the peculiar mitochondrial context.

## Material and Methods

**Materials.** Synthetic genes of yeast (21), *E. coli* (15) and human mt (10) tRNA$^{Asp}$ were previously cloned. *T. thermophilus* tRNA$^{Asp}$ transcript and *E. coli* AspRS were gifts from H. Becker and G. Eriani. Human mt-tRNA$^{Tyr}$ transcripts were gifts from L. Bonnefond and J. Rudinger-Thirion. Human mt-AspRS was cloned (plasmid pQE70-mt-AspRS), overexpressed and purified as described (10). T7 RNA polymerase was purified as described (17) and oligonucleotides synthesized and purified by Sigma Genosys. Restriction enzymes (*BamH*I, *Hind*III, and *BstN*I) were from New England Biolabs, T4 DNA ligase from Qbiogen, and L-[$^3$H] aspartic acid (37 Ci/mmol) from Amersham.

**tRNA gene mutagenesis and *in vitro* transcription.** Human mt-tRNA$^{Asp}$ variant genes were obtained by hybridization of 9 overlapping oligonucleotides, ligation between *BamH*I and *Hind*III sites of plasmid pTFMa (21), and transformation into TG1 cells. Mutants of mt-tRNA$^{Asp}$ genes were obtained with the QuikChange$^{TM}$ Site-Directed Mutagenesis Kit (Stratagene). The corresponding synthetic genes contain a hammerhead ribozyme (22) and tRNA sequences downstream the T7 polymerase promoter. A *BstN*I site was inserted coincidental with the 3'-end of the tRNA sequences, which allows synthesis of tRNAs ending with the expected CCA sequence.

Transcription using T7 RNA polymerase was as described (22).

**mt-AspRS mutagenesis and expression.** Plasmid pQE70-mt-AspRS was converted into pQE70-mt-AspRS/D269 using the QuikChange™ Site-Directed Mutagenesis Kit (Stratagene), and transformation of *E. coli* TOP10 strain. Mt-AspRS/G269D was expressed and purified on nickel affinity column (10). Protein concentration was determined from $OD_{280}$ using an extinction coefficient calculated with ProtParam from Expasy tools ($\varepsilon$=43540 $M^{-1}.cm^{-1}$).

**Aminoacylation.** Assays were performed as described (10), in 50 mM HEPES-KOH pH 7,5, 25 mM KCl, 12 mM $MgCl_2$, 2.5 mM ATP, 0.2 mg/mL BSA, 1 mM spermine, 32 µM [$^3$H]-aspartic acid (208 GBq/mmol) and adequate amounts of transcripts and enzyme. Transcripts were renatured by heating at 60°C for 90 sec in water and slow cooling down to room temperature before aminoacylation performed at 15°C. Low temperature is required to obtain reproducible results. Maximal charging levels were determined for 400 nM transcript and 500 nM synthetase (either wt mt-AspRS or mt-AspRS/G269D). Synthetase activity was measured under excess substrate (8 µM) and 50 nM enzyme. Kinetic parameters $k_{cat}$ and $K_M$ were derived from Lineweaver-Burk plots obtained under adequate ranges of tRNA (40 nM to 1 µM) and AspRS (50 nM to 200 nM). Experiments were at least duplicated on separate preparations of transcripts. Initial rates for duplicates varied up to 30%.

**Mt-AspRS 3D modeling.** A 3D model of human mt-AspRS was built using *modeller* (23) on the basis of the crystallographic structure of *E. coli* AspRS/tRNA$^{Asp}$ complex, associated with aspartyl-adenylate (PDB id: 1C0A) (24), and sequence alignment of mt and *E. coli* enzymes (10). The putative mt-targeting signal (aa 1–47) was not included and the current model encompasses residues 49–640 corresponding to *E. coli* residues 1–585. Amino acid main and side chains were locally adjusted to allow docking of the tRNA as observed in the bacterial complex. Model of the dimer shown in Figure 3A was generated by 2-fold symmetry. Structure manipulation, superposition and analysis were performed with *O*, *moleman2* and *lsqman* (25). Figures were prepared with PyMol (DeLano Scientific LLC, CA).

## Results and Discussion

### Aspartate identity set is simplified in human mitochondria

***Discriminator base G73 is not an identity element in human mt-*tRNA$^{Asp}$.** Wild-type and a series of human mt-tRNA$^{Asp}$ variants were prepared by *in vitro* transcription and assayed for their aspartylation capacity at 15°C in the presence of recombinant human mt-AspRS. The low temperature is required for stabilization of the flexible tRNA structure, a known situation for mammalian mt-tRNAs (7, 10, 26). A first series of variants comprised single point mutations at positions corresponding to major aspartate identity determinants in yeast, *E. coli* and *T. thermophilus*, namely G34, U35, C36 in the anticodon loop, and residue 73 in the acceptor stem. While mutations at position 34 moderately decrease aminoacylation, those at positions 35 and 36 strongly impair it (Fig. 2A). Replacing A73 by G73, C73 or U73 keep significant aspartylation activities, with rates close to those of wt aminoacylation (Fig. 2B). Kinetic parameters ($K_M$ and $k_{cat}$) for each of the three variants are similar to those of the wt transcript within the experimental error limits (Table 1). This suggests that position 73 is no longer a strong identity element, at opposite to what occurs in other aspartate systems.

***Search for other aspartate determinants.*** The contribution of nucleotides known as minor aspartate identity elements in bacterial systems was also investigated.

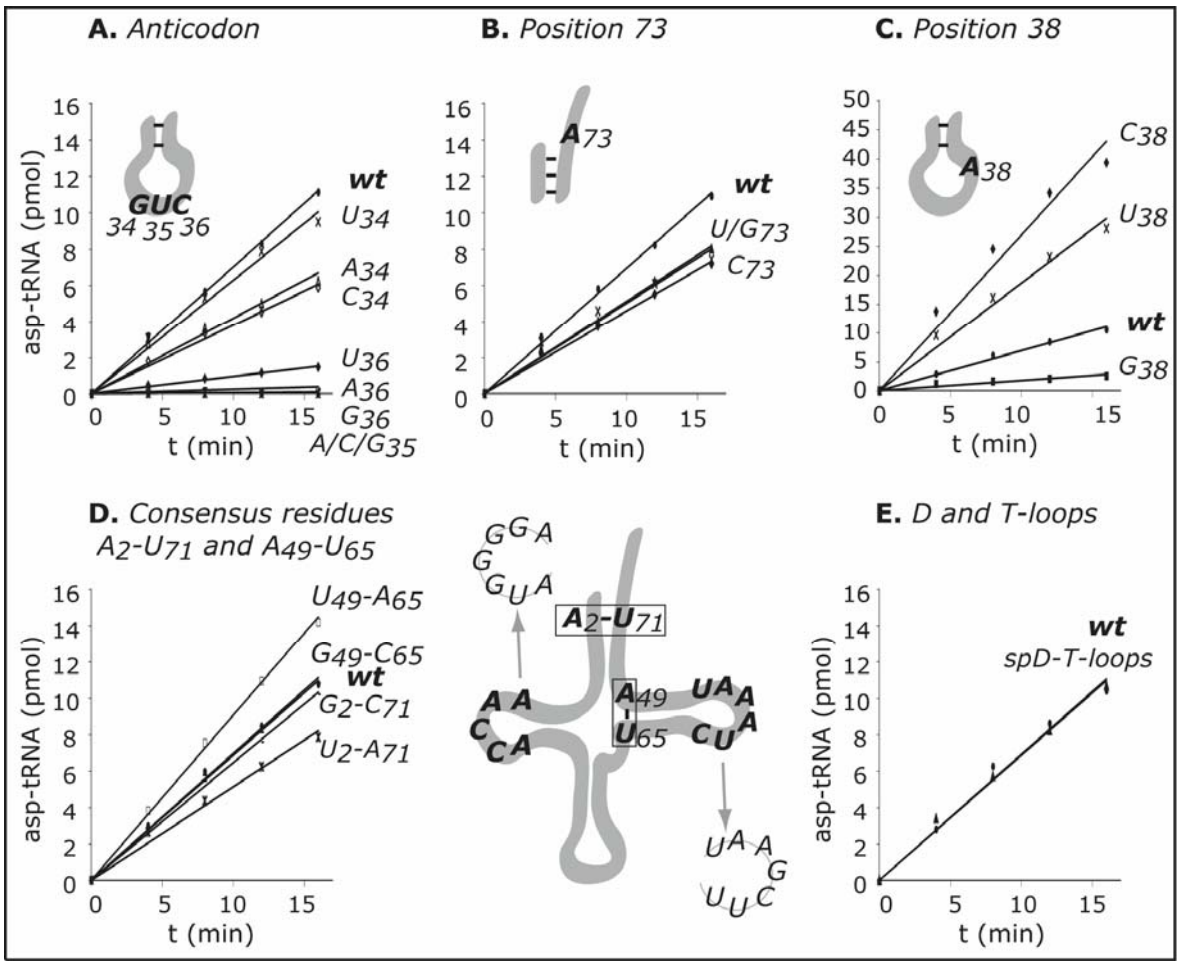



**Figure 2. Aspartylation of human mt-tRNA$^{Asp}$ variants with human mt-AspRS.**
Mutagenesis analysis on variants at N34, 35 or 36 (**A**), discriminator N73 (**B**), and anticodon loop N38 (**C**). Inserts recall localization of the tested positions in the tRNA. **D.** Variants with changes in the two conserved bp in 136 mammalian mt-tRNA$^{Asp}$ sequences (squared positions on the schematized cloverleaf). **E.** Structural variants of tRNA$^{Asp}$ with D- and T- loops from *S. pombe* tRNA$^{Asp}$ (see neighboring cloverleaf for sequences of these loops). Aminoacylation assays were performed as described under Materials and Methods, with 8 µM transcripts and 50 nM AspRS.

**Table 1. Kinetic parameters for tRNA aminoacylation by human mt-AspRS and mt-AspRS/G269D.**

| tRNA | $k_{cat}$ ($10^{-3}$ $sec^{-1}$) | $K_M$ (µM) | $k_{cat}/K_M$ ($10^{-3}$ $sec^{-1}.\mu M^{-1}$) |
|---|---|---|---|
| | ***wt mt-AspRS*** | | |
| ***wt tRNA$^{Asp}$*** | 5.3±1.7 | 0.58±0.26 | 9.1 |
| ***tRNA$^{Asp}$G73*** | 3.4±1.1 | 0.49±0.21 | 7 |
| ***tRNA$^{Asp}$C73*** | 3.0±1.0 | 0.31±0.17 | 9.7 |
| ***tRNA$^{Asp}$U73*** | 2.45±0.15 | 0.29±0.05 | 8.4 |
| | ***mt-AspRS/G269D*** | | |
| ***tRNA$^{Asp}$G73*** | 1.85±0.15 | 0.35±0.1 | 5.3 |
| ***wt tRNA$^{Asp}$*** | 0.6±0.15 | 0.3±0.1 | 2 |
| ***tRNA$^{Asp}$C73*** | 0.65±0.20 | 0.27±0.1 | 2.4 |
| ***tRNA$^{Asp}$U73*** | 0.7±0.25 | 0.58±0.3 | 1.2 |

Replacement of A38 by G38 decreases aminoacylation while replacement by C38 or U38 increases it (Fig. 2C). The surprising stimulation for variants C38 and U38 could either reveal insertion of a positive identity element, thus suggesting that the natural mt-tRNA sequence with A38 is not optimal for its cognate synthetase, or that a pyrimidine at

position 38 leads to structural rearrangements favoring synthetase recognition.

This last hypothesis has been tested by comparing solution structures of both wt mt-tRNA$^{Asp}$ and variant C38. The two transcripts share essentially similar reactivities against enzymatic probes, which is indicative of a common global secondary structure. They distinguish, however, from native mt-tRNA$^{Asp}$ (extracted from HeLa cell mitochondria), at the level of the D-domain (detailed data to be published elsewhere). Reactivities of nucleotides within the anticodon loop of the variants (Supporting Fig. 1) differ at subtle levels, those of variant C38 being at mid-way between those in the post-transcriptionnally modified native tRNA and in the "naked" wt transcript. This suggests that the increased aminoacylation capacity of variant C38 is in part linked to a structural rearrangement of its anticodon loop. The presence of C38 may mimic the structural role of post-transcriptional modifications present in the native tRNA. These modifications are not yet identified but are likely to occur as recently reported (Q34 and m$^1$G37) in nematode mt-tRNA$^{Asp}$ (27). Because of the importance of the anticodon loop in mt-tRNA$^{Asp}$ aminoacylation, the contribution of position 37 to aspartylation was also tested but none of the three nucleotides introduced in place of A37 had an effect (results not shown).

Base-pair G2–C71 is a minor identity element in *E. coli* tRNA$^{Asp}$ (14). In human mt-tRNA$^{Asp}$, the sequence is A2–U71, and R2–U71 in other mammalian mt-tRNA$^{Asp}$ (Fig. 1C) suggesting that it could be an important signal for mt-AspRS. Another strictly conserved signal is bp A49–U65 in the T-stem. Both pairs have been mutated and replaced by G–C or U–A pairs (Fig. 2D). Aminoacylation properties of the four new variants are close to those of the wt transcript showing that the corresponding wt nucleotides have no major contribution to mt aspartate identity.

An intriguing structural characteristic of mammalian mt-tRNAs is their unusual D- and T-loop ensemble with no classical size and no obvious signals for tertiary interactions. This has also been considered as a potential element specifying aspartylation (Fig. 2E). Both loops have been replaced at once by those of *Schizosaccharomyces pombe* cytosolic tRNA$^{Asp}$, where classical tertiary interactions can take place. Aspartylation of this variant is not affected (Fig. 2E).

In summary, aspartylation of human mt-tRNA$^{Asp}$ is restricted to the anticodon loop residues (with U35 and C36 as strongest elements, and G34 as moderate contributor). The discriminator residue 73, essential elsewhere (18), does not contribute. The structure of the anticodon loop is of importance. No additional identity element has been found (within the limit of the obvious candidates explored herein). Altogether, the mt aspartate identity set departs not only in a general manner from other aspartylation systems, but also more specifically from the *E. coli* system despite the high sequence similarity of *E. coli* and human mt-AspRSs. Base-pair 2–71 is not important and, within the anticodon triplet, the strength of the determinants is distinct: G34=U35>C36 in *E. coli*, and U35=C36>>G34 in human mitochondria. Finally, the mt enzyme is not sensitive to structural features at the level of the corner of its 3D structure since it aminoacylates efficiently a transcript transplanted with canonical D- and T-loops as well as *E. coli* tRNA$^{Asp}$ (10).

***Two synthetases of same evolutionary origin obey different aminoacylation rules.*** Non-conservation of major identity elements in aminoacylation systems of a same specificity has already been reported, and especially non-conservation of the discriminator base in different organisms (1). For example, specifity for prolylation requires A73 and C1–G72 in *E. coli* while the same positions are not involved in proline identity in the human cytosolic system (28). In the case of glycylation, U73 is a determinant for *E. coli* GlyRS while GlyRSs from archaea and eukaryotes require A73

(29). In these cases, variations in identity elements exist between a bacterial system at one hand and an archaeal/eukaryotic system at the other hand, where synthetases are rather distinct. In the present case, there are major changes in the identity elements for two AspRSs belonging to a same origin, the bacterial type human mt-AspRS and *E. coli* AspRS. A similar situation has been reported in *Arabidopsis thaliana*, where a prokaryotic-type mt-GlyRS is not sensitive to residue 73, while the *E. coli* enzyme is dependent on U73, a minor identity element (30).

**Understanding non-discrimination of residue 73 for recognition by mt-AspRS**

***Highly similar AspRSs with distinct functional behaviors***. Functional differences between *E. coli* and human mt-AspRSs have been searched for. Mt-AspRS was found to have a broad charging spectrum including, in addition to its cognate tRNA, $tRNA^{Asp}$ from *E. coli*, *T. thermophilus* and *S. cerevisiae* (Supporting Fig. 2A). In contrast, *E. coli* enzyme aminoacylates its cognate tRNA and $tRNA^{Asp}$ from human cytoplasm (not shown), but not the yeast (31) or the human mt-$tRNA^{Asp}$ (Supporting Fig. 2B). Non recognition of the mt-tRNA may simply be due to the absence of the major identity element G73 required for efficient activity of the *E. coli* enzyme. However, this enzyme does not recognize variant mt-$tRNA^{Asp}$ with G73 (charging level under 1%) (Supporting Fig. 2B) suggesting an unadapted structural scaffold and/or missing identity elements and/or presence of antideterminants in human mt-$tRNA^{Asp}$.

***Modeling the 3D structure of human mt-AspRS designates Gly269 as candidate for non-discrimination of G73.*** A 3D homology model of the mt enzyme has been built based on the crystallographic structure of *E. coli* AspRS in complex with its cognate tRNA (24). A surface representation of the mt-AspRS dimer is given in Figure 3A with different domains on one monomer cartooned and colored. In addition to the presence of modules featuring classical AspRSs (N-terminal anticodon binding domain / short hinge / C-terminal catalytic domain), the typical bacterial insertion (130 aa) and C-terminal extension (39 aa) domains are observed. The high level of sequence identity with the *E. coli* enzyme and the few major differences between the two enzymes are emphasized on a ribbon representation (Fig. 3A right).

Sequence comparison of mt and *E. coli* AspRSs (Fig. 3B) and the 3D model of the mt enzyme were used to identify differences potentially responsible for non-discrimination of position 73 by mt-AspRS. Residues of *E. coli* AspRS specifically contacting G73 are pictured in Figure 3C. Arginine (R222) and aspartate (D220) in motif 2 within the catalytic domain are conserved in all known bacterial AspRSs and establish specific contacts with G73. While R222 is conserved in mt-AspRS (R271), D220 is replaced by a glycine (G269), designating this residue as a potential candidate responsible for non discrimination of G73.

***Functional consequences.*** Based on the above conclusion, G269 was mutated and replaced by an aspartate residue, and aspartylation capacity of mt-AspRS/G269D tested in the presence of wt mt-$tRNA^{Asp}$ and the 3 variants of position 73 (C73, G73 and U73). In contrast to wt mt-AspRS which recognizes quite similarly these 4 substrates (Fig. 3D), mt-AspRS/G269D aminoacylates efficiently only variant G73 but poorly the 3 other tRNAs. Activity of mt-AspRS/G269D is ~10-fold decreased for wt mt-$tRNA^{Asp}$ but is clearly discriminative towards a substrate with G at position 73. Due to a $k_{cat}$ effect, variant G73 appears 2.5-fold more efficient for mt-AspRS/G269D than the 2 other variants and wt $tRNA^{Asp}$ (Table 1). Thus, the single mutation G269D converts mt-AspRS into a discriminating enzyme towards G73, most likely due to restoration of specific

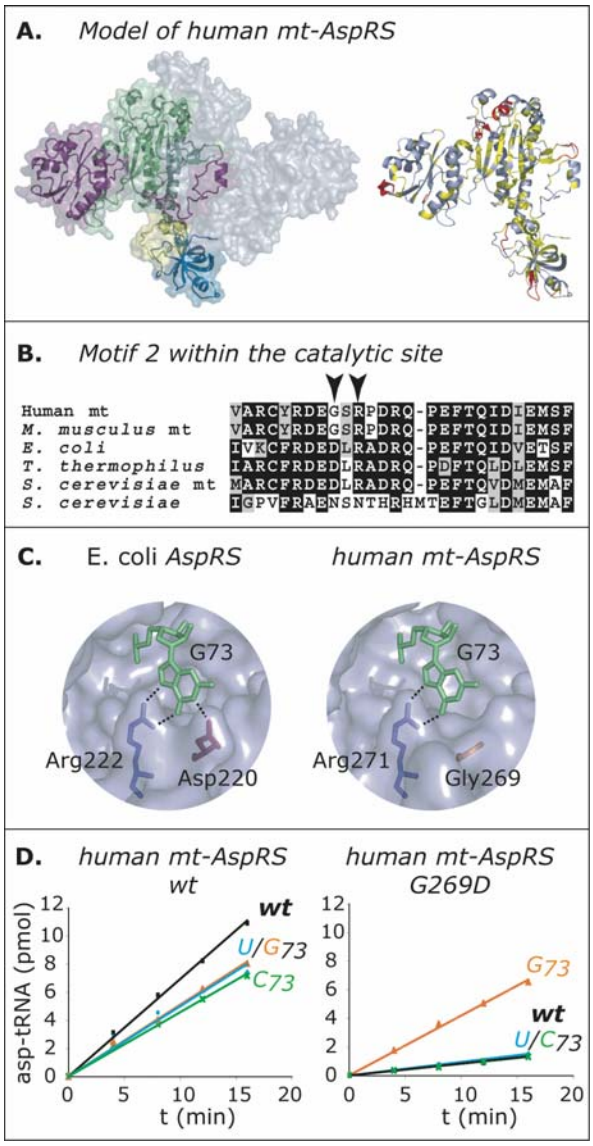



**Figure 3. Modeling the 3D structure of human mt-AspRS gives the clue of its relaxed specificity.**
**A.** 3D homology model of human mt-AspRS built from the crystallographic structure of *E. coli* AspRS in complex with its cognate tRNA (24). N-terminal anticodon binding domain (blue), short hinge (yellow), C-terminal active site (green), and bacterial insertion and C-terminal extension (purple) are cartooned on one monomer (left). Residues strictly conserved (in yellow) or regions strongly different (in red) from *E. coli* AspRS are emphasized on a cartoon representation of one monomer of mt-AspRS in which *E. coli* divergent loops have been superimposed (right). **B.** Sequence alignment of motif 2 of selected AspRSs (accession numbers: Human mt, translated from AL833633; *M. musculus* mt, NP_76623; *E. coli*, P21889; *T. thermophilus*, P36419; *S. cerevisiae* mt, NP_015221; *S. cerevisiae*, P04802). Black and gray boxes show strictly identical and similar residues, respectively. Arrowheads locate *E. coli* residues D220 and R222 specifically contacting base G73. **C.** Zooming into the binding pocket of residue G73 in the crystallographic structure of *E. coli* complex (left) and comparison with the same pocket in the human mt-AspRS model, in which the structure of *E. coli* $tRNA^{Asp}$ has been docked (right). **D.** Aminoacylation properties of wt (left) and mutated (right) mt-AspRSs, in the presence of *in vitro* transcribed wt and variant mt-$tRNA^{Asp}$ species. Reactions were performed under saturating tRNA concentrations (8 µM transcripts, 50 nM enzyme).

contacts with this nucleotide as supported by modeling. The presence of G269 in wt mt-AspRS reduces the number of direct contacts (H-bonds) with base 73 and thus may relax stringency with the protein region involved in mt-$tRNA^{Asp}$ 3'-end recognition. We note specificity alterations or recognition switches that were reported in non-mt aaRSs upon mutation of amino acids close to tRNA identity residues, either in the anticodon (*e.g.* 32-34) or like here the accepting stem (*e.g.* 35, 36).

***Mt-AspRS adaptation to mt-tRNA rapid evolution.*** Mitochondria originate from engulfment of bacteria by primitive eukaryotes with progressive transfer of endosymbiont genes to the nuclear genome, and the concomitant loss of other genes (8, 37). As a result, mammalian mt genomes became reduced to a set of genes coding for only 13 proteins, 2 rRNAs and 22 tRNAs. These genomes are subjected to a much higher mutational rate than nuclear genomes, due to the combination of an oxydative environment and limited DNA repair activity. Since mutational events appear to be more frequent in tRNA genes than in other parts of mt genomes (38), mammalian mt-tRNA show a characteristic sequence degeneracy leading to a strong polymorphism (39) and a diversity of pathology-related mutations (40, 41). In addition, the shift towards an increase of A/U content favors what could be called a structural polymorphism: mt-tRNA molecules display less stable 2D and more flexible 3D structures.
In contrast, mammalian mt-aaRSs are nuclear encoded and sustain a lower mutational rate. Their genes originate either from a transfer from the original endosymbiont, or from duplication of the initial nuclear gene coding

for the cytosolic synthetase of the host (42). In the present case, mt-AspRS is clearly of bacterial-type and its active site is close to that of *E. coli* (69% of the 89 aa are identical within a radius of 12 Å around the aspartate binding pocket). Hence, a limited number of mutations elsewhere along the tRNA interacting surface are probably sufficient to enable this mt enzyme to be more tolerant to sequence and conformation variations. This is first illustrated by the capacity of mt-AspRS to bind and charge a broad population of heterologuous tRNAs, whereas the *E. coli* ortholog cannot recognize mt-tRNAs. Second, the low sensitivity to the discriminator base (a phenotype which can be reverted by a single Gly-to-Asp mutation) is an indication that the rules of tRNA identity have been relaxed. In conclusion, the answer of the aspartate system to the mt context (sequence and structure polymorphisms) had consisted in the selection of a synthetase that works with a reduced set of identity elements and is more tolerant to structural fluctuations than homologous bacterial enzymes.

**Why is a restricted identity set sufficient in mitochondria?**

***Mischarging of mt-tRNAs by AspRS is restricted.*** Despite the fact that the mitochondrial translational machinery synthesizes only 13 proteins, it is assumed that mis-incorporation of amino acids into these proteins is as low as in bacteria or in eukaryotic cytosol. Restriction of the aspartate identity set to the anticodon loop in human mt-tRNA$^{Asp}$ raises questions about its sufficiency to guarantee specific aminoacylation in the organelle. There are only 22 tRNAs in mammalian mitochondria from which some share one or two nucleotides with tRNA$^{Asp}$ anticodon triplet (tRNA$^{Glu}$ with the strongest aspartate elements U35 and C36; tRNA$^{His}$, tRNA$^{Asn}$, tRNA$^{Tyr}$ with G34 and U35; and tRNA$^{Lys}$, tRNA$^{Val}$, and tRNA$^{Gln}$ with U35, C36 and U35, respectively). Attempts to aminoacylate *in vitro* transcribed human mt-tRNA$^{Glu}$, mt-tRNA$^{Tyr}$ or a variant mt-tRNA$^{Tyr}$ containing a GUC anticodon triplet, were not successful (results not shown). This suggests either the presence of antideterminants within these tRNAs, or limitations in conformational adaptation of these tRNAs required by mt-AspRS to fulfill non-cognate catalysis.

***Degeneracy of identity sets in mammalian mitochondria: a general trend?*** Only limited data are available on tRNA identity in mammalian mitochondria (43). However, it has been postulated that mt-tRNAs may contain less sequence-specific positive recognition elements in addition to anticodon elements because individual mt-tRNAs are so distinguishable that intricate recognition mechanisms might be unnecessary (9). This view becomes now experimentally validated for a few systems. In addition to the aspartate system, human bacterial-type mt-TyrRS does not discriminate bp 1–72 despite it is a strong identity element for *E. coli* TyrRS (44), and bovine mt-SerRS does not require the variable domain but rather the T-loop (especially A58) and nucleotides replacing the D-loop (45). Such restricted identity sets in mt-tRNAs may be correlated with the lower catalytic efficiency of mt aminoacylation systems as compared to bacterial or eukaryotic cytosolic systems (7).

## Evolution and mitochondria

Evolution of aminoacylation systems has been discussed at several instances (46, 47). Primordial systems were likely composed of minihelices (“ur” acceptor-stems) recognized by aaRS ancestors restricted to their catalytic core (48). An early “operational RNA code” typically formed of the 3 first tRNA acceptor stem base-pairs and of residue N73, fairly named “discriminator base”, may have preceded emergence of the genetic code (49). Subsequent evolution of aminoacylation systems was driven by insertion of additional tRNA domains and aaRS modules with concomitant *expansion* of tRNA identity sets, now composed of conserved major elements and idiosyncratic or species-specific minor

features. The present study on a mt aspartylation system reveals the loss of a strongly conserved major identity element, G73. This *restriction* in the identity set suggests a reverse evolution of this aminoacylation system in regard of its bacterial-type ancestor. Residue G73 was likely the primordial asparate identity element since it allows specific aminoacylation of minimalist tRNA substrates (50). Its loss as a major aspartate element from bacteria (*i.e. E. coli*) to bacteria-like (human mt) systems illustrates a rupture of the operational RNA code within afore mentioned evolutionary-related biological systems. This situation was likely driven by the high mutational rate of the mt genome, but also by a less stringent need of intricate identity rules due to a simplified tRNA world in mitochondria. The synthetase has adapted to allow for preservation of complementarity with its substrate by mutation of at least a critical aspartate residue into a glycine. We conjecture that due to high tRNA degeneracy, reverse evolution of mt aminoacylation systems is a general event in mammals.

## Acknowledgments

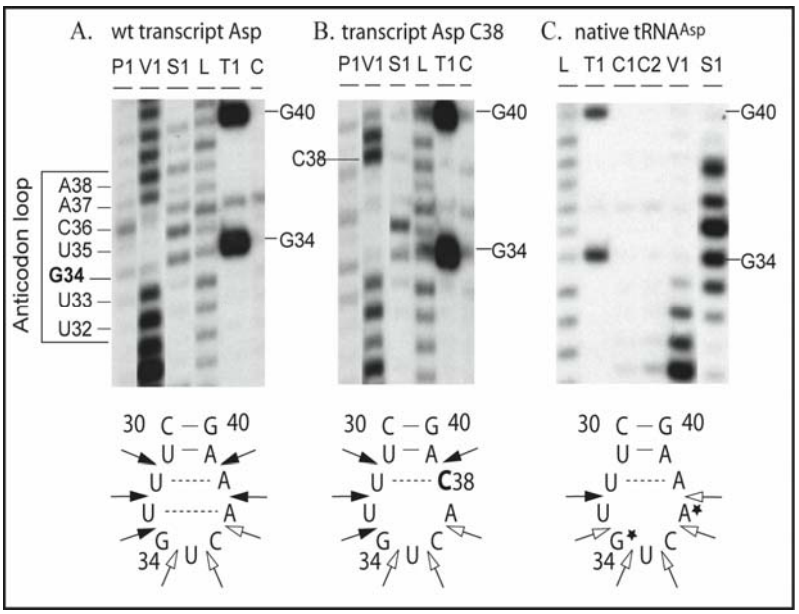
J. Rudinger-Thirion and L. Bonnefond are acknowledged for numerous discussions, A. Clénet for technical assistance, and G. Eriani, H. Becker and D. Kern for gifts of material. This work was supported by Centre National de la Recherche Scientifique (CNRS), Université Louis Pasteur Strasbourg (ULP), Association Française contre les Myopathies (AFM) and Action Concertée Incitative (ACI-BCMS 146). A.F. was supported by a fellowship from Ministère de l'Enseignement Supérieur et de la Recherche. C.S. was the recipient of a Marie Curie European Reintegration Grant (MERG-CT-2004-004898).

## References

1. Giegé, R. & Frugier, M. (2003) in *Translation Mechanisms*, eds. Lapointe, J. & Brakier-Gringas, L. (Landes Sciences, Georgetown, TX) pp. 1-24.
2. Saks, M. E. & Sampson, J. R. (1996) *EMBO J.* **15,** 2843-2849.
3. Frugier, M., Helm, M., Felden, B., Giegé, R. & Florentz, C. (1998) *J. Biol. Chem.* **273,** 11605-11610.
4. Giegé, R., Frugier, M. & Rudinger, J. (1998) *Curr. Opin. Struct. Biol.* **8,** 286-293.
5. Hauenstein, S., Zhang, C. M., Hou, Y. M. & Perona, J. J. (2004) *Nature Struct. Mol. Biol.* **11,** 1134-1141.
6. Helm, M., Brulé, H., Friede, D., Giegé, R., Pütz, J. & Florentz, C. (2000) *RNA* **6,** 1356-1379.
7. Florentz, C., Sohm, B., Tryoen-Tóth, P., Pütz, J. & Sissler, M. (2003) *Cell. Mol. Life Sci.* **60,** 1356-1375.
8. Margulis, L. (1981) *Symbiosis in cell evolution* (Freeman, W.H., San Francisco).
9. Kumazawa, Y., Himeno, H., Miura, K.-I. & Watanabe, K. (1991) *J. Biochem.* **109,** 421-427.
10. Bonnefond, L., Fender, A., Rudinger-Thirion, J., Giegé, R., Florentz, C. & Sissler, M. (2005) *Biochemistry* **44,** 4805-4816.
11. Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990) *Nature* **347,** 203-206.
12. Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. & Leberman, R. (1990) *Nature* **347,** 249-255.
13. Hasegawa, T., Himeno, H., Ishikura, H. & Shimizu, M. (1989) *Biochem. Biophys. Res. Comm.* **163,** 1534-1538.
14. Nameki, N., Tamura, K., Himeno, H., Asahara, H., Hasegawa, T. & Shimuzu, M. (1992) *Biochem. Biophys. Res. Comm.* **189,** 856-862.
15. Pütz, J., Puglisi, J. D., Florentz, C. & Giegé, R. (1991) *Science* **252,** 1696-1699.
16. Ryckelynck, M., Giegé, R. & Frugier, M. (2003) *J. Biol. Chem.* **278,** 9683-9690.
17. Becker, H., Giegé, R. & Kern, D. (1996) *Biochemistry* **35,** 7447-7458.
18. Giegé, R., Florentz, C., Kern, D., Gangloff, J., Eriani, G. & Moras, D. (1996) *Biochimie* **78,** 605-623.
19. Giegé, R. & Rees, B. (2005) in *Aminoacyl-tRNA Synthetases*, eds. Ibba, M., Francklyn, C. & Cusack, S. (Landes

Bioscience, Georgetown, TX) pp. 210-226.
20. Sprinzl, M. & Vassilenko, K. S. (2005) *Nucleic Acids Res.* **33,** D139-D140.
21. Perret, V., Garcia, A., Puglisi, J. D., Grosjean, H., Ebel, J.-P., Florentz, C. & Giegé, R. (1990) *Biochimie* **72,** 735-744.
22. Fechter, P., Rudinger, J., Giegé, R. & Théobald-Dietrich, A. (1998) *FEBS Lett.* **436,** 99-103.
23. Sali, A. & Blundell, T. L. (1993) *J. Mol. Biol.* **234,** 779-815.
24. Eiler, S., Dock-Bregeon, A.-C., Moulinier, L., Thierry, J.-C. & Moras, D. (1999) *EMBO J.* **18,** 6532-6541.
25. Kleywegt, G. J., Zou, J. Y., Kjeldgaard, M. & Jones, T. A. (2001) in *International Tables for Crystallography*, eds. Rossmann, M. G. & Arnold, E. (Kluwer Academic, Dordrecht, The Netherlands), Vol. F. Crystallography of Biological Macromolecules, pp. 353-356, 366-367.
26. Sissler, M., Helm, M., Frugier, M., Giegé, R. & Florentz, C. (2004) *RNA* **10,** 841-853.
27. Sakurai, M., Ohtsuki, T. & Watanabe, K. (2005) *Nucleic Acids Res.* **33,** 1653-1661.
28. Stehlin, C., Burke, B., Yang, F., Liu, H., Shiba, K. & Musier-Forsyth, K. (1998) *Biochemistry* **37,** 8605-8613.
29. Shiba, K., Motegi, H. & Schimmel, P. (1997) *Trends Biochem. Sci.* **22,** 453-457.
30. Duchêne, A. M., Peeters, N., Dietrich, A., Cosset, A., Small, I. D. & Wintz, H. (2001) *J. Biol. Chem.* **276,** 15275-83.
31. Moulinier, L., Eiler, S., Eriani, G., Gangloff, J., Thierry, J.-C., Gabriel, K., McClain, W. H. & Moras, D. (2001) *EMBO J.* **20,** 5290-5301.
32. Schmitt, E., Meinnel, T., Panvert, M., Mechulam, Y. & Blanquet, S. (1993) *J. Mol. Biol.* **233,** 615-628.
33. Auld, D. S. & Schimmel, P. (1995) *Science* **267,** 1994-6.
34. Auld, D. S. & Schimmel, P. (1996) *EMBO J.* **15,** 1142-1148.
35. Wakasugi, K., Quinn, C. L., Tao, N. & Schimmel, P. (1998) *EMBO J.* **17,** 297-305.
36. Connolly, S. A., Rosen, A. E., Musier-Forsyth, K. & Francklyn, C. S. (2004) *Biochemistry* **43,** 962-969.
37. Gray, M. W., Burger, G. & Lang, B. F. (2001) *Genome Biology* **2,** 1018.1-1018.5.
38. Lynch, M. (1996) *Mol. Biol. Evol.* **13,** 209-220.
39. Ingman, M., Kaessmann, H., Pääbo, S. & Gyllenstein, U. (2000) *Nature* **408,** 708-713.
40. Schon, E. A., Bonilla, E. & DiMauro, S. (1997) *J. Bioenerg. Biomemb.* **29,** 131-149.
41. Wallace, D. C. (1999) *Science* **283,** 1482-1488.
42. Sissler, M., Pütz, J., Fasiolo, F. & Florentz, C. (2005) in *Aminoacyl-tRNA synthetases*, eds. Ibba, M., Francklyn, C. & Cusack, S. (Landes Biosciences, Georgetown, TX), pp. 271-284.
43. Levinger, L., Mörl, M. & Florentz, C. (2004) *Nucleic Acids Research* **32,** 5430-5441.
44. Bonnefond, L., Frugier, M., Giegé, R. & Rudinger-Thirion, J. (2005) *RNA* **11,** 558-562.
45. Ueda, T., Yotsumoto, Y., Ikeda, K. & Watanabe, K. (1992) *Nucleic Acids Res.* **20,** 2217-2222.
46. Woese, C. R., Olsen, G. J., Ibba, M. & Söll, D. (2000) *Microbiol. and Mol. Biol. Reviews* **64,** 202-236.
47. Ribas de Pouplana, L. & Schimmel, P. (2000) *Cell Mol. Life Sci.* **57,** 865-670.
48. Schimmel, P., Giegé, R., Moras, D. & Yokoyama, S. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8763-8768.
49. Schimmel, P. (1995) *J. Mol. Evol.* **40,** 531-536.
50. Frugier, M., Florentz, C. & Giegé, R. (1994) *EMBO J.* **13,** 2218-2226.
51. Anderson, S., Bankier, A. T., Barrel, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, J. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1981) *Nature* **290,** 457-465.

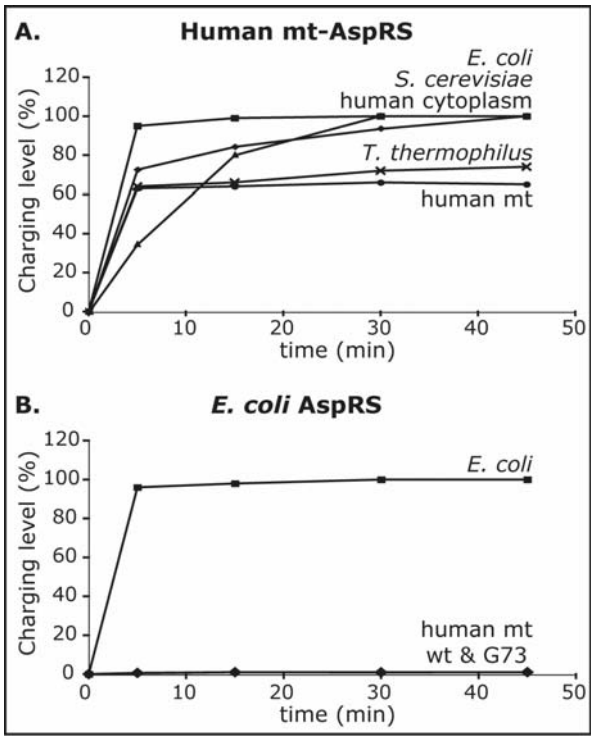
# Supporting Information




**Supp. Figure 1. Structural investigation of the anticodon loop of human mt native, *in vitro* transcribed and mutated tRNA$^{Asp}$.**

**Upper part.** Autoradiographs of denaturing 12% PAGE showing enzymatic digestion of 5' labeled (from nts 30 to 40) human wt mt tRNA$^{Asp}$ transcript (**A**), the related A/C38 variant (**B**) and native mt tRNA$^{Asp}$ (**C**). Digestions were performed with nucleases V1 (V1), S1 (S1) and with ribonuclease T1 (T1). Ladders were obtained after nuclease P1 (P1) or alkaline treatment (L). Control lanes (C) contained only RNA, but no added nuclease. Anticodon loop region is delineated at the left and G residues are indicated at the right side of each panel. **Lower part.** The mapping data are summarized on the corresponding 2D anticodon stem-loop models. Black and white arrowheads stand respectively for accessibilities by V1 and S1 nucleases. Possible post-transcriptional modifications are displayed by stars and additional potential base-pairings are highlighted by dotted lines.

**Experimental procedures.** The structure of wt and variant A38C as well as of native tRNA$^{Asp}$ extracted from HeLa cell mitochondria (pmolar amounts obtained by hybridization to a complementary oligonucleotide, Yokogawa *et al.* 1989 *Nucleic Acids Res.* **17,** 2623-2638.) was determined according to established procedures (Giegé *et al.* 1999 in *Prebiotic Chemistry, Molecular Fossils, Nucleosides, and RNA.*, eds. Söll, D., Nishimura, S. & Moore, P. (Pergamon, Oxford), Vol. 6, pp. 63-80). Briefly, the [$^{32}$P] end-labeled tRNA is submitted to statistical cleavage by enzymatic (S1 and V1 nucleases) or chemical (dimethylsulfate, diethylpyrocarbonate) probes. Reactivity of individual targets is detected after running the samples on 12% polyacrylamide/8M urea gels followed by autoradiography. The structures were probed under the same buffer conditions as those used for aminoacylation assays and the same process for renaturation.




**Supp. Figure 2. Broad substrate range of human mt-AspRS.** Charging spectrums of human mt (**A**) and *E. coli* (**B**) AspRSs. Aminoacylation assays were performed under conditions described in Material and Methods section, in the presence of 0.4 µM transcripts and 500 nM human mt-AspRS (**A**) or 200 nM *E. coli* AspRS (**B**).

# 1. Identité des ARNt$^{Asp}$ mt de mammifères

## 1.1. La perte du résidu 73 comme élément d'identité est spécifique des ARNt$^{Asp}$ mt de mammifères

Les nombreuses études concernant l'identité des ARNt canoniques ont montré que pour une spécificité d'aminoacylation donnée, les éléments d'identité sont globalement conservés au cours de l'évolution (Giegé et al., 1998). Dans le cas du système d'aspartylation, les éléments d'identité ont été déterminés pour les ARNt$^{Asp}$ d'*E. coli*, de *S. cerevisiae* et de *T. thermophilus* et sont les éléments majeurs G34U35C36 et G73, et les éléments mineurs G10-U/C25 et C38 (revue par (Giegé et al., 1996)). Nos travaux montrent que ce jeu n'est pas maintenu en l'état dans l'ARNt$^{Asp}$ mt humain ni dans ceux des autres ARNt$^{Asp}$ mt de mammifères.

Nous avons recherché la présence de ces éléments d'identité dans l'ensemble des séquences d'ARNt$^{Asp}$ disponibles dans la banque de Sprinzl (Sprinzl & Vassilenko, 2005), et compilées dans la banque mamit mise en place par Joern Pütz au laboratoire (http://mamit-trna.u-strasbg.fr). L'ensemble de ces informations sont exposées dans le Tableau 5. Globalement, les éléments d'identité aspartate sont conservés dans toutes les séquences de gènes d'ARNt$^{Asp}$, exceptés ceux d'organelles avec le cas extrême des gènes d'ARNt$^{Asp}$ de mitochondries de mammifères pour lesquels seuls 3 éléments d'identité sont strictement retrouvés (G34, U36 et U/C25). La variation de séquence observée au niveau de la position 35 concerne le gène de l'ARNt$^{Asp}$ mt chez les marsupiaux qui possède un C, édité partiellement en U après transcription et donne ainsi lieu à la fois à l'ARNt$^{Gly}$ (non édité) et à l'ARNt$^{Asp}$ (après édition) (Mörl et al., 1995). La perte du résidu G en position 73 ne concerne que les ARNt$^{Asp}$ mt de mammifères. La perte du C en position 38 ne concerne que ceux d'organelles d'eucaryotes supérieurs.

| **organismes** | 10 | 25 | **34** | **35** | **36** | 38 | **73** |
|---|---|---|---|---|---|---|---|
| *cyt pluricellulaires (8)* | G | U | G | U | C | C | G |
| *cyt unicellulaires (6)* | G | U | G | U | C | C | G |
| *archaea (6)* | G | U | G | U | C | C | G |
| *bactéries (20)* | G | U/C | G | U | C | C | G |
| *chloroplastes (13)* | G | C | G | U | C | **A*** | G |
| *mt plantes (6)* | G | U/C | G | U | C | **C/A** | G |
| *mt unicellulaires (11)* | G* | U* | G | U | C | **C*** | G |
| *mt mammifères (136)* | G* | U/C | G | U* | C | **A*** | A/U/C |

**Tableau 5. Recherche des éléments d'identité aspartate classiques dans 206 séquences d'ARNt$^{Asp}$.** Les chiffres entre parenthèses indiquent le nombre d'organismes dont la séquence du gène de l'ARNt$^{Asp}$ est disponible et a été pris en compte pour l'analyse. Les positions d'identité majeures sont indiquée en gras. *nucléotide majoritairement retrouvé à cette position. -/- indique que les nucléotides sont retrouvés de façon équivalente. cyt signifie cytoplasme.

## 1.2. Eléments d'identité complémentaires ?

L'absence de rôle pour le nucléotide 73 et la restriction du jeu d'éléments d'identité à la boucle de l'anticodon pose la question de la suffisance de ce jeu. Ce point a déjà été discuté dans la publication. Il reste possible que ce jeu soit complété par des éléments encore à découvrir.

Dans l'hypothèse d'un jeu d'identité alternatif, le bras D pourrait jouer un rôle complexe. En effet, le remplacement de la boucle D par celle d'un ARNt$^{Asp}$ classique (ici de *Schizosaccharomyces pombe*), non accompagnée de l'introduction de la boucle T correspondante, a un impact négatif non négligeable. La vitesse réactionnelle de l'AspRS mt en présence d'une quantité saturante d'ARNt variant (8 µM de transcrit et 50 nM d'enzyme) est réduite de 5 fois par rapport à la vitesse associée au substrat sauvage. Cependant, on ne sait si les effets de l'introduction de la boucle sont dus au changement de séquence ou tout simplement à un mauvais repliement de l'ARNt. L'étude de la structure de l'ARNt variant en solution devrait permettre de statuer. Un autre argument en faveur de l'implication de la boucle D dans l'identité aspartate vient des comparaisons de séquences de divers humains. De façon remarquable, 9 mutations polymorphiques (variations de séquences d'un individu à l'autre) ont été répertoriées dans la séquence de l'ARNt$^{Asp}$ mt humain (Figure 23). Aucune de

ces mutations n'a été trouvée dans le domaine D de l'ARNt. Ainsi, la séquence de ce domaine pourrait jouer un rôle important, non encore identifié.

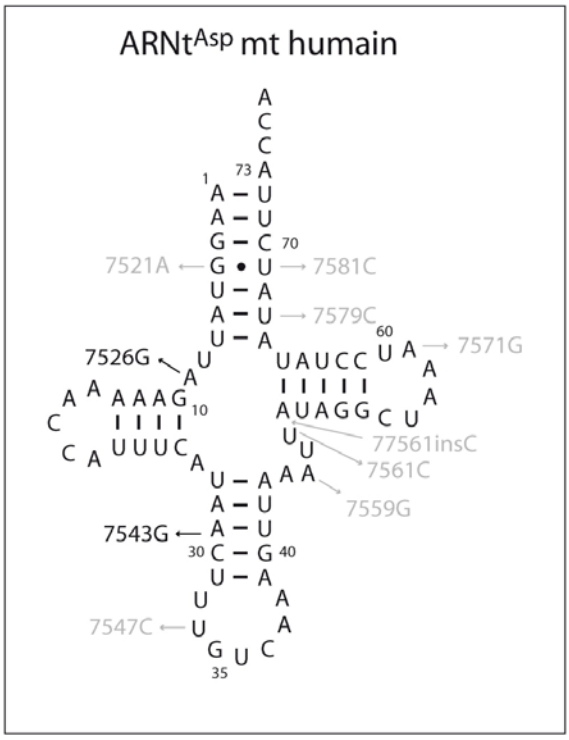



**Figure 23. ARNt$^{Asp}$ mitochondrial humain et mutations.** Localisation des mutations polymorphiques en gris (variations de séquence d'un individu à l'autre, (Ingman et al., 2000)), et pathogéniques en noir (corrélées à une maladie). La numérotation de ces mutations est selon leur localisation dans le génome mt (Anderson et al., 1981). La mutation 7543G initialement décrite comme pathogénique (Ingman et al., 2000) a été réassignée comme polymorphique (McFarland et al., 2004). La mutation 7526G vient d'être corrélée à une myopathie (Seneca et al., 2005).

Dans l'hypothèse d'un jeu d'identité restreint par rapport à ceux déterminés pour les systèmes classiques, le problème de spécificité au sein de la mitochondrie se pose. En effet, d'autres ARNt possèdent des anticodons très proches. Par exemple l'ARNt$^{Glu}$ présente également les résidus U35 et C36. De plus, il possède un U en position 34 qui, d'après nos tests expérimentaux est toléré dans l'ARNt$^{Asp}$ mt humain (la mutation G/U34 n'affecte pas significativement l'aminoacylation du variant par l'AspRS mt). Tous ces éléments font de l'ARNt$^{Glu}$ un substrat potentiel pour l'AspRS dans la mitochondrie. Néanmoins, nous avons pu montrer que l'AspRS mt n'est pas capable d'aminoacyler l'ARNt$^{Glu}$ d'*E. coli* transcrit *in vitro* ou natif. Ceci laisse supposer par exemple qu'il existe des antidéterminants au sein de l'ARNt$^{Glu}$ qui l'empêchent d'être reconnu par l'AspRS mt.

Finalement, l'approche *in vitro* ne nous a pas permis d'aborder l'implication des bases modifiées dans l'identité aspartate du système mt humain. Les modifications post-

transcriptionnelles des ARNt mt sont peu connues. Cependant, elles pourraient avoir des rôles importants. Ainsi, la structuration de l'ARNt$^{Lys}$ en feuille de trèfle dépend d'une modification dans le connecteur entre les bras D et T (Helm et al., 1999), et une sous-modification de la position 34 de l'anticodon de l'ARNt$^{Leu}$ est corrélée à une pathologie (Kirino et al., 2005). Dans notre cas, il sera intéressant de séquencer l'ARNt$^{Asp}$ afin de déterminer ses modifications post-transcriptionnelles.

## 1.3. ARNt$^{Asp}$ et mutations pathogéniques

Lorsque nous avons commencé l'étude de ce système, une mutation de l'ARNt$^{Asp}$ mt humain avait été décrite comme associée à une pathologie dont les symptômes sont graves et se traduisent par un retard général dans le développement (Shtilbans et al., 1999). Il s'agit de la mutation A7543G (numérotation dans le génome mt, (Anderson et al., 1981)). Celle-ci se trouve dans le bras de l'anticodon de l'ARNt et transforme la paire de bases A29-U41 en paire de bases G29•U41 (Figure 23). J'ai analysé son impact sur l'aminoacylation a été analysé. La vitesse maximale mesurée pour l'AspRS mt humaine en présence du transcrit *in vitro* de l'ARNt$^{Asp}$ muté est similaire à celle du transcrit sauvage. Ainsi, cette mutation n'affecte pas l'aminoacylation de l'ARNt *in vitro*. Elle pourrait toutefois altérer une autre étape de la vie de l'ARNt et ainsi causer la pathologie. Cependant, les critères pris en compte pour l'assignation d'une mutation comme étant pathologique ont été redéfinis récemment de manière plus stringente (McFarland et al., 2004). Dans ce contexte, la mutation a été réassignée comme polymorphique et donc sans conséquence pour la santé. Ceci rejoint nos résultats.

Récemment, une nouvelle mutation (A7526G) dans le gène de l'ARNt$^{Asp}$ mt humain (position 9 dans le connecteur entre le bras accepteur et D de l'ARNt), a été corrélée à une myopathie caractérisée par une intolérance aiguë à l'effort (Seneca et al., 2005) (Figure 23). Il serait intéressant d'analyser l'impact de cette mutation au niveau de l'aminoacylation mais aussi au niveau structural, puisque cette position est classiquement impliquée dans une interaction tertiaire dans les ARNt classiques.

## 2. Identité des autres systèmes mitochondriaux de mammifères

La découverte d'un jeu d'identité restreint pour l'ARNt$^{Asp}$ mt humain (et de mammifères) soulève la question de la situation dans les ARNt des autres systèmes mt d'aminoacylation. L'identité des systèmes mitochondriaux est encore peu étudiée (*e.g.* revue par (Levinger et al., 2004)). Elle a été explorée d'un point de vue fondamental au niveau des systèmes sérine (Ueda et al., 1992), tyrosine (Bonnefond et al., 2005a) et aspartate (Börner et al., 1996, article n°4). Dans le cas de l'ARNt$^{Ser(AGY)}$ de foie de bœuf, un rôle prépondérant de la boucle T (en particulier du nucléotide A58) et du bras D tronqué ont été mis en évidence. L'identité de l'ARNt$^{Tyr}$ mt humain repose majoritairement sur la base discriminatrice A73 et l'anticodon GUA (comme pour les autre systèmes classiques) mais ne met plus en jeu la paire de bases G1-C72 spécifiques des ARNt$^{Tyr}$ procaryotiques (Bonnefond et al., 2005a ; Bonnefond et al., 2005b). D'autre part, l'identité a été explorée via l'étude de l'impact des mutations pathologiques sur l'aminoacylation dans le cas des systèmes leucine (Sohm et al., 2003), isoleucine (Kelley et al., 2000 ; Kelley et al., 2001) et lysine (Sissler et al., 2004). L'analyse de 10 mutations pathogéniques différentes au sein de l'ARNt$^{Ile}$ a montré que les mutations T4274C (en position 12 du bras D), T4285C (position 27 du bras anticodon) et G4298A (position 40 du bras anticodon) affectent significativement l'aminoacylation (Kelley et al., 2000 ; Kelley et al., 2001). Au laboratoire, Bénédicte Sohm a montré que la mutation A3243T (en position 14 de la boucle D) associée au syndrome MELAS réduit de plus de 300 fois l'efficacité d'aminoacylation de l'ARNt$^{Leu(UUR)}$ par la LeuRS (Sohm et al., 2003), et Marie Sissler a mis en évidence l'effet délétère des mutations G8313A (en position 24 du bras D) et G8328A (en position 30 du bras anticodon) sur l'efficacité de l'ARNt$^{Lys}$ (Sissler et al., 2004). Ces différentes études pourraient suggérer que ces positions sont impliquées dans les jeux d'identité de ces ARNt.

Considérant l'origine endosymbiotique de la mitochondrie, la recherche des éléments d'identité déterminés pour les systèmes d'*E. coli* avaient été effectuée précédemment au laboratoire au sein des séquences consensus d'ARNt mt de 31 mammifères (Helm et al., 2000). Il avait été montré que les ARNt mt contiennent, au moins théoriquement, une bonne part des éléments d'identité majeurs d'*E. coli*. Grâce au séquençage de plus en plus de génome mt de mammifères, cette étude peut être étendue aujourd'hui à 136 séquences d'ARNt$^{Asp}$ mt et 71 séquences pour les ARNt d'autres spécificités (compilés par Joern Pütz au

laboratoire). Le Tableau 6 rappelle les jeux d'identité déterminés expérimentalement chez *E. coli* et donne les éléments effectivement retrouvés aux mêmes positions dans les séquences consensus des ARNt mt de mammifères.

De façon globale, il y a différents degrés de conservation des éléments d'identité parmi les systèmes mt de mammifères. Les cas opposés et extrêmes concernent la conservation de tous les éléments d'identité d'*E. coli* (cas du système Asn), à leur absence totale (cas du système Ala). Pour les systèmes Lys, Pro, Arg et His, seul un élément est manquant. Pour les autres systèmes, le nombre d'éléments conservés est très variable, avec par exemple le système Asp qui en conserve 4 sur 7, et le système Gly où seulement 1 des 6 éléments est retrouvé. Notons que les modifications post-transcriptionnelles ne sont pas ou peu connues et que par conséquent leur conservation ou non en tant qu'éléments d'identité n'a pu être déterminée.

Les éléments d'identité d'*E. coli* conservés dans les ARNt mt de mammifères concernent essentiellement les éléments d'identité aux deux extrémités de l'ARNt (extrémité acceptrice et triplet anticodon). Cependant, certains systèmes dévient fortement puisqu'ils ne présentent plus certains des éléments majeurs d'identité qui pourtant sont conservés au cours de l'évolution. Par exemple, notons la non conservation de la paire de bases G3•U70 qui spécifie l'alanylation des $ARNt^{Ala}$ classiques ou l'absence des éléments structuraux participant à l'identité des $ARNt^{Cys}$ (paire de Levitt G15•G48) et $ARNt^{Ser}$ (grande région variable) d'*E. coli*. Pour certains systèmes alanine mt non mammifère, la paire de bases G•U est soit conservée en position 3-70 (cas chez *Caenorhabditis elegans*, (Chihade et al., 1998)), soit déplacée en position 2-71 (cas chez *Drosophila melanogaster*, (Lovato et al., 2001)). Dans les deux cas, cette paire de bases conserve son rôle dans l'identité de l'$ARNt^{Ala}$. En ce qui concerne les éléments de l'anticodon, notons la non conservation du triplet anticodon dans les $ARNt^{Thr}$ et $ARN^{Trp}$ du fait qu'ils lisent des codons déviant par rapport au code génétique universel.

L'élément d'identité Ile U12, qui est non conservé dans les ARNt mt de mammifères, est retrouvé dans l'$ARNt^{Ile}$ mt humain et participe à l'identité (Kelley et al., 2000 ; Kelley et al., 2001). Un exemple du même type existe pour les système Leu, pour lequel l'élément d'identité A14 n'est pas conservé dans les ARNt mt de mammifères mais spécifie l'aminoacylation de l'$ARNt^{Leu}$ mt humain par la LeuRS (Sohm et al., 2003). Par contre, la paire de bases G2-C71, élément d'identité Tyr qui n'est pas conservé dans les $ARNt^{Tyr}$ mt de mammifères, mais qui est retrouvée dans l'$ARNt^{Tyr}$ mt humain n'est plus un élément d'identité (Bonnefond et al., 2005a). Enfin, nos travaux sur le système aspartate montrent que

les positions non conservées 73 et 38 n'ont plus de rôle dans l'identité suggérant soit que le jeu est simplifié et que l'aaRS n'a pas besoin d'autant de signaux pour reconnaître son ARNt de façon spécifique, soit que le jeu est alternatif. Notons que le cas d'un jeu alternatif a déjà été montré pour l'ARNt$^{Ser(AGY)}$ mt bovin qui présente un jeu d'identité distinct de celui déterminé pour l'ARNt$^{Ser}$ d'*E. coli* (Ueda et al., 1992).

Jeux d'identité alternatifs ou restreints, il n'existe pas encore de consensus sur cette question car il n'y a pas encore assez de travaux réalisés sur l'identité des ARNt mt de mammifères. Il a été suggéré que les ARNt sont suffisament particuliers pour que des identités simplifiées permettent d'assurer une reconnaissance spécifique entre aaRS et ARNt homologues (Kumazawa et al., 1991). Par exemple, les grandes variations observées au niveau de la longueur des boucles D et T, et/ou les modifications post-transcriptionnelles pourraient jouer des rôles clé dans l'identité. Cependant, d'autres travaux réalisés à partir d'ARNt mt de bœuf purifiés ont montré l'aminoacylation croisée des ARNt$^{Gln}$, ARNt$^{Ala}$ et ARNt$^{Asn}$ par la SerRS (clonée et purifiée) (Shimada et al., 2002). A présent que nous connaissons les gènes des aaRS mt humaines manquantes et que leur clonage, expression et purification sont en cours, une étude globale d'aminoacylations croisées pourra être menées. Ceci nous permettra d'avoir une vue d'ensemble de la spécificité d'aminoacylation au sein de la mitochondrie humaine.

| | *E. coli*[a] | mt de mammifères | Remarques |
|---|---|---|---|
| **Val** | A73<br>G3-C70, U4-A69<br>A35, C36 | **A73**<br>A3-U70, A4-U69<br>**A35, C36** | |
| **Ile** | A73<br>C4-G69<br>*L*/G34, A35, U36<br>$t^6$A37, A38<br>U12-A23, C29-G41 | R73<br>A4-U69<br>**G34, A35, U36**<br>R37, **G38**<br>Y12-R23, A29-U41 | |
| **Leu** | A73<br>U8•A14 | **A73**<br>**U**/A8•**A**14 | |
| **Met (fMet)**[§] | A73<br>(G2-C71, C3-G70)<br>U4-A69, A5-U68<br>C34, A35, U36<br>(C32, U33, A37) | R73<br>**G2-C71**, U3-A70<br>A4-U69, **A5**•G68<br>**C34, A35, U36**<br>**C32**, C33, **A37** | [§]le même ARNt est élongateur et initiateur dans la mitochondrie de mammifères |
| **Cys** | U73<br>G2-C71, C3-G70<br>G34, C35, A36<br>G15•G48, A13•A22 | Y73<br>**G2**-Y71, Y3-**G70**<br>**G34, C35, A36**<br>-[§], **A13**-U22 | [§]les nucléotides ne sont pas conservés et la structure n'est pas connue |
| **Tyr** | A73<br>G2-C71<br>U35 | **A73**<br>R2-Y71[§]<br>**U35** | [§]chez l'homme, la paire G2-C71 est retrouvée mais n'est pas élément d'identité (Bonnefond et al., 2005a) |
| **Trp** | G73<br>A1-U72, G2-C71<br>G3-C70<br>C34, C35, A36 | **G73**<br>**A1-U72**, R2-Y71<br>R3-Y70<br>U34[§], **C35, A36** | [§]variation due aux déviations du code génétique dans la mitochondrie |
| **Glu** | G1-C72, U2-A71<br>$s^4$U34, U35<br>A37<br>U11-A24<br>U13-G22•A46, Δ47 | R1-Y72, Y2-R71<br>**U34, U35**<br>R37<br>**U11-A24**<br>G13-C22•R46[§], **Δ47** | [§]l'interaction tertiaire est prédictive |
| **Gln** | G73<br>U1-A72, G2-C71<br>G3-C70<br>*Y*34, U35, G36<br>A37, U38<br>G10 | **G73**<br>**U1-A72**, R2-Y71<br>**G3-C70**<br>U34[§], **U35, G36**<br>R37, A38<br>**G10** | [§]la modification n'est pas connue |
| **Arg** | A/G73<br>C35, U/G36<br>A20 | **A73**<br>**C35, G36**<br>- | |

| | *E. coli*[a] | mt de mammifères | Remarques |
|---|---|---|---|
| **Ser**[§] | G73<br>C72, G2-C71<br>A3-U70<br>C11-G24, R4-Y69<br>grande région variable | R73<br>Y72, R2-U71<br>R3-Y70<br>R11-Y24, **R4-U69**<br>petite region variable | [§]il existe deux isoaccepteurs serine dans la mitochondrie de mammifère. Seule la sequence de l'ARNt$^{Ser(UCN)}$ a été analysée. |
| **Thr** | G1-C72, C2-G71<br>G34, G35, U36 | R1-Y72, Y2-R71<br>U34[§], **G35, U36** | [§]variation due aux deviations du code génétique |
| **Pro** | A73<br>G72<br>G35, G36<br>G15•C48 | **A73**<br>**G72**<br>**G35, G36**<br>- | |
| **Gly** | U73<br>G1-C72, C2-G71<br>G3-C70<br>C35, C36 | A73<br>R1-Y72, Y2-R71<br>Y3-R70<br>**C35**, Y36 | |
| **His** | C73<br>G-1<br>G34, U35, G36 | **C73**<br>-[§]<br>**G34, U35, G36** | [§]la présence de G en position -1 n'est pas connue |
| **Asp** | G73<br>G2:C71<br>G34, U35, C36<br>C38<br>G10 | N73<br>R2-U71<br>**G34, Y35[§], C36**<br>A38<br>**G10** | [§]position partiellement éditée chez l'ARNt$^{Asp}$ mt de marsupiaux pour donner U35 ou C35 (Mörl et al., 1995) |
| **Lys** | A73<br>U34, U35, U36<br>(*mnm*$^5$*s*$^2$-U)34 | **A73**<br>**U34, U35, U36**<br>[§] | [§]l'ARNt$^{Lys}$ mt humain possède un dérivé de la taurineméthyl en 34 position (Yasukawa et al., 2000b) |
| **Asn** | G73<br>G34, U35, U36 | **G73**<br>**G34, U35, U36** | |
| **Phe** | A73<br>G34, A35, A36<br>G27-C43, G28-C42<br>U20<br>G44, U45, U59, U60 | **A73**<br>**G34, A35, A36**<br>R27-U43, **G28**•Y42<br>-<br>A44, R45, - | |
| **Ala** | A73<br>G2-C71, G3•U70<br>G4-C69<br>G20 | R73<br>R2:Y71, R3:Y70<br>R4-Y69<br>- | |

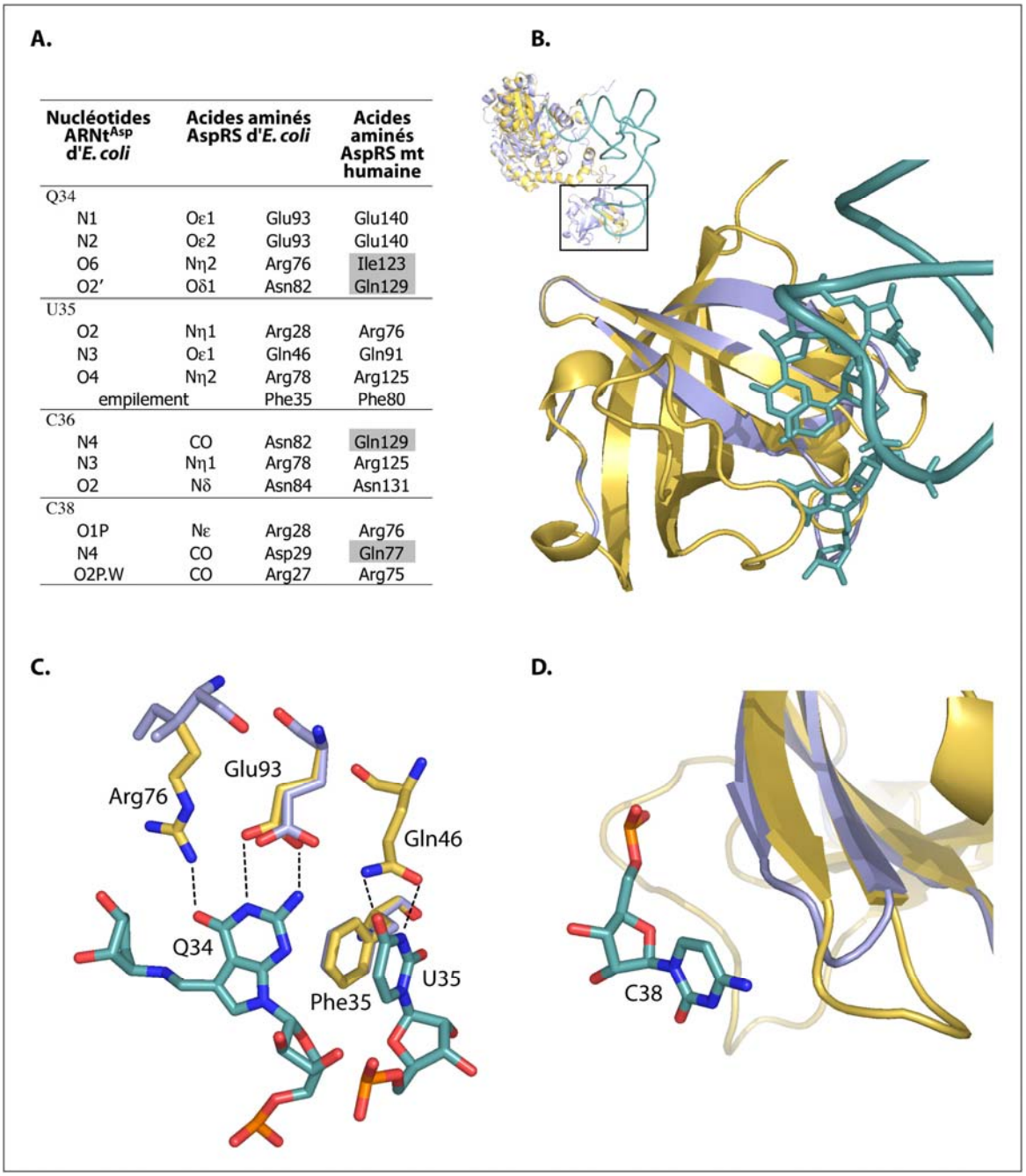
**Tableau 6. Recherche des éléments d'identité d'*E. coli* dans les séquences consensus des gènes d'ARNt mt de mammifères (136 pour l'ARNt$^{Asp}$ et 71 pour tous les autres).** Les séquences mt analysées ont été compilées par Joern Pütz. Les éléments d'identité d'*E. coli* ([a]revue par (Giegé et al., 1998)) conservés dans les ARNt mt de mammifères sont indiqués en vert. Les éléments strictement conservés dans tous les ARNt mt de mammifères d'une spécificité donnée sont soulignés. Les modifications post-transcriptionnelles sont en italique. Y et R signifient pyrimidine et purine, respectivement. Les systèmes ont été répartis selon leur appartenance aux classes et sous-classes d'aaRS (classe I : 1$^{er}$ tableau ; classe II : 2$^{nd}$ tableau).

## 3. Adapatation de l'AspRS mt humaine à son ARNt$^{Asp}$

Alors que le résidu G73 est nécessaire à la reconnaissance de l'ARNt$^{Asp}$ d'*E. coli*, l'AspRS mt tolère n'importe lequel des 4 nucléotides possibles à cette position. De type bactérienne, (~40 % d'identité de séquence avec l'AspRS d'*E. coli*), l'AspRS mt humaine est codée par le génome nucléaire (environnement stable), et s'est adaptée de manière ciblée à la dégénérescence de son ARNt (codé par le génome mt). Alors que l'AspRS bactérienne interagit spécifiquement avec l'élément G73 via deux acides aminés (Asp220 et Arg222), l'enzyme mt n'a conservé que l'arginine et possède une glycine à la place de l'aspartate. La substitution expérimentale de cette glycine par un aspartate restaure le pouvoir discriminant de l'enzyme vis-à-vis du résidu G73. La construction d'un modèle 3D de l'enzyme par Claude Sauter d'après la structure cristallographique du complexe d'*E. coli* (Eiler et coll. 1999 EMBO J.) et l'analyse d'alignements multiples de séquences d'AspRS de différents organismes ont été déterminants dans l'identification du résidu impliqué.

Le modèle 3D de l'AspRS mt et la comparaison des séquences ont également permis de comprendre ce qui se passe au niveau de l'interaction avec la boucle de l'anticodon de l'ARNt, et en particulier de savoir : (i) si les éléments d'identité de l'anticodon sont reconnus de la même façon que dans le complexe d'*E. coli*, et (ii) pourquoi l'élément en position 38 n'est plus un élément d'identité. Les acides aminés en contact direct avec le triplet anticodon du complexe d'E. coli sont globalement conservés dans l'AspRS

Maintenant ce qu'on aimerait savoir c'est ce qui se passe au niveau de l'interaction de l'AspRS avec la boucle de l'anticodon. Les éléments d'identité sont-ils reconnus de la même façon ? Que se passe-t-il au niveau de la position 38 qui n'est plus un élément d'identité dans l'ARNt$^{Asp}$ mt ? Pour ce faire, nous avons d'abord recherché si les résidus acides aminés en contact avec la boucle de l'anticodon dans le complexe d'*E. coli* sont retrouvés dans la séquence de l'AspRS mt. Grâce au modèle 3D de l'AspRS mt, nous avons pu placer ces résidus dans la structure et voir si les interactions avec la boucle de l'anticodon de l'ARNt$^{Asp}$ d'*E. coli* sont maintenues (Figure 24).




| Nucléotides ARNt$^{Asp}$ d'*E. coli* | Acides aminés AspRS d'*E. coli* | | Acides aminés AspRS mt humaine |
|---|---|---|---|
| Q34 | | | |
| N1 | Oε1 | Glu93 | Glu140 |
| N2 | Oε2 | Glu93 | Glu140 |
| O6 | Nη2 | Arg76 | Ile123 |
| O2′ | Oδ1 | Asn82 | Gln129 |
| U35 | | | |
| O2 | Nη1 | Arg28 | Arg76 |
| N3 | Oε1 | Gln46 | Gln91 |
| O4 | Nη2 | Arg78 | Arg125 |
| empilement | | Phe35 | Phe80 |
| C36 | | | |
| N4 | CO | Asn82 | Gln129 |
| N3 | Nη1 | Arg78 | Arg125 |
| O2 | Nδ | Asn84 | Asn131 |
| C38 | | | |
| O1P | Nε | Arg28 | Arg76 |
| N4 | CO | Asp29 | Gln77 |
| O2P.W | CO | Arg27 | Arg75 |

**Figure 24. Adaptation de l'AspRS mitochondriale humaine aux nucléotides de l'anticodon. A.** Récapitulatif des contacts entre nucléotides, éléments d'identité de l'ARNt$^{Asp}$ d'*E. coli*, et acides aminés de l'AspRS homologue. Les acides aminés correspondants dans l'AspRS mt humaine sont indiqués à droite, et ceux non conservés sont grisés. **B.** Agrandissement du domaine de liaison à l'anticodon. Le modèle 3D de l'AspRS mt (en bleu) et la structure cristallographique de l'AspRS d'*E. coli* (en jaune, (Eiler et al., 1999)) sont superposés. Les nucléotides 34, 35, 36 et 38 sont visibles sur la structure de l'ARNt$^{Asp}$ d'*E. coli* (en vert). **C.** Zone d'interaction avec les nucléotides 34 et 35. Les acides aminés de l'AspRS d'*E. coli* directement en contact sont ceux en jaune et indiqués, ceux correspondants dans l'AspRS mt sont en bleu. A noter, la Gln91 (mt) est parfaitement superposable à la Gln46 (*E. coli*), et l'Arg76 (*E. coli*) est remplacée par l'Ile123 (mt) d'orientation complètement différente. **D.** Environnement du résidu C38 de l'ARNt$^{Asp}$ d'*E. coli*. A noter, la variation de taille des boucles des AspRS mt (en bleu) et d'*E. coli* (en jaune).

Globalement, les acides aminés de l'AspRS d'*E. coli* en contact avec les éléments du triplet anticodon sont retrouvés dans la séquence de l'AspRS mt humaine. Notons la conservation du résidu Phe empilé sur le résidu U35 de l'anticodon, interaction conservée dans toutes les structures de complexes AspRS/ARNt$^{Asp}$ connues. Cependant, un des trois contacts spécifiques avec la base en position 34 est manquant. Le résidu Arg76 de l'AspRS d'*E. coli* impliqué est remplacé par l'Ile123 dans l'AspRS mt humaine. Ceci pourrait expliquer l'effet modéré sur l'aspartylation d'ARNt$^{Asp}$ mt mutés en position 34. Rappelons que la modification queosine du résidu G34 n'établie pas de contact dans le complexe d'*E. coli*. En ce qui concerne de la position 38 (occupée par un A dans l'ARNt$^{Asp}$ mt au lieu d'un C dans celui d'*E. coli*), sa non implication dans l'identité pourrait s'expliquer par la taille réduite d'une boucle de l'AspRS mt par rapport à celle d'*E. coli*, dans son environnement proche. La question est de savoir quelle est l'incidence de cette différence sur la reconnaissance de l'ARNt$^{Asp}$ mt.

# C. Unicité de reconnaissance de l'ARNt$^{Asp}$ mt humain

Il est connu depuis longtemps qu'il existe une barrière dans les possibilités d'aminoacylation croisées impliquant des systèmes mt. En effet, alors que les aaRS mt sont capables d'aminoacyler des ARNt de différentes origines, les ARNt mt ne sont pas reconnus par des synthétases autres que mitochondriales (Kumazawa et al., 1991). Nous avons exploré les signaux responsables de ceci au sein de l'ARNt$^{Asp}$ mt humain.

## *1. Aminoacylations croisées ou barrières d'espèces*

Les jeux d'identité sont globalement conservés au cours de l'évolution, ce qui autorise (au moins dans des expériences *in vitro*) l'aminoacylation croisée d'ARNt par des aaRS spécifiques mais d'organismes différents. Le système le plus représentatif de ce phénomène est celui de l'alanylation. En effet, l'aminoacylation spécifique de l'ARNt$^{Ala}$ dépend essentiellement de la paire de bases G3•U70 du bras accepteur, élément présent dans une large distribution d'ARNt de la bactérie à celui cytoplasmique humain (Chihade et al., 1998). Ainsi, l'expression de l'AlaRS cytoplasmique humaine sauve une souche de levure délétée de l'AlaRS endogène, et l'AlaRS d'*E. coli* charge des substrats minimalistes dérivés des ARNt$^{Ala}$ de *S. cerevisiae*, *Bombyx mori* et cytoplasmique humain (Hou & Schimmel, 1989 ; Ripmaster et al., 1995). Cependant, il existe des systèmes pour lesquels l'aminoacylation croisée ne se fait pas et ainsi des barrières d'espèces ont été définies. Par exemple, la glycylation de l'ARNt$^{Gly}$ bactérien dépend de la base discriminatrice U73, alors que celle des ARNt$^{Gly}$ eucaryotiques et d'archaea est définie par un A à cette position. Cette différence de séquence explique à elle seule l'absence d'aminoacylation croisée. L'échange de cette base inverse la spécificité de l'ARNt et permet ainsi l'aminoacylation hétérologue (Shiba et al., 1994a ; Hipps et al., 1995). Notons qu'en accord avec la barrière d'espèces, on observe deux groupes structuralement distincts de GlyRS. En particulier, les enzymes bactériennes possèdent une organisation oligomérique de type $\alpha 2\beta 2$ alors que les GlyRS eucaryotiques et d'archaea sont

de type α2 (Shiba et al., 1997). Dans ce cas, on parle de barrière d'espèces inter-phylum (ou entre les différents règnes du vivant)

Dans le cas des systèmes mt de mammifères, nous sommes confrontés à un problème similaire de barrière d'espèces. Une étude réalisée par (Kumazawa et al., 1991) a montré l'unilatéralité de l'aminoacylation entre les systèmes mt bovins et bactériens (*E. coli* et *T. thermophilus*) pour 5 spécificités différentes (Phe, Thr, Arg, Lys et Ser). En effet, alors que les synthétases mt sont capables d'aminoacyler à la fois les ARNt mt et bactériens, les aaRS bactériennes n'aminoacylent pas les ARNt mt. La particularité ici réside dans le fait que certaines aaRS mt sont de type bactérien, confirmant la théorie de l'endosymbiose et le transfert de gènes de la bactérie primitive vers le noyau. Ainsi on a ce qu'on pourrait définir une barrière d'espèce « intra-phylum », c'est à dire entre systèmes biologiques reliés par l'évolution (bactérie et pseudo-bactérie).

## 2. Non reconnaissance d'un ARNt dégénéré ?

Une hypothèse évidente quant à la barrière d'espèce entre les systèmes mt de mammifères et bactériens, viendrait de la grande différence de séquence et de structure entre les ARNt. En effet, les gènes des ARNt mt de mammifères sont les cibles d'un taux de mutation très élevé (100 fois supérieur à celui de leurs homologues nucléaires (Saccone et al., 1999)). Par conséquent, ils possèdent des séquences et structures dégénérées par rapport aux ARNt canoniques et notamment bactériens (cf Introduction C.2.1.). Cette hypothèse semble d'autant plus probable que les ARNt mt de levure, qui sont restés classiques, peuvent être reconnus par les aaRS d'*E. coli*. Pour exemple, la TyrRS d'*E. coli* complémente une souche de levure dont le gène de la TyrRS mt est soit muté, soit délété, et ce malgré la différence significative entre les séquences de $ARNt^{Tyr}$ bactérien et mt de levure (Edwards & Schimmel, 1987).

Nous avons montré précédemment que l'AspRS mt humaine aminoacyle les $ARNt^{Asp}$ d'*E. coli*, de *T. hermophilus* et de *S. cerevisiae* alors que l'$ARNt^{Asp}$ mt n'est pas un substrat de l'enzyme bactérienne (voir l'article n°4). L'ARNt mt n'est pas non plus reconnu par les AspRS de levure et de *T. thermophilus* (AspRS 1 et 2). Ceci pourrait s'expliquer par (i) l'absence d'éléments d'identité nécessaires à ces AspRS (c'est à dire les éléments majeur G73

et mineur C38, et dans le cas de l'enzyme d'*E. coli* la paire de base G2-C71), (ii) les particularités structurales de l'ARNt$^{Asp}$ mt, et/ou (iii) la présence d'antidéterminants. Afin de statuer entre ces différentes hypothèses, notre objectif a été de convertir l'ARNt mt en un substrat de l'enzyme bactérienne. Pour ce faire, nous avons (i) introduit dans la séquence de l'ARNt$^{Asp}$ mt humain des éléments d'identité aspartate d'*E. coli* manquants, soit de manière individuelle, soit dans leur ensemble, (ii) introduit les boucles D et T d'un ARNt à structure 3D définie (celui de *S. pombe*), (iii) et/ou ciblé et muté des éléments potentiellement antidéterminants vis-à-vis de l'AspRS d'*E. coli*. Les différents mutants (Figure 25) ont été testés en présence de l'AspRS d'*E. coli* d'une part et de l'AspRS mt humaine d'autre part. Les tests ont été effectués dans les mêmes conditions réactionnelles que celles utilisées pour l'AspRS mt. Bien que ces conditions ne soient pas optimales pour l'AspRS d'*E. coli* (Hasegawa et al., 1989 ; Nameki et al., 1992), elles nous permettront de réaliser des tests comparatifs pour les deux enzymes (mt et bactérienne).

## 3. Nécessité des éléments d'identité aspartate d'E. coli

Chacun des éléments d'identité de l'ARNt$^{Asp}$ d'*E. coli* absent dans l'ARNt$^{Asp}$ mt humain (le majeur G73, et les mineurs C38 et G2-C71) a été introduit de façon indépendante dans l'ARNt mt. Les variants sont les ARNt$^{Asp}$G73, ARNt$^{Asp}$C38 et ARNt$^{Asp}$G2-C71. Leurs taux de charge ont été testés en présence de l'AspRS mt et d'*E. coli* (Figure 26). Alors qu'ils sont tous les trois aminoacylés par l'AspRS mt avec des taux de charge équivalents à celui du substrat sauvage (taux de charge de 45 à 70 %), aucun ne devient un meilleur substrat pour l'enzyme bactérienne (taux de charge inférieur à 1%).

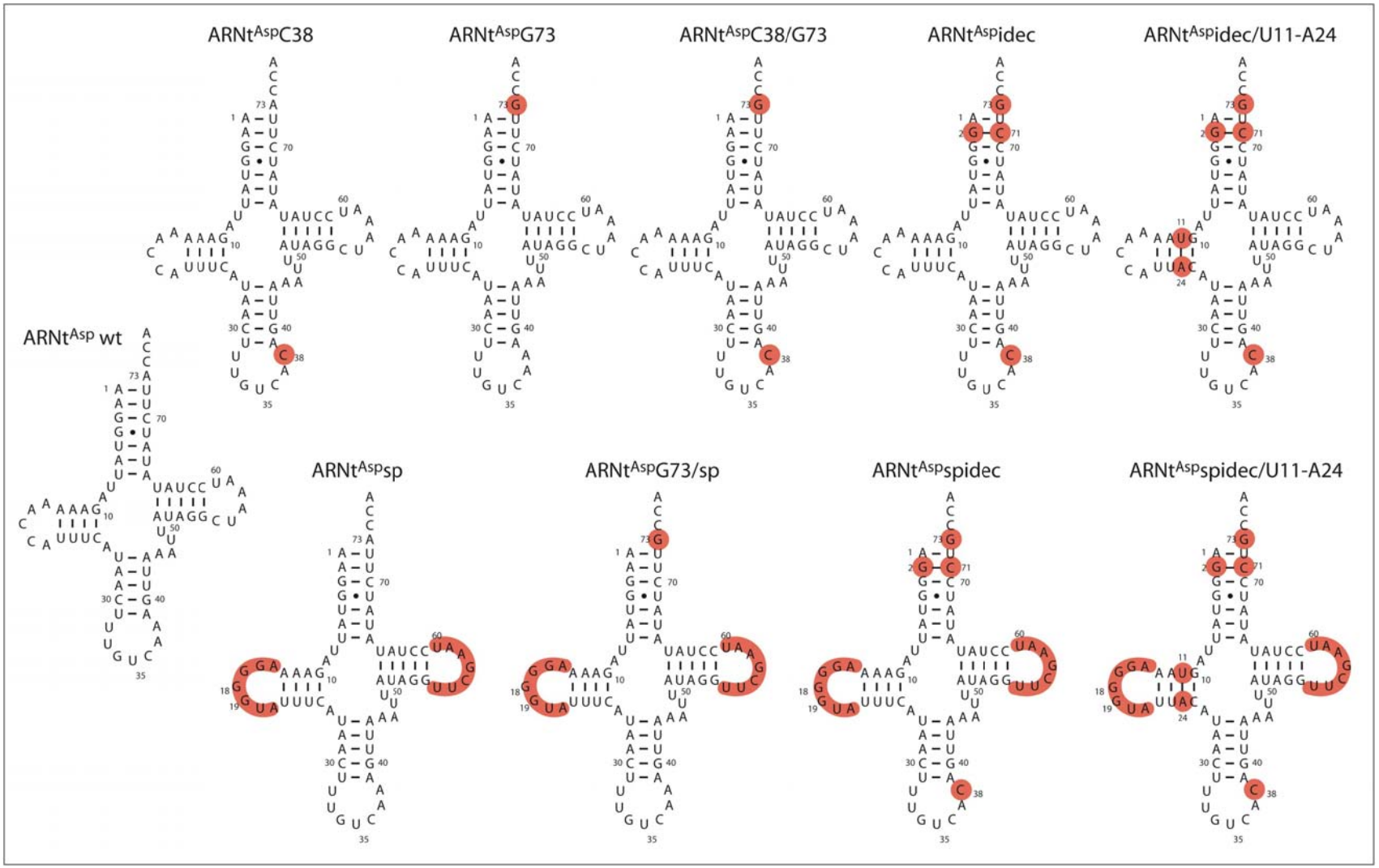



**Figure 25. Variants de l'ARNt$^{Asp}$ mitochondrial humain.** Les mutations introduites sont indiquées en rouge. L'appellation des molécules obtenues est précisée.

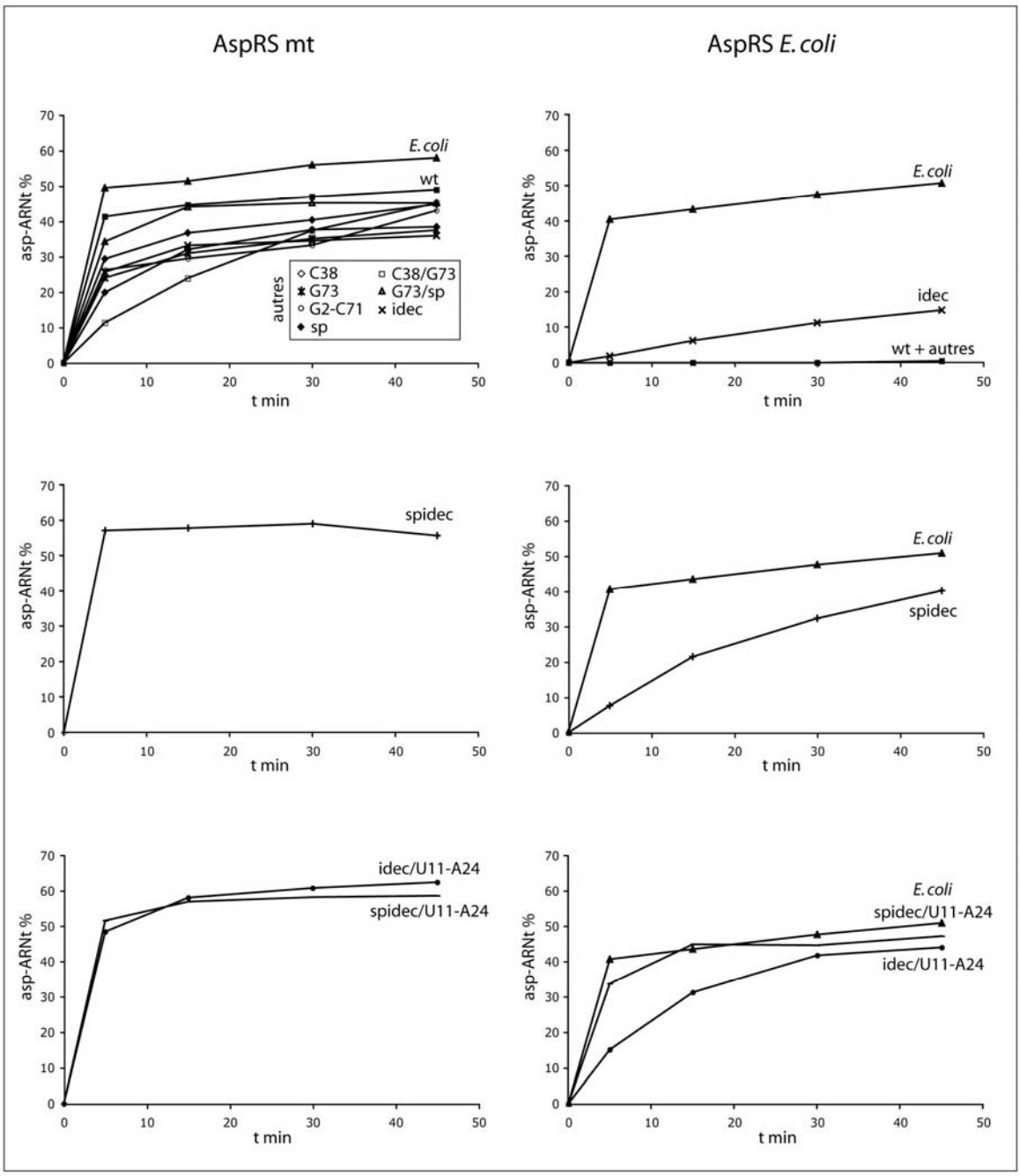



**Figure 26. Plateaux d'aspartylation** des transcrits *in vitro* des ARNt$^{Asp}$ mt sauvage (wt) et dérivés, et d'*E. coli* en présence de 500 nM d'AspRS mt (graphiques de gauche) ou 200 nM d'AspRS d'*E. coli* (graphiques de droite). La concentration en transcrits est de 0,4 µM. Les variants de l'ARNt$^{Asp}$ mt sont indiqués par la mutation introduite (voir Figure 25).

Nous avons donc combiné les mutations au sein de l'ARNt$^{Asp}$ mt. Les mutants multiples obtenus sont les ARNt$^{Asp}$C38/G73 et ARNt$^{Asp}$idec (pour la transplantation des éléments C38, G73 et G2-C71, équivalant au jeu d'éléments d'identité complet d'*E. coli*). Les deux mutants restent de bons substrats pour l'AspRS mt avec des taux de charges de 45 % et 36 %, respectivement. Alors que l'ARNt$^{Asp}$C38/G73 n'est pas reconnu par l'AspRS d'*E. coli*

(taux de charge <1%), l'introduction de l'ensemble des éléments d'identité d'*E. coli* améliore clairement les capacités d'aminoacylation de l'ARNt$^{Asp}$ mt (le mutant ARNt$^{Asp}$idec est chargé à 15 %, Figure 26). Les paramètres cinétiques $k_{cat}$ et $K_M$ associés à l'ARNt$^{Asp}$idec et à l'ARNt$^{Asp}$ d'*E. coli* (qui sert de référence) ont alors été déterminés (Tableau 7). L'ARNt$^{Asp}$ d'*E. coli* et le variant ARNt$^{Asp}$idec possèdent des efficacités catalytiques ($k_{cat}/K_M$) de 3700 et 0,5 $10^{-3}$ $sec^{-1}.\mu M^{-1}$ respectivement.. L'ARNt$^{Asp}$ mt de séquence sauvage étant un très mauvais substrat de l'AspRS *d'E. coli*, son efficacité catalytique n'a pu être mesurée et par conséquent le gain d'activité du variant n'a pu être calculé. Toutefois, l'ARNt$^{Asp}$idec reste un substrat 7 400 fois moins efficace pour l'enzyme bactérienne que l'ARNt$^{Asp}$ d'*E. coli*, du fait à la fois d'un effet $k_{cat}$ et $K_M$.

Ainsi, l'introduction indépendante des éléments d'identité manquants ne suffit pas à convertir l'ARNt$^{Asp}$ mt en un substrat de l'AspRS d'*E. coli*. Seule l'introduction conjointe de tous ces éléments permet l'obtention d'un variant (ARNt$^{Asp}$idec) aspartylable par l'AspRS hétérologue mais avec une efficacité qui reste significativement réduite par rapport à celle du substrat homologue. Ceci témoigne de la coopérativité des éléments d'identité dans le contexte de l'ARNt mt. Etant donné que le jeu d'identité d'*E. coli* est admis comme étant complet (sa transplantation permet l'aspartylation efficace des ARNt$^{Val}$ et ARNt$^{Asn}$, (Nameki et al., 1992)), nous pensons qu'il ne manque plus d'éléments déterminants. Plusieurs hypothèses restent possibles : (i) soit l'architecture globale de l'ARNt est défavorable à la présentation et l'adaptation optimale des éléments d'identité, (ii) soit il y a des éléments antidéterminants qui empêchent la reconnaissance par l'AspRS d'*E. coli*, (iii) soit une combinaison des deux.

| **ARNt** | **$k_{cat}$ ($10^{-3}$ $sec^{-1}$)** | **$K_M$ (µM)** | **$k_{cat}/K_M$ ($10^{-3}$ $sec^{-1}.\mu M^{-1}$)** | **P**[a] |
|---|---|---|---|---|
| *E.coli* AspRS | | | | |
| ARNt$^{Asp}$ d'*E. coli* | 2590 | 0,7 | 3700 | 1 |
| ARNt$^{Asp}$ mt wt | nd | nd | nd | nd |
| ARNt$^{Asp}$idec | 39 | 69,3 | 0,5 | 7400 |
| ARNt$^{Asp}$spidec | 34 | 17,5 | 2,0 | 1850 |
| ARNt$^{Asp}$idec/U11-A24 | 43 | 7,5 | 5,7 | 650 |
| ARNt$^{Asp}$spidec/U11-A24 | 23 | 1,6 | 14,4 | 260 |

**Tableau 7.** Paramètres cinétiques de l'aspartylation des transcrits *in vitro* de l'ARNt$^{Asp}$ d'*E. coli*, mt sauvage (wt), et dérivés, en présence de l'AspRS d'*E. coli*. [a] La perte d'efficacité **P** est définie comme le rapport $k_{cat}/K_M$ (ARNt$^{Asp}$ d'*E. coli*) / $k_{cat}/K_M$ (ARNt$^{Asp}$ mt ou variant). nd pour non déterminable. L'erreur expérimentale est estimée à 30 %.

## 4. Implication de l'architecture centrale de l'ARNt

Nous avons testé l'influence de la structure de l'ARNt$^{Asp}$ mt sur l'aspartylation par l'AspRS d'*E. coli*. Pour ce faire les boucles D et T d'un ARNt$^{Asp}$ canonique (ici de *S. pombe*) ont été introduites dans l'ARNt mt sauvage ou dans des variants. Les mutants obtenus (Figure 25) sont les ARNt$^{Asp}$sp, ARNt$^{Asp}$G73/sp et ARNt$^{Asp}$spidec (où sp signifie « boucles D et T de *S. pombe* »). Ces boucles possèdent les signaux « classiques » impliqués dans les interactions tertiaires à longue distance et conduisent au repliement 3D sous forme de L de l'ARNt. Nous admettons que l'ARNt$^{Asp}$sp et ses variants se replient sous la forme en L et possèdent les interactions entre les résidus G18G19 et U55C56 qui verrouillent l'ensemble. Ceci conduit à remplacer les éléments structuraux les plus « bizarres » de l'ARNt mt par des éléments plus conventionnels.

Les taux de charge en présence de l'AspRS mt et d'*E. coli* ont été déterminés (Figure 26). Les 3 ARNt restent aspartylés par l'enzyme mt avec un taux de charge similaire à celui de l'ARNt mt sauvage (de 45 à 70 %). Le variant ARNt$^{Asp}$spidec devient un substrat de l'AspRS d'*E. coli*, atteignant un taux de charge de 40 % après 45 min de réaction. Les deux autres mutants restent inactifs (taux de charge <1%). Les paramètres cinétiques associés au variant ARNt$^{Asp}$spidec sont reportés dans le Tableau 7. La perte d'efficacité catalytique, calculée par rapport à la référence ARNt$^{Asp}$ d'*E. coli*, est de 1 840. Ainsi, l'introduction de boucles « canoniques » améliorent l'activité du mutant ARNt$^{Asp}$idec de 4 fois. Ce gain est essentiellement dû à un effet sur le $K_M$, les ARNt$^{Asp}$idec et ARNt$^{Asp}$spidec possédant des $k_{cat}$ similaires (39 et 34 $10^{-3}$ $sec^{-1}$, respectivement).

En conclusion, l'introduction de boucles D et T canoniques améliore les capacités d'aminoacylation croisée de l'ARNt$^{Asp}$ mt déjà transplanté de l'ensemble des éléments d'identité aspartate d'*E. coli*. Cependant, ce contexte ne suffit pas en tant que tel puisque lorsqu'il est associé à l'élément aspartate majeur G73 (qui suffit pourtant à l'aspartylation de minihélices jusqu'à 30 %, (Frugier et al., 1994a)), l'ARNt n'est pas un substrat de l'AspRS d'*E. coli*. Ceci montre que le contexte structural de l'ARNt mt ne joue pas un rôle clé dans les signaux d'unicité d'aminoacylation.

## 5. Recherche d'éléments antidéterminants

Malgré la présence de tous les éléments d'identité aspartate d'*E. coli* et malgré l'introduction de boucles D et T canoniques, l'ARNt$^{Asp}$ mt n'est pas un substrat optimal pour l'AspRS d'*E. coli* (il reste 1850 fois moins efficacement aspartylé). Nous avons alors exploré la possibilité de présence d'antidéterminants dans l'ARNt mt empêchant sa reconnaissance optimale par l'AspRS d'*E. coli*. Un antidéterminant peut être un nucléotide, une paire de bases ou un élément structural, et peut agir directement en empêchant l'interaction avec l'aaRS ou indirectement en altérant la structure de l'ARNt et/ou la transmission du signal au site catalytique. De tels éléments ont été recherchés au sein du variant ARNt$^{Asp}$spidec en comparant sa séquence à celle de tous les ARNt efficacement aspartylés par l'AspRS d'*E.coli* : l'ARNt$^{Asp}$ homologue, les ARNt$^{Asp}$ cytoplasmiques de *S. cerevisiae* (transplanté des deux paires de bases G1-C72 et G2-C71 (Moulinier et al., 2001)) et humain (article n°4), les ARNt$^{Val/Asp}$ et ARNt$^{Asn/Asp}$ d'*E. coli* (transplantés du jeu d'identité aspartate) (Nameki et al., 1992).

L'alignement des séquences est donné dans la Figure 27. Il met en évidence les nucléotides strictement conservés (en orange) ou semi-conservés (Y ou R, en jaune) et fait ressortir les différences avec l'ARNt$^{Asp}$spidec par rapport aux 5 autres séquences. Les variations de séquence observées dans l'ARNt$^{Asp}$spidec mt concernent des positions semi-conservées (Y ou R) dans les autres séquences d'ARNt aspartylables. Ces positions sont la paire de bases 11-24, les nucléotides 16 de la boucle D, 13 et 22 du bras D et 46 de la boucle variable. D'après la structure cristallographique du complexe ARNt$^{Asp}$/AspRS d'*E. coli* (Eiler et al., 1999), la boucle D se situe sur la face de l'ARNt qui n'interagit pas avec l'enzyme. Ainsi l'élément à la position 16 ne produit probablement pas de gênes stériques empêchant l'interaction avec l'enzyme. Les positions 13, 22 et 46 dans les ARNt canoniques sont impliquées dans une interaction triple. La co-variation de séquence à ces positions dans l'ARNt$^{Asp}$ mt permettrait en fait de maintenir l'interaction triple, du moins en théorie. Ainsi, nous considérons la position 11-24 comme un bon candidat dans le rôle d'antidéterminant et ce pour deux raisons : (i) il s'agit d'une paire de bases semi-conservée dans les ARNt canoniques sous la forme Y11-R24 (Dirheimer et al., 1995) et qui devient A11-U24 (R11-Y24) dans l'ARNt$^{Asp}$ mt humain, et (ii) l'élément U11 de l'ARNt$^{Asp}$ d'*E. coli* est en contact direct avec l'AspRS (Eiler et al., 1999). Afin de vérifier le rôle de cette paire de bases dans la non reconnaissance de l'ARNt mt par l'AspRS d'*E. coli*, la paire de bases A11-U24 a été

remplacée par celle naturellement présente dans l'ARNt$^{Asp}$ d'*E. coli*, c'est à dire U11-A24, dans les variants ARNt$^{Asp}$idec et ARNt$^{Asp}$spidec. Les deux nouveaux ARNt obtenus sont les ARNt$^{Asp}$idec/U11-A24 et ARNt$^{Asp}$spidec/U11-A24 (Figure 25). Ils ont été testés en présence de l'AspRS mt et de l'AspRS d'*E. coli* (Figure 26, Tableau 7).

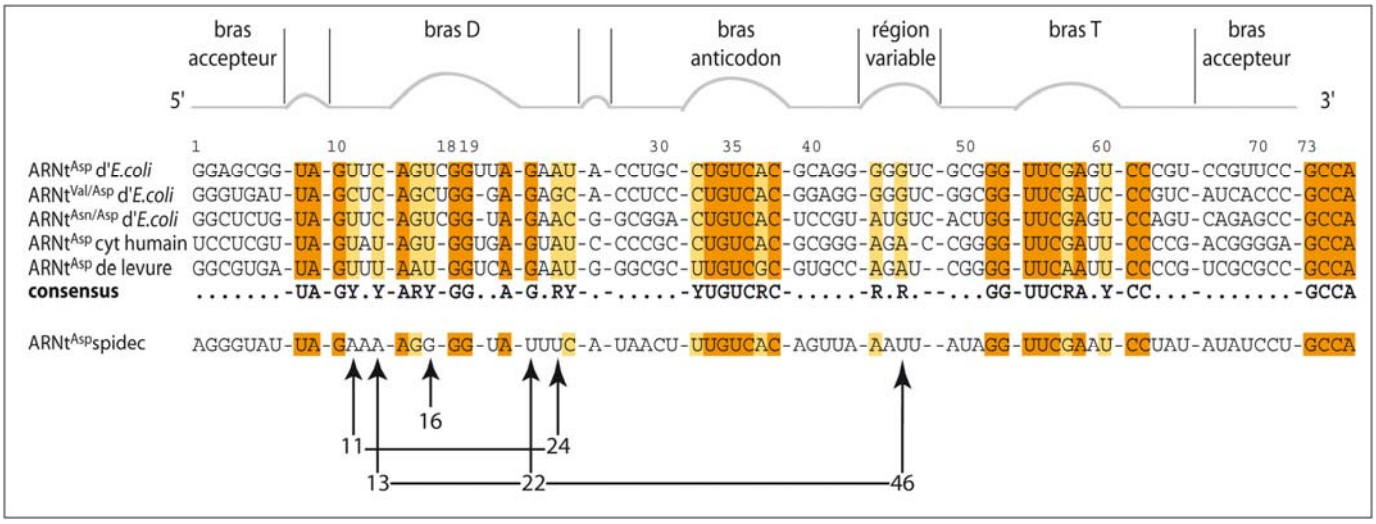



**Figure 27. Consensus des séquences d'ARNt efficacement aspartylés par l'AspRS d'*E. coli* et comparaison à la séquence du variant ARNt$^{Asp}$/spidec.** Les éléments strictement conservés sont en orange et semi-conservés (Y ou R) en jaune. Le consensus de ces séquences est indiqué en gras. Les éléments de ce consensus, non conservés dans le variant de l'ARNt mt sont mis en évidence par des flèches (reliées en cas d'interaction de type paire de bases ou tertiaire). Les domaines de structure secondaire sont indiqués au-dessus des séquences primaires.

La présence de la paire de bases U11-A24 améliore significativement les activités des deux variants en présence de l'AspRS d'*E. coli*. L'ARNt$^{Asp}$idec/U11-A24 est aspartylé à 44 % et ne possède plus qu'une perte d'efficacité de 650 par rapport à l'ARNt$^{Asp}$ d'*E. coli*. L'ARNt$^{Asp}$spidec/U11-A24 est encore un meilleur substrat pour l'enzyme bactérienne, avec un comportement équivalent à celui de l'ARNt$^{Asp}$ homologue en plateau de charge (taux de charge de 45 %) et une perte d'efficacité réduite à 260 fois. Ainsi, l'introduction de la paire de bases U11-A24 apporte un gain d'efficacité de ~11 fois pour l'ARNt$^{Asp}$idec et ~7 fois pour l'ARNt$^{Asp}$spidec. Ces gains sont essentiellement dus à un effet sur le $K_M$ puisque les $k_{cat}$ associés à ces deux nouveaux mutants restent similaires à ceux associés aux ARNt$^{Asp}$idec et ARNt$^{Asp}$spidec. La différence d'activité entre les deux molécules est de ~2,5 fois. Notons que les variants ARNt$^{Asp}$idec/U11-A24 et ARNt$^{Asp}$spidec/U11-A24 restent des substrats de l'AspRS mt humaine avec des taux de charge similaires à celui du substrat sauvage (de 45 à 70 %).

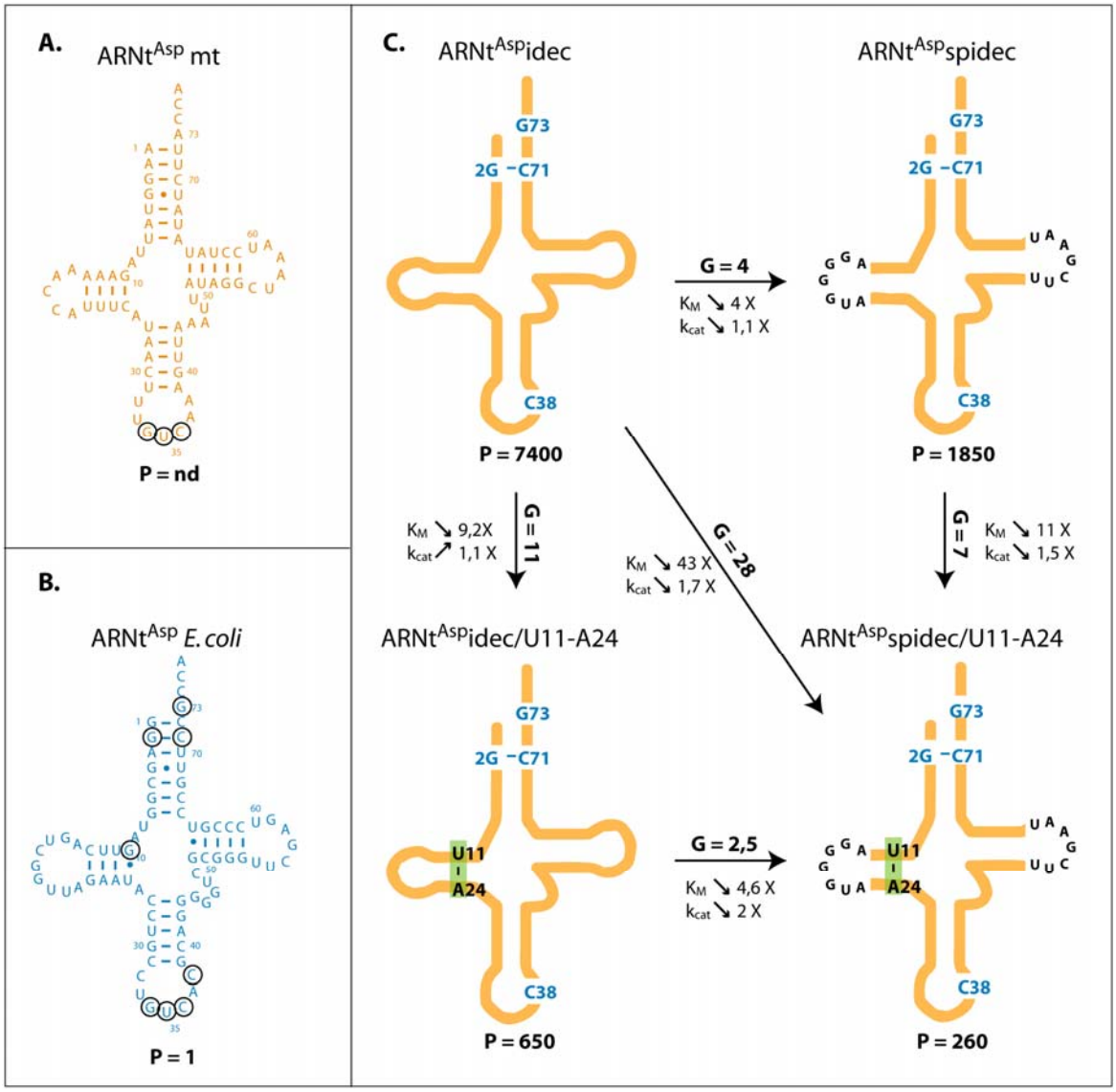



**Figure 28. Conversion de l'ARNt$^{Asp}$ mitochondrial humain en un substrat de l'AspRS d'*E. coli*.** Séquences des ARNt$^{Asp}$ mt (en orange, **A.**) et d'*E. coli* (en bleu, **B.**) sur lesquels les éléments d'identité sont encerclés. **C.** Relation entre les 4 variants aspartylables par l'enzyme d'*E. coli*. Les nucléotides restaurant l'identité d'*E. coli* (bleu), la paire de bases U11-A24 (vert), ou la séquence des boucles D et T de *S. pombe* (noir) sont indiqués sur des représentations schématiques des ARNt. Les relations fonctionnelles entre deux molécules (Gain G) résultent d'augmentation ou de diminution des valeurs de $k_{cat}$ et de $K_M$ relatifs. Les valeurs P, associées à toutes les molécules, indiquent les pertes d'efficacité catalytique des variants comparées à celle de l'ARNt$^{Asp}$ d'*E. coli*, en présence de l'AspRS d'*E. coli*. nd pour non déterminable.

## 6. Effet additif des mutations

La Figure 28 récapitule les résultats obtenus et indique que (i) la contribution des boucles (architecture) est de 2,5 à 4 fois (les horizontales), (ii) la contribution de la levée de l'antidéterminant A11-U24 est de 7 à 11 fois (les verticales) (notons qu'elle est de plus grande importance), et (iii) les deux contributions ont un effet additif. En ce qui concerne cette dernière notion, la relation entre des mutations, définie par (Pütz et al., 1993), est régie par l'équation suivante : $G_{mm} = \sum G_{ms} \times R$ (où G est le gain d'efficacité catalytique, mm signifie mutant multiple et ms mutant simple, et où R est le facteur de couplage). L'effet de plusieurs mutations dans un même variant peut être additif (R=1), coopératif (R>1) ou anti-coopératif (R<1). Dans notre cas, les gains des mutants simples sont 11 (levée de l'antidéterminant) et 4 (introduction des boucles classiques), et le gain du mutant multiple est de 28 fois. Le facteur de couplage est ainsi égal à -0,6, ce qui reste proche de 1 et traduit donc un effet additif.

## 7. Conclusion et perspectives

Les aaRS mt de mammifères sont capables de reconnaître des ARNt de différents organismes, alors que les ARNt mt ne sont pas reconnus par des synthétases autres. Une hypothèse émise quant à cette barrière d'espèce intra-phylum (bactérie et pseudo-bactérie), viendrait de la grande différence de séquence et de structure entre les ARNt, les uns étant très structurés et les autres moins structurés et plus flexibles. L'élucidation partielle de l'unilatéralité d'aminoacylation pour le système aspartate mt humain a permis d'entrevoir un mode d'antidétermination au sein de l'ARNt$^{Asp}$ mt. Outre l'absence des éléments d'identité, nous avons montré que c'est une paire de bases dans le bras D (A11-U24) qui empêche l'aminoacylation de l'ARNt$^{Asp}$ mt par l'AspRS d'*E. coli*, et ce à un degré plus important que la structure centrale de l'ARNt (boucles D et T non canoniques). En effet, l'inversion de la paire de bases 11-24 apporte un gain d'aminoacylation de l'ARNt mt plus important (G=11) que l'introduction d'une structure canonique (G=4). Ceci montre que le contexte structural de l'ARNt mt ne joue pas le rôle clé dans les signaux responsables de l'unilatéralité d'aminoacylation que l'on pensait. La paire de base A11-U24 agirait en tant

qu'antidéterminant au sein de l'ARNt$^{Asp}$ mt vis-à-vis de l'AspRS d'*E. coli*, en produisant une gêne stérique (sa mutation a un effet sur le $K_M$). La possibilité que la paire de bases en position 11-24 soit un élément d'identité supplémentaire (encore non identifié à ce jour) pour l'AspRS d'*E. coli* est peu probable puisque cette position n'est que semi-conservée au sein des substrats aspartylables (Figure 27). Par ailleurs, nous avons montré que l'introduction de la paire de bases U11-A24 améliorait l'activité de l'ARNt$^{Asp}$ mt en présence de l'AspRS de levure (résultat non illustré) pour laquelle cette position n'est pas un élément d'identité (Puglisi et al., 1993). Ainsi, pour l'AspRS de levure, la paire A11-U24 constituerait également un antidéterminant.

Toutefois, le variant de l'ARNt$^{Asp}$ mt (ARNt$^{Asp}$spidec) reste 260 fois moins bien aspartylé que le substrat homologue d'*E. coli*. Ainsi, il reste des éléments dans cet ARNt empêchant sa reconnaissance parfaite par l'enzyme d'*E. coli*. Il reste à rechercher quels sont ces éléments de séquence et de structure. Nous proposons que la stabilité de l'ARNt puisse avoir un rôle à jouer. Ceci pourra être testé par la substitution de paires de bases A-U par des paires de bases G-C.

Une manière plus radicale et plus générale de détection des éléments « perturbateurs » consistera en la transplantation des éléments d'identité de l'ARNt$^{Asp}$ d'*E. coli* dans d'autres ARNt mt humains de séquences et structures bizarres (par exemple ARNt$^{Lys}$, ARNt$^{Tyr}$, ARNt$^{Leu}$ disponibles au laboratoire). Ceci permettra de vérifier (i) si ces ARNt deviennent des substrats efficaces, confirmant alors que la paire de bases A11-U24 est bien un antidéterminant spécifique de l'ARNt$^{Asp}$ mt, ou (ii) ne le deviennent pas et présentent alors d'autres antidéterminants. De façon globale, des analyses supplémentaires devront être réalisées au sein d'autres systèmes pour déterminer les raisons fines de la non aminoacylation des ARNt mt par des aaRS hétérologues. Celles-ci incluent des essais d'aminoacylations croisées mais également par des approches structurales comme la cristallisation de complexes ARNt/aaRS.

# D. Essais de cristallisation

Une grande partie de mon travail de thèse a consisté en l'étude fonctionnelle du système d'aspartylation mt humain. Entre autres, nous avons montré que les éléments majeurs d'identité de l'ARNt$^{Asp}$ mt humain se trouvent dans le triplet anticodon et que la base discriminatrice en 73 ne joue plus de rôle dans la spécificité d'aminoacylation. L'utilisation du modèle 3D de l'AspRS mt, réalisé à partir de la structure cristallographique de l'AspRS d'*E. coli* par Claude Sauter, a été un outil puissant pour avoir des indices quant aux mécanismes sous-jacents à la reconnaissance de l'ARNt par la synthétase. Cependant, le mode exact d'interaction entre l'ARNt et l'enzyme reste inconnu, et notamment le mode d'adaptation de l'enzyme à son substrat structuralement dégénéré. Marie Messmer, qui a réalisé son Diplôme d'Etude Approfondie et commence sa première année de thèse au laboratoire, a débuté une analyse structurale par des études en solution de la structure de l'ARNt$^{Asp}$ mt humain soit libre soit en complexe avec l'AspRS mt. Pour ma part, j'ai tenté d'obtenir des cristaux de l'AspRS mt libre ou en complexe avec l'ARNt homologue. Ceci a été réalisé en collaboration avec Bernard Lorber (IBMC).

Ces essais de cristallisation ont pour objectif non seulement d'enrichir les connaissances structurales au niveau des systèmes d'aminoacylation mt de mammifères (rappelons que seule la SerRS mt bovine en complexe avec le séryl-adénylate a été cristallisée et sa structure résolue à ce jour, (Chimnaronk et al., 2004; Chimnaronk et al., 2005)), mais aussi d'ajouter au catalogue de structures connues d'AspRS libre ou en complexe. Notons qu'il existe une structure cristallographique d'AspRS pour chacun des règnes vivants mais aucune d'enzyme d'organelles.

Nous avons tiré profit du clonage et de la surproduction de l'AspRS pour préparer des quantités suffisantes (dans la mesure du possible) d'enzyme et avons effectué de nombreuses purifications sur gels de polyacrylamide dénaturants préparatifs de l'ARNt$^{Asp}$ mt transcrit *in vitro* en vue de co-cristallisation.

## 1. Approche préliminaire et stratégies

Les conditions d'expression et de purification (50 mM $NaH_2PO_4$ pH 7,5, 300 mM NaCl, 20 mM imidazole, 10 % (v/v) glycérol, 10 mM β-mercaptoéthanol), mises au point précédemment, conduisent à l'obtention de l'AspRS mt humaine pure à plus de 95 %, à une concentration d'environ ~0,6 mg/ml et sous la forme d'un dimère homogène dans le tampon de conservation (50 mM $KH_2PO_4/K_2HPO_4$ pH 7,5, 150 mM KCl, 1 mM EDTA, 10 mM β-mercaptoéthanol, 50 % (v/v) glycérol) (voir article n°3). De plus, l'enzyme est stable à 20°C. Elle reste active après 71 h d'incubation à cette température et sa migration sur gel dénaturant ne révèle pas de produits de dégradation. Ainsi, les essais de cristallisation ont été réalisés dans une chambre tempérée à 20°C.

Nous avons dans un premier temps choisi d'aborder nos essais de façon simple, en présence ou en absence de l'$ARNt^{Asp}$ mt transcrit, avec l'enzyme obtenue comme indiqué ci-dessus mais dans un tampon de conservation dont le pourcentage de glycérol a été ramené à 20 % (une trop forte concentration en glycérol pourrait diminuer les chances de cristallisation en augmentant de façon trop importante la solubilité de la protéine). Cependant, deux aspects sont négatifs : (i) la protéine est faiblement concentrée (~0,6 mg/ml), (ii) elle est dans un tampon contenant du phosphate ($KH_2PO_4/K_2HPO_4$) or la combinaison des ions phosphates et de cations divalents (par exemple $Mg^{2+}$) donnent des sels très peu solubles formant des cristaux plus facilement que la protéine. Ainsi, nous avons essayé de concentrer l'échantillon (voir Matériel et méthodes) et/ou de remplacer le tampon de conservation par dialyse.

Ces premières tentatives ont toutes conduit à la précipitation de la protéine (visible à l'œil nu). D'autres stratégies ont été entreprises : (i) modification des conditions d'expression de la protéine afin de produire l'AspRS mt dans un contexte plus favorable à son repliement, (ii) utilisation de différents additifs lors de la purification pour pallier au remplacement du tampon contenant du phosphate.

## 2. Recherche de conditions de cristallisation

Les principes généraux de cristallisation des macromolécules biologiques et, en particulier, de la technique de diffusion de vapeur, sont donnés dans la partie Matériel et Méthodes. Les essais de cristallisation ont été réalisés avec l'AspRS mt seule (à partir d'un stock de protéine obtenu d'après les conditions de purification décrites précédemment) ou en présence de l'ARNt$^{Asp}$ mt (transcrits *in vitro* et stockés à –20°C dans le tampon 10 mM MES pH 6 à la concentration de 156 µM). Nous avons choisi la méthode de diffusion de vapeur (les gouttes sont assises ; volume du réservoir 70 µl et goutte de 6 µl, soit 3 µl d'échantillon et 3 µl de solution du réservoir pour les boîtes remplies à la main). Les boîtes sont mises à équilibrer en chambre tempérée à 20°C pendant quelques jours ou plusieurs semaines, et les gouttes sont régulièrement observées à l'aide d'une loupe binoculaire (grossissement 6 à 50 x).

La recherche de conditions de cristallisation est extrêmement délicate du fait du grand nombre de paramètres (pH, agent précipitant, additifs, concentration en molécules, impuretés, etc) pouvant influencer la nucléation ou la croissance cristalline (Ducruix & Giegé, 1999). Ainsi, comme aucune loi ne permet, *a priori*, de prédire la composition de la solution la plus favorable, l'approche généralement utilisée consiste à tester un grand nombre de conditions au travers d'une « matrice clairsemée ». Les matrices peuvent réunir jusqu'à 96 combinaisons différentes de substances tampons, d'agents précipitants et d'additifs, qui ont pour la plupart permis la cristallisation de molécules. L'utilisation de tampon permet de maintenir le pH du milieu de cristallisation et de stabiliser les molécules présentes. L'agent précipitant est employé pour diminuer la solubilité des molécules et de permettre leur cristallisation. Les composés additifs, généralement des sels, permettent l'apport d'ions nécessaires au maintien de la conformation des macromolécules (par exemple le $Mg^{2+}$). Ils ont aussi pour fonction de moduler la solubilité des molécules à cristalliser. Ici, nous avons tenté de cristalliser l'AspRS mt seule ou en complexe avec l'ARNt$^{Asp}$ mt transcrit en utilisant les matrices commerciales Wizard$^{TM}$ I et II (commercialisées par Emerald BioStrutures) et une matrice de 24 conditions préparées par Bernard Lorber. Les compositions de l'ensemble des solutions de cristallisation testées sont données dans la partie Matériel et Méthodes. Les matrices Wizard$^{TM}$ I et II empiriques réunissent 96 conditions, permettant de tester16 agents cristallisants et 11 tampons différents (de pH 4,5 à pH 10,5) pour la cristallisation de macromolécules biologiques (protéines, acides nucléiques ou complexes). La matrice empirique réalisée par Bernard

Lorber rassemble 24 conditions basées sur le mélange des agents précipitants et additifs les plus couramment utilisés pour la cristallisation de protéines. La composition des solutions sont données dans la partie « Essais de cristallisation » de Matériel et Méthodes.

Un grand nombre de conditions testées ont conduit à la formation de cristaux. Cependant, tous ces cristaux sont uniquement composés de sels. En effet, la solution contenant l'AspRS mt seule est composée d'un tampon contenant du phosphate, connu pour générer des cristaux de sels avec les solutions généralement utilisées pour la cristallisation de macromolécules biologiques et utilisées ici. En ce qui concerne les essais de cristallisation du complexe, seules les matrices Wizard$^{TM}$ I et II ont été testées et pour l'instant aucun résultat probant n'a pu être obtenu.

## 3. Echange de tampon et effet sur la solubilité de l'AspRS

Dans le but de se défaire du tampon contenant du phosphate, nous avons réalisé différentes tentatives de dialyse de l'AspRS en fin de purification contre de nouveaux tampons (Tableau 8). Différents paramètres ont été testés : la nature et le pH du tampon, la nature et la concentration en sels. Nos choix ont aussi été orientés par les résultats déjà obtenus au labboratoire. Ainsi, le tampon Tris-HCl a été montré comme adéquat pour d'autres aaRS mt au laboratoire, et en particulier pour la GluRS, alors que celui composé de Hepes-NaOH est nécessaire à la solubilisation de la TyrRS mt. Parmi les différents types de sels testés, le sulfate d'ammonium a déjà été utilisé pour l'AspRS de levure (Sauter et al., 1999). Malheureusement, l'ensemble de ces tests d'échanges de tampons effectués a conduit à la précipitation de l'AspRS mt.

Remarque : l'EDTA qui est présent dans le tampon de conservation permet de complexer les ions $Ni^{2+}$ éventuellement décrochés de la matrice lors de la purification. Cet additif aurait pu améliorer la solubilité de notre protéine et aurait éventuellement du être ajouté de façon systématique.

| Substance tampon | Sel | Autres ingrédients | Solubilité |
|---|---|---|---|
| 50 mM $KH_2PO_4/K_2HPO_4$ pH 7,5 | 150 mM KCl | 10 mM β-Me<br>50 % (v/v) glycérol<br>1 mM EDTA | + |
| 50 mM $KH_2PO_4/K_2HPO_4$ pH 7,5 | 150 mM KCl | 10 mM β-Me<br>20 % (v/v) glycérol<br>1 mM EDTA | + |
| 10 mM Hepes-NaOH pH 7,2 | | | - |
| 50 mM Hepes-NaOH pH 7,2 | 10 mM NaCl | | - |
| 50 mM Tris-HCl pH 6,8 | 10 mM NaCl | | - |
| 50 mM Tris-HCl pH 6,8 | 50 mM NaCl | | - |
| 50 mM Tris-HCl pH 8,1 | 50 mM NaCl | | - |
| 50 mM Tris-HCl pH 8,1 | 100 mM NaCl | | - |
| 100 mM Tris-HCl pH 8,1 | 150 mM NaCl | | - |
| 100 mM Tris-HCl pH 8,1 | 800 mM sulfate d'ammonium | | - |
| 100 mM Tris-HCl pH 8,1 | 200 mM sulfate d'ammonium | 10 % (v/v) glycérol | - |
| 50 mM Hepes-KOH pH 7,2 | 40 mM $MgCl_2$ | 10 % (v/v) glycérol<br>3 mM β-Me | - |

**Tableau 8**. **Composition des différentes solutions de dialyse testées pour l'AspRS mt humaine.** Différents paramètres ont été variés : nature, concentration et pH du tampon ; nature et concentration du sel ; présence d'additifs. β-Me pour β-mercaptoéthanol.

## 4. Conditions d'expression et de purification

Suite à l'impossibilité de remplacer le tampon de l'AspRS mt humaine en sortie de colonne de purification, nous avons choisi de modifier les conditions des étapes précédentes, à savoir celles de surexpression ou de purification. L'objectif a été d'optimiser les conditions d'obtention de la protéine afin d'avoir un échantillon homogène. Nous avons également essayé de resolubiliser la protéine précipitée en utilisant des matrices clairsemées (voir plus loin). Les résultats sont donnés dans la partie suivante.

Les différents tests d'expression de l'AspRS mt humaine ont été effectués dans la souche TOP10 d'*E. coli*. Nous avons essayé de jouer sur plusieurs paramètres, de façon indépendante, afin de permettre une expression ralentie des protéines permettant ainsi le

repliement correct. Pour ce faire, nous avons : (i) réduit la température d'induction à 18°C, (ii) ajouté une quantité minimale d'IPTG de 50 µM, ou (iii) ralenti la croissance des bactéries en ajoutant 20 % (p/v) de sucrose dans le milieu de culture. Cependant aucune des artifices testés ne permet une meilleure expression de l'AspRS mt humaine, et la protéine obtenue n'est pas plus soluble après purification.

Dans la cellule, il existe des chaperons protéiques qui permettent le bon repliement des protéines synthétisées (Bukau & Horwich, 1998). Une façon simple d'induire l'expression de telles protéines chaperonnes est d'effectuer un choc thermique à 42°C avant l'induction de l'expression de la protéine recombinante. Une autre méthode consiste à ajouter des chaperons chimiques, remplaçant les chaperons protéiques. Par exemple, l'ajout de glycérol lors de la période d'induction de l'expression de la protéine a déjà porté ses fruits pour l'expression de protéines humaines dans *E. coli* (Leandro et al., 2001), ou dans la levure (Figler et al., 2000). Les deux méthodes ont été testées mais n'ont pas apporté d'amélioration significative dans la production de notre protéine.

D'autre part, nous avons essayé de jouer sur les conditions de purification en ajoutant des additifs tels que du glycérol (10-20 % (v/v)) ou des détergents (triton, zwittergent) (Tableau 9). Ces additifs sont largement utilisés pour accroître la solubilité des protéines. Malheureusement, les rendements de purification dans les nouvelles conditions ont été globalement moins bons et la solubilité de la protéine n'est pas accrue (précipitation lors de la concentration). De plus, l'AspRS mt n'est plus active en présence des détergents que nous avons testés.

En conclusion, aucune des conditions d'expression ou de purification n'a permis une meilleure production de l'AspRS mt. La protéine ne se concentre toujours pas mieux. Mais, l'ensemble des essais a apporté des informations quant à la solubilité réduite de l'AspRS. Le tampon contenant du phosphate restant celui optimal pour l'obtention de l'AspRS mt homogène (dimérique) à faible concentration, il est probable qu'il faudra persévérer en essayant de rajouter l'ARNt (qui est composé d'une chaîne polyphosphate) lors des échanges de tampons et de la concentration.

| Substance tampon | Sel | Autres ingrédients |
|---|---|---|
| **Conditions optimales** | | |
| 50 mM $NaH_2PO_4$ pH 7,5-9,0 | 300 mM NaCl | 20 mM imidazole |
| 50 mM $NaH_2PO_4$ pH 7,5-9,0 | 300 mM NaCl | 20 mM imidazole<br>10-20 % glycérol<br>10 mM β-Me |
| **Autres conditions testées** | | |
| 50 mM $NaH_2PO_4$ pH 8,0 | 300 mM NaCl | 10 % glycérol<br>10 mM β-Me<br>0,1 % triton |
| 100 mM Hepes-NaOH pH 7,5 | 300 mM NaCl<br>20 mM $MgCl_2$ | 10 % glycérol<br>10 mM β-Me<br>0,1 % triton |
| 100 mM Hepes-NaOH pH 7,5 | 300 mM NaCl<br>20 mM $MgCl_2$ | 10 % glycérol<br>10 mM β-Me<br>0,2 mM Zwittergent 3-14 (Fluka) |

**Tableau 9. Composition des différentes solutions de purification testées pour l'AspRS mt humaine.** Différents paramètres ont été variés : composition du tampon, nature du sel et des additifs ajoutés. Notons que le glycérol ou les détergents (à faible concentration) permettent de solubiliser les macromolécules biologiques. β-Me pour β-mercaptoéthanol.

## 5. Conclusions et perspectives

Bien que les résultats soient restés négatifs pour le moment, les potentialités de succès restent grandes. En effet, les essais de cristallisation de Luc Bonnefond et Marie Sissler de la TyrRS et de la GluRS mt humaines, respectivement, sont très encourageants. De plus, l'ajout de DTT pour l'obtention de la SerRS mt de bœuf, première aaRS mt de mammifère cristallisée et dont la structure a été résolue récemment (Chimnaronk et al., 2004; Chimnaronk et al., 2005), a été également utile à la production de la GluRS non agrégée. Nous pensons en effet que les conditions permettant de solubiliser et de cristalliser une aaRS mt humaine pourrait aider à résoudre les problèmes de solubilité et de cristallisation des autres. Des indices sont en faveur de l'association des ribosomes à la membrane interne de la mitochondrie. Ceci suggère que les aaRS mt pourraient être localisées de la même façon, et par conséquent partageraient des propriétés physicochimiques communes de celles retrouvées pour les protéines membranaires.

En perspective, l'AspRS mt devra être obtenue à un degré de pureté plus élevé, en effectuant d'autres étapes de chromatographie, car les impuretés peuvent conduire à la mauvaise cristallisation, voir l'empêcher. D'autre part, de nouveaux essais devront être faits, en présence par exemple de l'ARNt$^{Asp}$ mt humain surproduit dans *E. coli* (qui sera modifié et probablement de conformation plus homogène) ou de l'ARNt$^{Asp}$ d'*E. coli* (qui est un très bon substrat de l'AspRS mt humaine), et/ou en présence des autres substrats de l'enzyme, tels que l'aspartate ou l'ATP (ou des analogues non hydrolysables de l'adénylate). De nombreux cristaux de synthétases ont en effet été obtenus en présence de l'un et/ou l'autre de ces substrats. L'addition d'autres composés doit également être envisagée. Enfin, l'acquisition du robot Mosquito dans l'institut, capable de réaliser des milliers de nanogouttes et des boîtes de cristallisation en moins de 5 min, permettra de tester en masse d'autres matrices clairsemées à la fois dans le but de solubiliser la protéine (il faudra peut-être tester des conditions moins riches ou ne contenant pas d'agent précipitant pour la cristallisation) et de la cristalliser.

Finalement, Il faudra également tester d'autres organismes tels que la levure ou les cellules d'insectes en utilisant baculovirus pour exprimer l'AspRS mt. Ce dernier système a déjà fait ses preuves pour l'expression d'aaRS mt humaine (Jørgensen et al., 2000). La production d'une protéine eucaryotique peut s'avérer plus efficace dans un système eucaryotique (par exemple, les modifications de la protéine dans l'organisme d'origine, ont plus de probabilité d'être effectuées dans un organisme producteur proche phylogénétiquement). Par exemple, il a été montré que l'ArgRS cytoplasmique de hamster produite dans la levure est plus active que celle produite dans *E. coli* (Guigou et al., 2004).

# CONCLUSIONS & PERSPECTIVES

Le travail de cette thèse s'inscrit dans le cadre de l'étude des règles qui régissent la spécificité d'aminoacylation des ARNt par les aaRS. La précision de cette réaction est cruciale puisqu'elle détermine la fidélité de la traduction de l'information génétique et la synthèse de protéines fonctionnelles. De nombreuses études portant sur la compréhension des systèmes d'aminoacylation ont déjà été réalisées, cependant leur exploration est loin d'être achevée. En effet, les connaissances accumulées concernent surtout les systèmes procaryotiques (par exemple *E. coli, T. thermophilus*) et dans une moindre mesure ceux de la levure *S. cerevisiae*. Par contre, les systèmes eucaryotiques supérieurs ou d'organelles restent peu connus, en particulier ceux humains. Les différentes lignes de recherche de cette thèse ont porté sur 5 points majeurs : (i) l'activité relative d'ARNt isoaccepteurs dans la levure, (ii) les relations structures/fonctions d'ARNt de différentes spécificités dans la levure, (iii) la mise en place de nouveaux systèmes d'aminoacylation mt humains, (iv) les signaux d'identité permettant l'aminoacylation spécifique dans un système mt humain, (iv) la non-reconnaissance des ARNt mt par les aaRS bactériennes et l'aminoacylation croisée d'ARNt de différentes origines par les aaRS mt. Différentes techniques de la biologie moléculaire ont été mises à profit. Ainsi l'utilisation de transcrits *in vitro* d'ARNt et d'enzymes clonées, surproduites et purifiées nous ont permis de réaliser de multiples analyses mutagéniques. Les aspects structuraux et fonctionnels ont également été explorés par des essais de cristallisation et par une approche *in vivo* utilisant une souche « knock-out » de levure.

## A. Systèmes de levure

### Comportements idiosyncrasiques d'une famille d'ARNt

Pour un système donné, il existe souvent plusieurs ARNt isoaccepteurs. Etant tous reconnus par la même aaRS, il était aisé de proposer des règles de reconnaissance et d'aminoacylation identiques pour l'ensemble des molécules d'une même famille. Or cela n'avait pas encore été vérifié expérimentalement, ce que nous avons choisi de faire pour les 4 isoaccepteurs du système arginine de *S. cerevisiae*. Les signaux majeurs d'identité arginine

C35 et G/U36 avaient été déterminés dans l'$ARNt_3^{Arg}$ précédemment au laboratoire et se retrouvent dans la séquence des 3 autres $ARNt^{Arg}$. Cependant, en dehors de ces éléments, les séquences des $ARNt^{Arg}$ sont largement moins conservées puisqu'elles ne partagent que 37 % d'identité. Nous avons montré que ces différences peuvent générer des comportements idiosyncrasiques :

(i) au niveau de la reconnaissance homologue par l'ArgRS. Les efficacités relatives d'arginylation des 4 $ARNt^{Arg}$ sont variables, un facteur 20 séparant le meilleur substrat ($ARNt_1^{Arg}$) du moins efficace ($ARNt_3^{Arg}$). La présence d'éléments mineurs propres à chaque isoaccepteur est largement suspectée.

(ii) au niveau des moyens mis en œuvre pour se protéger de la reconnaissance non homologue. Nos travaux ont montré que les 4 $ARNt^{Arg}$ se protègent différemment de l'aspartylation croisée. Après transplantations des éléments d'identité aspartate, les $ARNt_1^{Arg}$ et $ARNt_4^{Arg}$ deviennent efficacement aspartylables, alors que les $ARNt_2^{Arg}$ et $ARNt_3^{Arg}$ restent inactifs à cause de la présence d'antidéterminants. De façon intéressante, la nature de ces antidéterminants n'est pas la même pour les 2 molécules.

Ainsi, des ARNt ne se distinguent pas seulement les uns des autres par le codon à lire, mais encore, nous avons montré qu'ils possèdent chacun des propriétés propres, probablement relatives aux variations de séquence. Il sera intéressant d'étendre cette observation aux autres « familles nombreuses » tels que les systèmes leucine et sérine. Un cas particulier concerne les deux systèmes de la mitochondrie humaine où il existe 2 isoaccepteurs pour chacun de ces systèmes, très différents structuralement les uns des autres. De plus, les fréquences relatives de codons à lire par chacun d'eux sont largement disproportionnées (les $ARNt^{Leu(CUN)}$ et $ARNt^{Leu(UUR)}$ lisent respectivement 86 % et 14 % des codons leucine, et les $ARNt^{Ser(UCN)}$ $ARNt^{Ser(AGY)}$ lisent respectivement 81 % et 19 % des codons sérine).

## Les $ARNt_4^{Arg}$ et $ARNt^{Asp}$ descendent d'un ancêtre commun

Dès 1973, un lien fonctionnel a été établi au laboratoire entre les systèmes arginine et aspartate de *S. cerevisiae*, avec l'isolement du complexe hétérologue formé de l'ArgRS et de l'$ARNt^{Asp}$. Plus tard, il a été montré que l'$ARNt^{Asp}$ transcrit est un substrat efficace de l'ArgRS et se protège d'une telle erreur dans la cellule par une modification post-transcriptionnelle. L'existence d'un 4ème isoaccepteur arginine ressemblant fortement à l'$ARNt^{Asp}$ et se distinguant des 3 autres $ARNt^{Arg}$ nous a permis d'approfondir la relation entre les deux systèmes d'aminoacylation.

Les propriétés d'aspartylation de l'ARNt$_4^{Arg}$ sauvage ou muté, transcrit *in vitro* ou exprimé dans la levure, ont été explorées et nous ont amenés à proposer que les ARNt$_4^{Arg}$ et ARNt$^{Asp}$ soient des molécules paralogues. Ils proviendraient de la duplication d'un gène unique, dont le produit aurait été un ARNt à spécificité ambiguë (aspartylable et arginylable). La différenciation vers des ARNt à spécificité stricte serait le résultat de la mise en place (au cours de l'évolution de ces systèmes) de verrous moléculaires (e.g. mutations des nucléotides aux positions 38 et 73), combinée à de phénomènes épigénétiques (e.g. modifications post-transcriptionnelles) et/ou de compétition dans le contexte cellulaire. Cette réduction d'amibiguïté est d'avantage illustrée par des expériences menées *in vivo*. En effet, la complémentation d'une souche de levure délétée du gène endogène de l'ARNt$_4^{Arg}$ par le gène de l'ARNt$_4^{ArgC38/G73}$ n'est plus double spécifique dans un contexte cellulaire.

La comparaison d'un grand nombre de séquences d'ARNt$^{Arg}$ et ARNt$^{Asp}$ de différentes levures a montré que la relation évolutive entre les gènes *ARNt$_4^{Arg}$* et *ARNt$^{Asp}$* est étendue à un sous-groupe de levure (proche de *S. cerevisiae*). Le dendrogramme généré à partir des séquences de tous les gènes d'ARNt chez *S. cerevisiae* a montré d'un des 3 des isoaccepteurs de la thréonine (ARNt$^{Thr(UGU)}$) est proche de l'ARNt$^{Asn}$ et divergent des deux autres. Ceci suggère qu'il pourrait exister le même type de relation évolutive que celle montrée pour les ARNt$_4^{Arg}$ et ARNt$^{Asp}$.

La mise en évidence d'une origine différente pour l'ARNt$_4^{Arg}$ par rapport aux 3 autres ARNt$^{Arg}$ renforce nos résultats à propos des comportements idiosyncrasiques de ces molécules. Ceci illustre bien que l'occurrence d'ARNt dans une même famille isoacceptrice n'est pas forcément due à une origine commune mais peut être le résultat, dans le cas au moins de l'ARNt$_4^{Arg}$, d'une convergence fonctionnelle.

# B. Systèmes mitochondriaux de mammifères

Les systèmes mt de mammifères sont particulièrement intéressants car leurs ARNt, codés par le génome mt, possèdent des structures dégénérées alors que les aaRS mt, codées par le génome nucléaire, ressemblent aux synthétases classiques. De plus, dans le cas des systèmes mt humains, de nombreuses mutations retrouvées dans les gènes d'ARNt sont corrélées à diverses pathologies. L'accumulation de connaissances fondamentales permettra

de comprendre mieux ces systèmes atypiques et peut-être d'en contrecarrer les disfonctionnements.

### Vers le jeu complet de gènes d'aaRS mitochondriales humaines

Au début de ce travail, les gènes des 22 ARNt nécessaires à la traduction mt étaient séquencés alors que seulement 7 gènes d'aaRS mt avaient été identifiés, clonés et les protéines correspondantes caractérisées (IleRS, PheRS, LyRS, LeuRS, TrpRS, SerRS et MetRS) et deux autres gènes avaient été caractérisés mais les protéines non étudiées (ceux codant pour la GlyRS mt et l'HisRS mt). Grâce aux avancées dans le séquençage du génome humain, nous avons pu assigné les gènes manquants (à l'exception de celui de la GlnRS) et ainsi avoir une vue d'ensemble des aaRS mt humaines. De façon globale, (i) les aaRS mt sont codées par un gène distinct de ceux des enzymes cytoplasmiques (excepté les LysRS et GlyRS pour lesquels un seul gène code pour les enzymes mt et cytoplasmiques), et (ii) les protéines suivent toutes la répartition prévue en deux classes de synthétases selon la nature structurale de leur site catalytique. L'accès à ces dix nouvelles séquences ouvre de nombreuses perspectives telles que la caractérisation structurale et fonctionnelle des couples aaRS/ARNt, la recherche de partenaires, de produits d'épissage alternatif, etc…

### Evolution reverse de l'identité de l'ARNt$^{Asp}$ mt et adaptation focalisée de l'AspRS

Parmi ces nouveaux systèmes, nous avons considéré le système aspartate comme modèle d'étude des mécanismes sous-jacents à la spécificité de reconnaissance ARNt/aaRS. La séquence de l'ARNt$^{Asp}$ est dégénérée alors que l'AspRS mt est classique et de type bactérien (avec plus de 40 % d'identité de séquence avec l'AspRS *d'E. coli*). Nous avons montré que le jeu d'identité aspartate est restreint par rapport aux systèmes classiques. Alors que les éléments de l'anticodon restent des éléments majeurs, la position 73 pourtant universellement représentée par un G dans les systèmes aspartate non mitochondriaux de mammifères ne participent plus à l'identité de l'ARNt$^{Asp}$ mt. L'AspRS mt humaine s'est adpatée de manière ciblée à la dégénérescence de son substrat. Alors que l'AspRS bactérienne interagit spécifiquement avec l'élément G73 par le biais de deux acides aminés (Asp220 et Arg222), l'enzyme mt n'a conservé que l'Arg et possède une Gly à la place de l'Asp. La substitution expérimentale de cette Gly par un Asp (mutant AspRSG269D) restaure le pouvoir discriminant de l'enzyme vis-à-vis du résidu G73. La construction d'un modèle 3D de l'enzyme et l'analyse d'alignements multiples de séquences d'AspRS de différents

organismes ont été déterminants dans l'identification du résidu impliqué. Ils seront également très utiles pour la recherche d'autres particularités fines de la synthétase qui font que bien qu'étant très proche de l'enzyme bactérienne, elle en est fonctionnellement distincte. Enfin, si l'on considère que les systèmes d'aminoacylation primordiaux étaient constitués d'aaRS et de substrats minimalistes, réduits respectivement aux seuls domaines catalytiques et branches acceptrices, G73 faisait probablement partie du jeu d'identité aspartate primordial. Sa perte en tant qu'élément spécifiant l'aspartylation dans la mitochondrie témoigne d'une évolution réverse des systèmes d'aminoacylation mt.

### Unicité d'aminoacylation de l'ARNt$^{Asp}$

Les aaRS mt de mammifères aminoacylent des ARNt d'organismes variés, alors que l'aminoacylation des ARNt mt de mammifères par des synthétases autres n'a pas encore été observée. Il a été suggéré que les particularités structurales (longueur variables des boucles D et T) étaient responsables d'une telle unilatéralité d'aminoacylation. Nous avons montré que l'AspRS mt humaine est capable d'aminoacyler une grande gamme d'ARNt$^{Asp}$ (d'*E. coli*, de *T. thermophilus*, cytoplasmiques humain et de levure), mais qu'au contraire l'ARNt$^{Asp}$ mt humain n'est reconnu que par l'enzyme mt. Une étude approfondie des séquences nous a permis de créer des mutants de l'ARNt$^{Asp}$ mt devenant progressivement des substrats efficaces pour l'AspRS d'*E. coli*. De façon intéressante, nous avons montré que l'inversion de la paire de bases A11-U24 dans le bras D (Y11-R24 dans les ARNt classiques), en combinaison avec l'introduction des éléments d'identité aspartate nécessaires à l'enzyme d'*E. coli*, a un impact plus important sur l'aminoacylation hétérologue de l'ARNt$^{Asp}$ mt que l'introduction d'une structure canonique. Cette paire de bases serait impliquée dans la voie d'antidétermination de l'ARNt mt humain.

Cette découverte ne représente qu'une réponse partielle à la question de la non aminoacylation croisée des ARNt mt par des aaRS hétérologues. Il s'agira maintenant non seulement de mener à terme la compréhension du système aspartate, mais encore de vérifier la situation dans les autres et en particulier de discerner si un mécanisme d'antidétermination intervient de manière systématique ou non.

## Perspectives liées à l'étude du système d'aspartylation mt humain

### *Jeux d'identité simplifiés dans la mitochondrie de mammifère : généralité ?*

Il sera intéressant, maintenant que tous les gènes des aaRS sont identifiés, de réaliser une analyse globale des systèmes d'aminoacylation mt humains afin d'établir des règles de spécificité dans le compartiment mitochondrial. Notamment, ceci pourrait permettre de confirmer que les jeux d'identité sont simplifiés et de savoir par des tests d'aminoacylation croisées s'ils sont suffisants en tant que tels pour maintenir la précision d'aminoacylation. Divers arguments sont en faveur de la simplification des jeux d'identité. Les travaux de Luc Bonnefond effectués sur le système tyrosine vont également dans ce sens. En effet, l'aminoacylation de l'ARN$^{Tyr}$ mt humain par la TyrRS correspondante ne requiert plus la présence de la paire de bases G1-C72, qui pourtant constitue un élément d'identité fondamental dans les ARNt$^{Tyr}$ de procaryotes. En contrepartie, les éléments de l'anticodon auraient un rôle plus fort. L'analyse des séquences des ARNt mt de mammifères disponibles montre en effet que les éléments d'identité d'*E. coli* les plus conservés sont essentiellement ceux retrouvés aux deux extrémités des ARNt. Ainsi et de façon générale, la simplification du génome mt pourrait s'accompagner de jeux d'identité constitués d'un plus petit nombre d'éléments.

### *Analyse structurale de couples aaRS/ARNt mt humains*

L'obtention de structures cristallographiques d'aaRS mt humaines libres ou en complexe avec l'un ou l'autre de leurs substrats reste une priorité. Ceci apportera des connaissances considérables sur les systèmes d'aminoacylation mt de mammifères, à l'heure où les données structurales restent encore très pauvres puisque seule la structure de la SerRS mt de bœuf est connue et ce depuis peu (Chimnaronk et al., 2004 ; Chimnaronk et al., 2005). Un grand projet de génomique structurale, initié au laboratoire par Marie Sissler, a déjà permis de cloner tous les gènes des aaRS mt nouvellement assignés. Malheureusement, les protéines mt humaines restent difficiles à surproduire dans *E. coli*. Toutefois, l'obtention récente de la GluRS mt sous forme non agrégée et monodisperse nous encourage vivement à poursuivre, et notamment à persévérer dans la production et la purification de l'AspRS mt. Par ailleurs, des indices sont en faveur de l'association des ribosomes à la membrane interne de la mitochondrie, suggérant une localisation semblable pour les aaRS mt. Par conséquent,

elles pourraient partager des propriétés physico-chimiques de celles retrouvées dans les protéines membranaires. Nos recherches tiendront compte de cela.

### *Recherche de partenaires des systèmes mt humains*

Les particularités fonctionnelles des systèmes mt (propriétés physicochimiques défavorables, activités réduites,…) laissent supposer qu'il pourrait exister des co-facteurs ou chaperons qui stabilisent l'ARNt ou améliorent l'activité enzymatique de l'aaRS. La contribution de protéine partenaire a déjà été décrite pour certaines aaRS non mitochondriales (*e.g.* l'association de Arc1p qui améliore l'activité de la MetRS cytoplasmique de levure). Selon cette idée, l'interaction de co-facteurs pourrait aussi permettre la formation d'un complexe multisynthétasique comme c'est le cas des enzymes cytoplasmiques. Un argument en faveur de cette hypothèse est que l'un des facteurs retrouvés dans le complexe cytoplasmique semble être aussi adressé vers la mitochondrie (il s'agit de la protéine p43).

Les des anticorps dirigés contre l'AspRS mt produits dans ce travail seront rapidement utilisés pour tenter de répondre à cette question. En effet, des expériences de co-immunoprécipitation des protéines liées à l'AspRS mt à partir de lysats mt pourront être réalisées. Une approche complémentaire de TAP-tag, basée sur l'expression *in vivo* des aaRS fusionnées à deux peptides (permettant la rétention sélective sur deux résines différentes) et de la séquence d'adressage vers la mitochondrie pourra être effectuée. Cette étude sera menée en parallèle pour au moins deux aaRS (AspRS et GluRS) pour lesquelles des anticorps spécifiques ont été produits.

En parallèle, l'utilisation des anticorps permettra de vérifier la localisation cellulaire des aaRS mt par hybridation *in situ*, et notamment de savoir si elles sont associées à la membrane interne de la mitochondrie. La réponse à ce dernier point permettra d'affiner encore les recherches de conditions d'expression et de solubilisation des aaRS mt.

### *Vérification de l'épissage alternatif de l'AspRS mt*

L'aboutissement du séquençage du génome humain et le recensement de seulement 32 000 gènes pour un besoin en protéines fonctionnelles bien plus important révèle la complexité du rapport génome/protéome humain. Ceci renforce le rôle majeur de mécanismes d'épissage alternatif des pré-ARNm. La possibilité de tels évènements au niveau des gènes nouvellement identifiés des aaRS mt pourra être abordée. Des analyses et expériences préliminaires sont prometteuses en ce sens. Pour exemple, l'obtention d'un fragment supplémentaire et plus court après PCR (ciblée sur le gène de l'AspRS) à partir d'une banque

d'ADNc humains soutient cette hypothèse. Ce fragment correspond exactement à la séquence de l'AspRS délétée d'un exon complet. Par ailleurs, durant le clonage des autres gènes d'aaRS (projet de génomique structurale réalisé par Marie Sissler) d'autres produits « tronqués » ont été obtenus (ThrRS délétée de l'exon 8 ou des exons 8 et 9, et LysRS tronquée des exons 4 et 5). Ces premiers constats positifs sont très encourageant. Si la réalité biologique de ces formes tronquées se trouvent avérée (par expériences de northern blot, RT-PCR), les protéines seront exprimées en vue de leur caractérisation structurale et de l'identification de leur rôle biologique.

Ce projet ouvre de nouvelles perspectives vers un rôle alternatif des modules structuraux des aaRS. De telles utilisations de parties d'aaRS pour des fonctions non conventionnelles ont déjà été répertoriées. Or, des partenaires des systèmes d'aminoacylation (molécules chaperonnes et/ou cofacteur) sont pressentis et pourraient correspondre à des produits d'épissage alternatif des gènes d'aaRS.

### *Mécanismes moléculaires sous-jacents aux pathologies mitochondriales à ARNt*

Les travaux menés au cours de cette thèse ont été axés pour l'essentiel vers des aspects fondamentaux, répondant au besoin de connaître le plus largement possible les systèmes mt humains. Bien évidemment, l'ensemble de ces nouvelles connaissances est une contribution importante dans la compréhension de la traduction dans le compartiment mitochondrial, prérequis indispensable à la compréhension des mécanismes moléculaires sous-jacents aux pathologies mitochondriales liés aux ARNt, et pourront être exploitées dans ce sens.

# Matériel & Méthodes

## A. Matériel

### 1. Produits chimiques

| | |
|---|---|
| [$\gamma^{32}$P]-ATP / [$\alpha$ $^{32}$P]-ATP / [$\alpha$ $^{32}$P]-pCp | Amersham Pharmacia |
| Acrylamide 40 % | Bio-Rad |
| Ampicilline | Carl Roth |
| Bleu de Coomassie R250 | Carl Roth |
| Chloramphénicol | Amersham Pharmacia |
| EDTA | Boehringer Mannheim |
| Glycine | Sigma-Aldrich |
| IPTG | Euromedex |
| Kanamycine | Sigma-Aldrich |
| Arginine | Sigma-Aldrich |
| Aspartate | Sigma-Aldrich |
| [$^{3}$H]-arginine | Amersham Pharmacia |
| [$^{3}$H]-aspartate | Amersham Pharmacia |
| Pipérazine DiAcrylamide (PDA) | Bio-Rad |
| Rotiphorèse Gel (acrylamide and N,N'-methylène-bis-acrylamide) | Carl Roth, GmbH |

### 2. Enzymes commerciales

| | |
|---|---|
| *BamH*I | New England Biolabs |
| *Hind*III | New England Biolabs |
| *Nsi*I | New England Biolabs |
| *BstN*I | New England Biolabs |
| *Bgl*II | New England Biolabs |
| *Dpn*I | New England Biolabs |
| Calf Intestinal Phosphatase (CIP) | Boehringer Mannheim |
| T4 ADN ligase | Qbiogen |
| T4 polynucléotide kinase | New England Biolabs |
| Dynazyme EXT polymérase | Finnzymes |
| *Pfu Turbo*® ADN polymérase | Stratagen |

## 3. *Enzymes non commerciales*

**T7 ARN polymérase :** cette enzyme permet la transcription de gènes sous contrôle du promoteur spécifique. Elle est couramment utilisée pour la synthèse *in vitro* d'ARN et notamment d'ARNt. La T7 ARN polymérase ADN dépendante a été purifiée au laboratoire selon le protocole mis au point par (Becker et al., 1996).

**ArgRS de levure, et AspRS de levure, d'*E. coli* et AspRS 1 et 2 de *T. thermophilus* :** les cinq enzymes ont été clonées, surproduites dans *E. Coli* et purifiées par chromatographie d'affinité nickel (Ni-NTA). Ce type de colonne retient sélectivement l'extension poly-histidines (poly-His) incorporée à l'extrémité N-ou C-terminale de l'enzyme. La protéine est ensuite éluée par gradient d'imidazole (compétition poly-His et imidazole). L'ArgRS de levure et l'AspRS d'*E. coli* ont été généreusement fournies par G. Eriani (IBMC), l'AspRS de levure par M. Frugier (IBMC), et les AspRS 1 et 2 de *T. thermophilus* par H. Becker (IBMC).

**ARNt :** les $ARNt^{Asp}$ natif d'*E. col*i et $ARNt^{Asp}$ transcrit *in vitro* de *T. thermophilus* ont été généreusement fournies par Daniel Kern et Hubert Becker (IBMC).

## 4. *Souches cellulaires*

### 4.1. Souches d'*E. coli*

**TG1 :** *F' [traD36 proAB$^+$ lacI$^q$ lacZΔM15] supE Δ(hsdM-mcrB)5 thi Δ(lac-proAB)*. Cette souche a servie pour le clonage des ARNt.

**TOP10 :** *F$^-$ mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 deoR recA1 araD139, Δ (ara-leu)7697, galU, galK, rpsL, endA1, nupG.* Cette souche a été utilisée pour le clonage des ARNt et de l'AspRS mt humaine.

**TOP10F' :** *F' [lacI$^q$ Tn10 (Tet$^r$)] mcrA Δ(mrr-hsdRMS-mcrBC) ϕlacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (Str$^R$) endA1 nupG.* Cette souche est fournie compétentes dans le kit « TA-cloning » (Invitrogen) et a servi de réceptrice pour les clonages de produit de PCR dans le vecteur pCR®2.1.

**XL1Blue :** *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI$^q$ZΔM15Tn10 (Tet$^r$)].* Cette souche a servi de réceptrice pour la transformation par électroporation du milieu de mutagénèse par PCR.

## 4.2. Souche de *S. cerevisiae*

**YAL5 :** Δ*tR4::His3, ura$_3^-$, lys$_2^-$, trp$_1^-$, his$_3^-$, leu$_2^-$, ade$_2^-$, ade$_3^-$*, possédant le plasmide pAL5 (tR4$^+$,Ura$^+$, Ade$_3^+$). Elle a été disruptée du gène *tR4* codant pour l'ARNt$^{Arg}{}_4$ au laboratoire (Geslain et al., 2003). Le plasmide pAL5 permet de compenser la disruption létale du gène *tR4* et confère à la souche un caractère rouge. La souche YAL5 a été utilisée pour l'étude de l'aminoacylation *in vivo* de mutants de l'ARNt$^{Arg}{}_4$.

# 5. Plasmides

**pCR®2.1 :** contient une origine de réplication, les gènes de résistance à l'ampicilline (Amp$^R$) et à la kanamycine (Kan$^R$), le promoteur *lac* et le gène *lacZ*α. Il est fourni linéairisé au niveau du gène *lacZ*α avec aux extrémités 5' et 3' des T sortants (kit « TA Cloning » Invitrogen). Il permet le clonage des produits de PCR qui possèdent à l'extrémité 3' un A sortant. L'intérêt de ce vecteur est d'identifier les clones recombinants par le système de sélection blanc/bleu. Ce système est basé sur l'interruption ou non du gène *lacZ*α codant pour la ß-galactosidase par l'insertion ou non du produit de PCR.

**pQE70 :** permet l'expression dans *E. coli* de protéines avec en C terminal une queue poly-His, qui permettra leur rétention sur colonne Ni-NTA. Le vecteur possède le gène de résistance à l'ampicilline (Amp$^R$).

**pTFMa :** dérivé du vecteur pUC18 (Yanisch-Perron et al., 1985) comprenant une origine de réplication, le gène de résistance à l'ampicilline (Amp$^R$), et le gène de l'ARNt$^{Asp}$ de levure entre les sites *BamH*I et *Hind*III.

**pTSM$_3^{Arg}$ :** ce plasmide résulte du plasmide pTFMa pour lequel le gène de l'ARNt$^{Asp}$ de levure a été excisé et remplacé par celui de l'ARNt$^{Arg}{}_3$ de levure.

**pRS314 :** les vecteurs de la série pRS (Sikorski & Hieter, 1989) sont des plasmides navettes capables de se répliquer dans *E. coli* et la levure grâce à leurs 2 origines de réplication. Ils contiennent un gène de résistance à l'ampicilline pour la sélection dans *E. coli* et un marqueur pour le maintien dans la levure, *TRP1* pour le vecteur de type pRS314. Le gène sauvage *tR4* correspondant à l'ARNt$^{Arg}_{4}$ a été cloné dans le site *Hind*III entouré de ses séquences flanquantes nécessaires à la transcription.

# B. Milieux de culture

Les souches d'*E. coli* sont cultivées à 37°C dans le milieu LB (Luria Bertani) contenant par litre 10 g de bactotryptone, 5 g d'extraits de levure et 10 g de NaCl. Les milieux gélosés renferment en plus 20 g de bactoagar. Le milieu est additionné, si nécessaire et après autoclavage, de 100 µg/ml d'ampicilline, 170 µg/ml de chloramphénicol, ou de 50 µg/ml de kanamycine.

Les souches de *S. cerevisiae* sont cultivées à 30°C en milieu complet ou en milieu minimum complémenté en différents métabolites. Un litre de milieu complet YPG contient 10 g de bactopeptone, 10 g d'extraits de levure et 20 g de glucose. Le milieu minimum YNB contient 6,7 g de « yeast nitrogen base » et de 20 g de glucose par litre ; ce milieu est complémenté, après stérilisation, par les différents acides aminés pour lesquels la souche est auxotrophe. Ici, le milieu YNB a été complémenté par 20 µg/ml d'uracile, 30 µg/ml de lysine, 60 µg/ml de leucine et une concentration minimale d'adénine 2 µg/ml (permettant l'observation de la coloration rouge des colonies $ade_2^-$).

Les milieux de culture sont stérilisés par autoclavage.

# C. Techniques courantes

## 1. Préparation de bactéries compétentes et transformation

### Méthode au $CaCl_2$

La technique consiste à exposer des cellules en phase exponentielle de croissance au $CaCl_2$ qui perméabilise les membranes et permet ainsi l'entrée de l'ADN dans les bactéries.
Une aliquote de cellules *E. coli* est mise en préculture dans 5 ml de milieu LB, pendant la nuit à 37°C, sous agitation. 250 µl de la préculture sont transférés dans 250 ml de milieu LB. Les cellules sont mises en culture à 37°C et sont récoltées en phase exponentielle de croissance ($A_{700nm}$ ~0,7) par centrifugation 10 min à 4000 g et à 4°C. Le culot est repris dans 12,5 ml $CaCl_2$ 50 mM froid et incubé au moins 20 min dans la glace. Après centrifugation, les bactéries compétentes obtenues sont reprises dans 2 ml $CaCl_2$ 50 mM, 2 ml glycérol 80 %, aliquotées par volume de 200 µl et soit utilisées directement soit conservées à –80°C.

La transformation des bactéries ainsi rendues compétentes est effectuée de la façon suivante : après incubation 1 h 30 de 200 µl de cellules en présence de 10 µl de milieu de ligation (voir §D.5.), les bactéries sont soumises à un choc thermique de 1 min à 42°C et laissées dans la glace 5 min. Le mélange est étalé sur boîte de Pétri contenant du milieu LB gélosé renfermant les antibiotiques adéquats.

### Préparation et transformation de bactéries électrocompétentes

Une aliquote de cellules *E. coli* est mise en préculture dans 5 ml de milieu LB, pendant la nuit à 37°C, sous agitation. 1 ml de la préculture sont transférés dans 200 ml de milieu LB. Le cellules, mises en culture à 37°C, sont récoltées en phase exponentielle de croissance ($A_{700nm}$ ~0,7) par centrifugation 10 min à 4000 g et à 4°C. Le culot de bactéries est repris dans 200 ml Hepes-NaOH 1 mM pH 7,0 froid, puis centrifugé 10 min à 4000 g. Le culot obtenu est dissous dans 200 ml du même tampon et à nouveau centrifugé. Ensuite le culot est repris dans 8 ml de glycérol 10 % froid et centrifugé. Les bactéries récoltées sont reprises dans 800 µl de glycérol 10 % froid, aliquotées par volume de 40 µl et conservées à –20°C.

40 µl de bactéries ainsi obtenues sont additionnées de 2 à 4 µl d'ADN dépourvu de tout sel. Le mélange est transféré dans une cuve à électroporation et soumis à un champ électrique (2,5 kV, 25 µF, 200 Ω). Celui-ci va perturber la membrane de la bactérie et ainsi permettre l'entrée de l'ADN. Immédiatement après, 90 µl de milieu LB sont ajoutés au milieu de transformation. Différentes dilutions sont étalées sur milieu gélosé contenant les antibiotiques adéquats.

# 2. Etude des acides nucléiques

## 2.1. Détermination de la concentration d'une solution d'acides nucléiques

Les acides nucléiques absorbent la lumière UV et présentent un pic d'absorption à 260 nm. La concentration d'une solution d'ADN est estimée par la relation : concentration (mg/ml) = 0,05 x $DO_{260}$, tandis que la concentration en ARN est estimée par la relation : concentration (mg/ml) = 0,04 x $DO_{260}$. Sur un spectre d'absorption réalisé entre 220 et 320 nm, une solution pure d'acide nucléique présente un maximum d'absorption à 256-260 nm et un rapport $A_{260}/A_{280}$ compris entre 1,8 et 2. Une solution contaminée par des protéines présentera un pic d'absorption à 280 nm.

## 2.2. Extraction phénolique

Cette étape permet de séparer les acides nucléiques des protéines. L'échantillon à traiter est additionné d'un volume de phénol, agité vigoureusement pendant 1 min et centrifugé 5 min à 10 000 g. La phase aqueuse est soigneusement prélevée, de façon à ne pas entraîner les protéines dénaturées dans la phase phénolique. Puis la phase aqueuse est lavée deux par addition de 2 volumes d'éther saturé en eau, agitée vigoureusement pendant 1 min et centrifugation 1 min à 10 000 g. Après élimination de la phase éthérée (en surface), les traces résiduelles d'éther sont évaporées. Les acides nucléiques contenus dans la phase aqueuse sont généralement précipités.

## 2.3. Précipitation des acides nucléiques

La solution d'acides nucléiques (ADN/ARN) à précipiter est additionnée de 0,1 volume d'acétate de sodium 3 M (pH 5,5-6.0) et de 3 volumes d'éthanol absolu 100% et placée à –20°C pendant 1 h ou 20 min à –80°C. Le culot est récupéré par centrifugation à 12 00 rpm à 4°C pendant 30 min, puis lavé avec de l'éthanol 70% et récupéré par une seconde étape de centrifugation dans les mêmes conditions. Après élimination de la phase aqueuse, le culot est séché sous vide et repris dans le tampon adéquat.

## 2.4. Amplification d'ADN plasmidique

L'ADN plasmidique est amplifié par culture de la souche bactérienne choisie dans 100 ml de milieu LB en présence de l'antibiotique adéquat. La purification de l'ADN plasmidique est réalisée à l'aide d'un kit commercialisé (Plasmid MAXI kit, Genomed), suivant le protocole fourni. De façon succincte, après lyse des cellules, et élimination des débris par centrifugations, l'ADN plasmidique (présent dans le surnageant) est retenu sur colonne échangeuse d'anions. L'ADN est ensuite élué et précipité par ajout de 0,7 volumes d'isopropanol.

## 2.5. Migration des acides nucléiques sur gel

### *Electrophorèse sur gel d'agarose*

Le gel d'agarose contient habituellement 1 % d'agarose (1 g pour 100 ml), mélangé à chaud dans du tampon TBE 1x (0,1 M Tris-base, 0,1 M acide borique, 2 mM EDTA, pH 8,3). Les échantillons d'ADN, après addition d'un volume de dépôt (80 % (v/v) formamide, 10 mM EDTA pH 8,0, 0,025 % (p/v) xylène cyanol et bleu de bromophénol) sont déposés sur le gel et séparés par migration à 120 V dans une solution de TBE 1x. Les acides nucléiques sont visualisés sous UV après un bain dans une solution de bromure d'éthidium à 0,5 µg/ml, TBE 1x.

### *Electrophorèse sur gel de polyacrylamide*

L'électrophorèse sur gel de polyacrylamide est une méthode largement utilisée pour la purification et l'analyse d'ADN et d'ARN. Le gel de polyacrylamide est un copolymère constitué d'acrylamide et de N-N'-méthylène bis-acrylamide dont le rapport est de 19/1 respectivement. Le gel est constitué d'acrylamide/bis-acrylamide de 6 à 22 %, et de TBE 1x. Dans le cas des gels dénaturants, de l'urée est ajoutée à une concentration de 8 M. La polymérisation du gel est déclenchée par l'ajout de 1 % (v/v) de solution de peroxodisulfate d'ammonium à 5 % et de 0,1 % (v/v) de TEMED. L'épaisseur du gel varie de 0,4 mm pour les gels analytiques ou de séquençage à 3 mm pour les gels « préparatifs » (pour la purification de grande quantité d'ARN). Avant dépôt, les échantillons sont additionnés d'un volume d'une solution de dépôt (20 % (p/v) saccharose, 8 M urée, 0,025 % (p/v) xylène cyanol et bleu de bromophénol). L'électrophorèse est réalisée dans le tampon TBE 1x. La migration des échantillons s'effectue sur mini-gel (0,075x8,5x10,5 cm) à 150 V, sur gel de séquence ou gel de purification (0,075x30x40 cm) à 1 800 V et 55 mA, et sur gels préparatifs (0,2x30x40) à 650 V et 55 mA. La durée de migration varie en fonction de l'ADN ou de l'ARN.

## 2.6. Elution des ARN à partir d'un gel de polyacrylamide

Après migration sur gel de polyacrylamide, les ARN sont visualisés par ombrage aux UV sur un écran de fluorographie. Les transcrits radiomarqués sont visualisés par application d'un film autoradiographique. La bande de gel contenant la molécule de taille correcte est découpée et les ARN sont extraits soit de manière passive, soit par électroélution.

### *Elution passive*

Le fragment de gel dans lequel se trouve l'ARN est incubé 3 h dans 500 µl de tampon d'élution (0,5 M acétate d'ammonium, 10 mM acétate de magnésium, 0,1 mM $Na_2EDTA$ et 0,1 % SDS), sous agitation et à 4°C. La phase aqueuse est récupérée et l'ARN est précipité.

### *Electroélution*

Cette méthode est utilisée lorsque la bande de gel contenant l'ARN est épaisse, comme c'est le cas des gels « préparatifs ». Les bandes sont placées dans un appareil à électroéluer

« Biotrap » (Schleicher & Schuell) muni de membranes semi-perméables dans le tampon TBE 1x stérile. Sous l'influence du courant électrique, les ARN sont extraits du gel et récupérés dans la chambre anodique après 1 h 30 de migration à 150 V, 4°C Cette opération est réalisée deux fois. Les échantillons sont ensuite précipités pour être concentrés.

## 2.7. Northern blot

Les northern blot ont été effectués en utilisant un kit commercialisé par Ambion® (« NorthernMax® »), selon le protocole fourni. L'ensemble des solutions utilisées (bleu de dépôt, tampon dénaturant pour le gel d'agarose, tampon de migration, solutions d'hybridation et de lavages) sont fournies dans le kit.

### *Séparation des ARN sur gel dénaturant agarose*

30 µg d'ARN total humain de muscle squelettique (commercialisé par Panomics), aditionnés de 20 µl de bleu de dépôt contenant 10 µg/ml BET, ont été séparés sur gel dénaturant (formaldéhyde) 1 % (p/v) agarose de 6 mm d'épaisseur et 20 cm de long, dans du tampon MOPS[1] 1x. La migration est réalisée à 5 V/cm de gel de l'anode à la cathode. La migration des ARN est suivie par visualisation aux UV des ARNr.

### *Transfert sur membrane de nylon*

Les ARN sont transférés du gel d'agarose à une membrane de nylon chargée positivement de type Hybond-XL (Amersham Pharmacia Biosciences) selon le schéma Figure 29. Le flux de tampon de transfert (commercial) va entraîner par capillarité la migration des ARN vers la membrane déposée juste en dessous du gel. La durée du transfert est directement proportionnelle à l'épaisseur du gel (15-20 min/mm). Le transfert des ARN peut être contrôlée par la visualisation du gel aux UV (disparition du signal correspondant aux ARNr). Les ARN sont ensuite fixés covalemment sur la membrane par une exposition sous UV de 1 min.

[1] MOPS : acide 3-(*N*-morpholino) propanesulfonic.

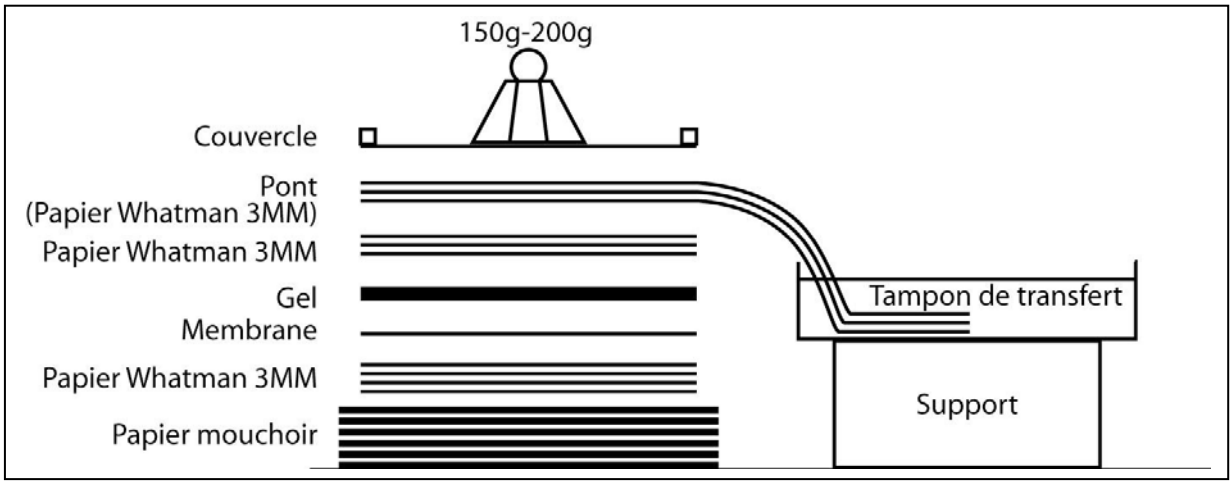



**Figure 29. Assemblage permettant le transfert d'acides nucléiques sur membrane de nylon par capillarité.**

### *Synthèse des sondes radiomarquées*

Les sondes ont été synthétisées en utilisant le kit « Nonaprimer™ DNA abelling Kit I » (Qbiogen), selon le protocol fourni. Cette technique se base sur la synthèse, en présence d'un des dNTP radioactif ([$\alpha^{32}$P]), de courts fragments aléatoires d'ADN (9-mer) en utilisant comme matrice la séquence du gène cible obtenu par PCR et purifié sur gel au préalable. Ceci permet, lors d'expériences de northern blot, de cibler toute la séquence de l'ARNm d'intérêt et ainsi d'amplifier le signal. Brièvement, les sondes sont synthétisées en utilisant 100 ng de matrice, adsorbées sur une résine permettant l'élimination des dNTP en surplus par lavage, et éluées.

### *Hybridation des sondes radioactives et révélation*

L'étape de préhybridation de la membrane se fait au moins 30 min à 42°C dans 10 ml de solution de préhybridation (préchauffée à 68°C) par $cm^2$ de membrane. L'hybridation se fait à la même température pendant 12 h, dans la solution de préhybridation et 200 µl de sonde radioactive (contenant $10^7$ cpm et 4 µl EDTA 0,5 M).

Après hybridation, la membrane est lavée 2 fois pendant 5 min à 37°C dans 20 ml d'une solution de lavage à faible stringence pour 100 $cm^2$ de membrane, puis 2 fois 15 min à 42°C dans le même volume d'une solution à forte stringence. La membrane est emballée dans un plastique scellé. Le signal est observé par autoradiographie de la membrane. Si la membrane doit être réhybridée, il est impératif qu'elle soit maintenue humide dans un sac en

plastique scellé ; un traitement de 10 min à 90°C dans une solution 1 % SDS permet de déshybrider totalement les sondes radioactives de la membrane.

# 3. Etude des protéines

## 3.1. Détermination de la concentration d'une solution de protéines

Les protéines peuvent être quantifiées par dosage colorimétrique selon la méthode de Bradford (réactif Bio-Rad) en utilisant l'albumine de sérum bovin (BSA) comme protéine étalon. La concentration en protéines pure peut également être calculée en fonction du coefficient d'extinction molaire de celle-ci, selon la relation de Beer-Lambert : c (mg/ml) = $DO_{280}$ x $E^{-1}$ x $l^{-1}$, où E est le coefficient d'extinction molaire ($M^{-1}.cm^{-1}$) et l la distance optique (1 cm). Le coefficient d'extinction molaire de l'aspartyl-ARNt synthétase mt humaine délétée des 47 premiers acides aminées est de 43 540 $M^{-1}.cm^{-1}$.

## 3.2. Précipitation des protéines

La solution de protéines à précipiter est additionnée de 10 volumes d'acétone 100 % et placée à –20°C pendant la nuit. Le culot est récupéré par centrifugation à 12 00 rpm à 4°C pendant 30 min, puis lavé avec de l'acétone et récupéré par une seconde étape de centrifugation dans les mêmes conditions. Après élimination de la phase aqueuse, le culot est séché sous vide et repris dans le tampon adéquat.

## 3.3. Electrophorèse en conditions dénaturantes (SDS-PAGE)

Cette technique permet d'évaluer la pureté et le poids moléculaire apparent des protéines en solution. Le gel de polyacrylamide est constitué (i) d'un gel de concentration à faible maillage qui permet la concentration des protéines de l'échantillon et (ii) d'un gel de résolution qui permet ensuite la séparation des protéines les unes des autres en fonction de

leur poids moléculaire apparent. Les deux parties du gel sont constituées d'acrylamide et de N-N'-méthylène bis-acrylamide dont le rapport est de 37,5/1. Le gel de concentration est constitué de 5 % d'acrylamide/bis-acrylamide, 0,1 % de SDS et de 125 mM de tampon Tris-HCl pH 6,8. Le gel de résolution est constitué de 6 à 15 % d'acrylamide/bis-acrylamide, 0,1 % de SDS et de 375 mM de tampon Tris-HCl pH 8,8. La migration des protéines s'effectue dans un tampon formé de 25 mM de Tris-base, 140 mM de Glycine et 0,1 % SDS. Avant dépôt sur le gel, les échantillons sont additionnés d'un volume de solution de dépôt (50 mM Tris-HCl pH 6,8, 100 mM DTE, 2 % SDS, 0,1 % bleu de bromophénol, 20 % glycérol) et chauffés 5 min à 100°C. Le voltage appliqué pour la migration est d'abord de 80 V pendant 30 min puis de 150 V. Les protéines sont visualisées par coloration pendant 2 h au bleu de Coomassie (0,25 % (p/v) bleu de Coomassie R250, 30 % méthanol, 10 % acide acétique), puis décoloration jusqu'à visualisation des bandes par une solution contenant 30 % d'éthanol, 10 % d'acide acétique.

# D. Préparation d'ARNt par clonage et transcription *in vitro*

## *1. Principe*

Les gènes synthétiques comportent le promoteur de la T7 ARN polymérase ADN dépendante suivi (i) d'une séquence ribozyme, (ii) du gène de l'ARNt et (iii) du site de restriction *BstN*I ou *Nsi*I (pour les séquences qui ont un site *BstN*I interne). Ces sites, après linéarisation du plasmide par l'enzyme correspondante, introduisent un arrêt de transcription et libèrent des fragments dont la séquence 5' du brin qui sera transcrit (GGT) code pour l'extrémité 3' (CCA) de l'ARNt. La transcription des gènes synthétiques permet la production d'ARN (nommés « transzymes »), qui après coupure auto-catalytique libèrent l'ARNt d'une part et le ribozyme d'autre part (Fechter et al., 1998).

## 2. Construction d'un gène synthétique

Les gènes synthétiques insérés dans le plasmide sont issus de l'hybridation d'un ensemble de 9 oligonucléotides chevauchants (synthétisés par Thermohybaid, Genset ou Sigma genosys). Ces fragments de synthèse possèdent une extrémité 5'OH qu'il est indispensable de phosphoryler pour leur ligation par la T4 ADN ligase.

La phosphorylation se fait dans le milieu réactionnel de 50 µl suivant : 50 mM Tris-HCl pH 7,5 , 10 mM $MgCl_2$, 5 mM DTE, 100 µM ATP, 400 pmoles d'oligonucléotide, 10 U de T4 polynucléotide kinase (New England BioLabs). La réaction se déroule à 37°C pendant 15 min puis est arrêtée par inactivation de l'enzyme par chauffage à 65°C pendant 10 min.

L'hybridation se fait dans un volume de 40 µl, dans les conditions suivantes : 16 pmoles de chacun des oligonucléotides phosphorylés, 50 mM Tris-HCl pH 7,5, 10 mM $MgCl_2$, 5 mM DTE. Une première incubation de 10 min à 80°C permet la dénaturation des oligonucléotides, puis l'étape d'hybridation se déroule durant le refroidissement lent du milieu jusqu'à température ambiante (3 à 4 h).

## 3. Mutagénèse dirigée

Les gènes synthétiques peuvent être mutés soit par échange de cassette ou par PCR. La technique d'échange de cassette consiste à reconstruire le gène synthétique *de novo* (voir paragraphe précédent) en échangeant les oligonucléotides de séquence sauvage par ceux comportant la ou les mutations. L'intérêt est de pouvoir réaliser des mutations multiples tout le long du gène en une seule étape.

La technique de mutagénèse dirigée par PCR requiert l'utilisation du kit « QuickChange™ Site-Directed Mutagenesis » (Stratagene). Cette technique est basée sur l'élongation de deux oligonucléotides d'environ 30 mers complémentaires portant la mutation au milieu de leurs séquences, et s'hybridant de part et d'autre des brins parentaux. La PCR est réalisée dans un volume de 50 µl dans le tampon fourni et en présence de 100 ng de matrice double brin circulaire, 200 ng de chacun des oligonucléotides et de 2,5 U de Pfu Turbo® ADN polymérase. La réaction est précédée par 30 sec à 95°C, puis se déroule durant 30 cycles de 30 sec à 95°C, 2 min à 47-55°C (selon la température de fusion des

oligonucléotides), 2 min par kilobases d'ADN à répliquer à 72°C, et se finit par 10 min à 72°C. Les brins parentaux sont ensuite digérès par ajout de 20 U de *Dpn*I directement dans le milieu de PCR, pendant 1h à 37°C. Après précipitation, l'ADN est repris dans 50µl d'$H_2O$ ultrapure. 40 µl de cellules *E. coli* compétentes sont transformées avec 2 µl d'ADN par électroporation (25 µF, 2,5 kV, 200 Ω) puis des dilutions sont étalées sur milieu LB gélosé comprenant l'antibiotique adéquat.

## 4. Clonage dans pTFMa

La linéarisation du plasmide pTFMa par les enzymes de restriction *BamH*I *et Hind*III est réalisée à 37°C pendant 6 h dans le tampon commercial NEB2 (New England Biolabs). Le volume réactionnel de 150 µl comporte 50 µg de plasmide, 50 U de chacune des enzymes (New England BioLabs), 1,5 µg BSA. Le vecteur est purifié par extraction phénolique puis est précipité. Le culot de centrifugation est lavé à l'éthanol 70 %, séché et repris dans le tampon de déphosphorylation afin d'obtenir une solution à 1 µg d'ADN/µl.

La déphosphorylation du vecteur linéarisé est effectuée dans 50 µl de milieu réactionnel contenant 2 U de phosphatase alcaline d'intestin de veau (CIP, Biolabs), 5 µl de tampon commerciale de la phosphatase alcaline 10x concentré. La réaction se fait pendant 30 min à 37°C et est arrêtée par extraction phénolique suivie d'une précipitation à l'éthanol. La déphosphorylation empêche la recircularisation ultérieure du vecteur.

L'insertion du gène synthétique dans le plasmide linéarisé et déphosphorylé se fait par ligation par la T4 ADN ligase (Qbiogene) soit pendant 16 h à 16°C, soit 15 min aux ultrasons (en présence de glace). Le milieu réactionnel de 17 µl est le suivant : 1 µg de plasmide linéarisé, 10 µl du milieu d'hybridation, 5 µl d'ATP 10 mM, 5 U de T4 ADN ligase.

40 µl de cellules *E. coli* TG1 rendues compétentes au $CaCl_2$ sont transformées avec 10 µl d'ADN, puis des dilutions sont étalées sur milieu LB gélosé comprenant l'antibiotique adéquat.

## 5. *Sélection de clones positifs*

Les clones positifs sont identifiés par séquençage, selon la méthode des didésoxyribonucléotides (Sanger et al., 1977). L'ADN plasmidique, extrait de cellules en culture par minipréparation (kit commercialisé QIAprep Spin, QIAGEN), est dénaturé en présence de 0,2 N NaOH, 0,2 mM EDTA pendant 10 min à température ambiante, puis précipité à l'éthanol. Le séquençage est réalisé selon le protocole d'utilisation de la Sequenase® 2.0 (United States Biochemicals) et les produits de séquençage radiomarqués ([$\alpha^{35}$S]dATP) sont séparés sur gel (0,075x30x40 $cm^3$) 8% acrylamide/bisacrylamide dénaturant (19/1), 8 M Urée, TBE 1x. La migration est faite à 1800 V, 55 mA pendant 3 h, puis le gel est séché et les séquences sont révélées par autoradiographie.

## 6. *Transcription* in vitro

L'ADN plasmidique, après son amplification et sa purification (voir §C.2.4.), est tout d'abord linéarisé par les enzymes de restriction *BstN*I ou *Nsi*I afin de créer un signal d'arrêt de transcription. Pour ce faire, 100 µg d'ADN sont digérés dans un volume de 250 µl en présence du tampon associé à l'enzyme (NEB2 pour *BstN*I et tampon *Nsi*I pour l'enzyme correspondante), 0,01 mg/ml de BSA et 125 U d'enzyme, pendant 2h à 60°C pour *BstN*I ou 12h à 37°C pour *Nsi*I. La réaction est arrêtée par extraction phénolique et l'ADN digéré est précipité à l'éthanol, puis repris dans de l'eau afin d'avoir une solution à 1µg/µl. Cet ADN sera utilisé comme matrice dans la réaction de transcription *in vitro* par la T7 ARN polymérase.

Le milieu réactionnel de transcription de 500 µl contenant 50 µg d'ADN matrice est composé de : 22 mM $MgCl_2$, 40 mM Tris-HCl pH 8,1 à 37°C, 0,01 % triton, 1 mM spermidine, 5 mM DTE, 4 mM de chacun des 4 NTP et 24,5 µg de T7 ARN polymérase. La réaction s'effectue pendant 4-5h à 37°C, puis est arrêtée par une extraction phénolique suivie d'une précipitation à l'éthanol. L'ajout de 5 mM GMP dans le milieu de transcription des ADNt clonés sans la séquence du ribozyme permet la production de transcrits avec des extrémités 5' monophosphate. La concentration de $MgCl_2$ est amenée à 30 mM dans le milieu

de transcription des ARNt produits sous forme de transzyme, afin d'optimaliser la coupure autocatalytique par le ribozyme (Fechter et al., 1998).

## 7. Purification de transcrits sur gel préparatifs

L'utilisation d'une construction transzyme nécessite de séparer l'ARNt du fragment ribozyme et de l'ARN transzyme éventuellement non clivé. De plus, la T7 ARN polymérase peut ajouter, de façon aléatoire, 1 ou 2 nucléotides en 3' des transcrits (Milligan et al., 1987). Il est donc nécessaire de purifier les transcrits à extrémité CCA-3' correcte des autres molécules. La purification se fait par la taille sur un gel préparatif (0,2x30x40 $cm^3$) 12% acrylamide/bisacrylamide dénaturant (19/1), 8 M Urée, TBE 1x. La migration se fait 16 h à 650V, 55 mA. La bande correspondant au transcrit de bonne taille est détectée par ombrage aux ultraviolets, puis découpée. L'ARN est ensuite récupéré par électroélution. L'éluat est alors précipité à l'éthanol, et le culot est repris dans 100 µl d'eau.

## 8. Analyse de l'extrémité 3' des ARNt transcrits

Cette étape permet de vérifier l'extrémité 3' acceptrice des ARNt transcrits. Les ARNt sont radiomarqués en 3', purifiés et hydrolysés. Les nucléotides obtenus sont séparés en deux dimensions sur plaque mince de chromatographie. L'autoradiographie permet l'observation des nucléotides radiomarqués. Le profil de migration des nucléotides par rapport à des références permet de déterminer la nature du nucléotide révélé.

**Marquage de l'extrémité 3' de l'ARNt et purification sur gel.** 50 pmol d'ARNt transcrits sont d'abord radiomarqués en 3' dans le milieu suivant de 20 µl : 10 µCi pCp[$\alpha^{32}$P], 20 U T4 ARN ligase, 10 % DMSO, 0,1 mM ATP, tampon T4 ARN ligase 1x. Le milieu est incubé la nuit à 4°C. Puis les ARNt sont purifiés sur un gel de polyacrylamide 8%/Urée 8 M. Ceci permet d'éliminer les nucléotides radioactifs non incorporés et les fragments de dégradation. Les bandes correspondant aux ARNt sont détectées par

radiographie puis sont excisées. Les ARNt sont récupérés par élution passive puis précipités et repris dans de l'eau.

**Hydrolyse totale des ARNt radiomarqués purifiés.** L'équivalent de 5 000 cpm d'ARNt radiomarqués (comptage Cerenkov) sont hydrolysés par 13,5 % pipéridine et en présence de 10 µg d'ARNt total d'*E. coli*, le tout dans un volume de 20 µl. La réaction est effectuée dans un capillaire pendant 45' à 95°C.

**Migration en 2 dimensions du mélange de nucléotides.** Le milieu hydrolysé est déposé µl par µl (séchage après chaque dépôt) sur le coin inférieur gauche d'une plaque de cellulose (TLC-Ready Foils, F1440, 20x20cm$^2$, Cellulose Avial). La migration en première dimension s'effectue en trempant le bord inférieur de la plaque de chromatographie, tenue à la verticale, dans un bain de solvant contenant 66 vol. acide butyrique, 1 vol. ammoniaque 25 %, 32 vol. eau stérile. Il faut veiller à ce que le niveau du solvant n'atteigne pas le point de dépôt, de façon à ne pas dissoudre l'échantillon dans le bain. La migration se fait par capillarité pendant 3 h à température ambiante. La migration suivant la deuxième dimension s'effectue en plaçant la plaque de chromatographie à la verticale dans un bain de solvant mais cette fois-ci en faisant tremper le bord gauche. Là aussi, il faut veiller à ne pas tremper le point de dépôt de l'échantillon dans le bain. Le sovant utilisé contient 68 vol. isopropanol, 17,6 vol. HCl 37 %, 13,4 vol. d'eau stérile et 1 vol. EDTA 100 mM. La migration est effectuée pendant 4 h à température ambiante. La plaque est séchée et les nucléotides sont révélés par autoradiographie. Les signaux sont quantifiés en utilisant Fuji Bioimager Bas2000.

# E. Clonage, surexpression et purification de l'AspRS mitochondriale humaine

## *1. Clonage du gène à partir d'une banque d'ADNc humains*

Le gène de l'AspRS mt humaine a été amplifié par PCR à partir d'une banque d'ADN complémentaires humains (généreusement fournie par Franck Martin, IBMC). Les oligonucléotides utilisés comme amorces sont longs de 31 mers (amorce 5') et 30 mers (amorce 3') et comprennent chacun le site de restriction *Bgl*II à leur extrémité. La réaction de PCR a été réalisée dans le volume total de 50 µl et dans le milieu suivant : 2 U Dynazyme EXT (Finnzymes), 1 µM de chaque oligonucléotide amorce, 50 ng matrice, 200 µM de chaque dNTP et 1x tampon Dynazyme EXT contenant 1,5 mM $Mg_2Cl$. Après 2 min à 95°C, 30 cycles de PCR ont été réalisés de la façon suivante : 30 sec à 95°C, 45 sec à 52°C et 2 min à 72°C. La réaction d'élongation a été finalisée par une étape supplémentaire de 10 min à 72°C. Le fragment d'ADN obtenu a été directement utilisé pour le clonage dans le vecteur pCR2.1 et la transformation dans la souche TOP10F', selon le protocole fourni dans le kit « TA-cloning » (Invitrogen). Les clones positifs ont été identifiés selon la sélection blanc/bleu. En effet, l'insertion du fragment de PCR dans le vecteur pCR2.1 interrompt le gène codant pour la β-galactosidase, enzyme capable de généré un produit bleu en présence d'IPTG (inducteur de l'enzyme) et du substrat analogue X-gal. Les plasmides des clones blancs ont été préparés par minipréparation et séquencés (service de séquençage de l'IBMP).

Une fois qu'un clone positif a été identifié, son ADN plasmidique est amplifié et purifié par maxipréparation (voir paragraphe C.2.4.). 16 µg d'ADN plasmidique sont digérés dans un volume de 20 µl en présence de 20 U *Bgl*II et dans le tampon NEB3 (New England Biolabs). Le fragment généré est purifié sur un gel contenant 1,2 % d'agarose à faible température de fusion. Après incubation dans un bain de Bromure d'Ethidium et révélation aux UV, le fragment est découpé et extrait du gel grâce au kit « QIAquick gel extraction » qui utilise les propriétés de faible température de fusion du gel (Qiagen). La ligation du fragment dans le vecteur d'expression pQE70 (Qiagen), préalablement digéré par *Bgl*II et

déphosphorylé selon la méthode décrite précédemment (D.4.), et la transformation de cellules TOP10 ont été effectuées selon le protocole décrit dans le paragraphe C.1.

## 2. Expression et purification

La souche TOP10 contenant le plasmide pQE70-mt-AspRS est mise en culture dans 200 ml de milieu LB renfermant 50 µg/ml ampicilline et 2 % glucose, la nuit à 37°C. Après dilution au 10$^{ème}$ dans 1800 ml du même milieu, les cellules ont été remises en culture à 37°C jusqu'à $A_{700nm}$ = 0,7, puis récupérées par 10 min de centrifugation à 4000 g et à 20°C, lavées dans du milieu LB frais et à nouveau centrifugées. Les cellules culottées sont reprises dans 2000 ml de milieu LB contenant 50 µg/ml ampicilline et 500 µM IPTG (permettant l'induction de l'expression de notre protéine), et incubées à 25°C pendant la nuit. Elles sont finalement récupérées par centrifugation 10 min à 4000 g, à 4°C.

Le culot de cellules obtenu est repris dans 45 ml de tampon A (50 mM $NaH_2PO_4$ pH 7,5, 300 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoéthanol) et les cellules sont brisées aux ultrasons dans la glace (120 variations/sec pendant 40 sec, 7 fois). Les débris cellulaires sont culottés par ultracentrifugation 45 min à 30 000 g et à 4°C. Le surnageant est ensuite chargé sur une colonne Ni-NTA d'un volume de 2 ml (Qiagen), préalablement équilibrée dans du tampon A. Les ions nickel de la résine permettent la rétention sélective de l'AspRS mt humaine qui contient une queue poly-His à son extrémité C-terminale. Après élimination de la majorité des contaminants par une étape de lavage (passage de 50 ml de tampon A), l'AspRS mt est éluée grâce à un gradient linéaire d'imidazole de 30 ml (réalisé à partir du mélange des tampons A et B (tampon A contenant 500 mM imidazole)). Tout au long de la purification, le débit est maintenu à 1 ml/min. Les effluents sont collectés par fractions de 2 ml et l'absorption à 280 nm est enregistrée (système Biologic Duoflow, BioRad). Le profil d'absorption permet de suivre le déroulement des différentes étapes, et en particulier l'élution de la protéine. Des aliquotes de 10 µl des fractions correspondant au pic d'élution sont déposés sur un gel dénaturant de protéines. L'AspRS mt correspond à une bande majoritaire (>95 %). Les fractions contenant la protéine sont regroupées et dialysées une nuit à 4°C contre le tampon de conservation 50 mM $KH_2PO_4/K_2HPO_4$ pH 7,5, 150 mM KCl, 1 mM EDTA, 10 mM β-mercaptoéthanol et 50 % glycérol. L'enzyme est conservée à –

20°C. Sa concentration est déterminée par test de Bradford et mesure de l'absorbance à 280 nm. Le coefficient d'extinction de l'AspRS recombinante calculé par le programme ProtParam d'Expasy est de 43 540 $M^{-1}.cm^{-1}$.

# F. Aminoacylation *in vitro*

## *1. Détermination des taux de charge et des constantes cinétiques*

La réaction d'aminoacylation, quantifiée *in vitro* en utilisant l'acide aminé radioactif, permet de déterminer 3 paramètres : le taux de charge et les constantes cinétiques $k_{cat}$ et $K_M$. Ces paramètres sont déterminés dans des conditions expérimentales distinctes : le premier en condition d'équilibre (atteinte d'un plateau de charge en excès d'enzyme), et les derniers en condition de vitesses initiales (accumulation du produit de façon linéaire en fonction du temps). Le taux (ou plateau) de charge représente l'équilibre entre la forme aminoacylée et non aminoacylée de l'ARNt. Cet équilibre dépend des vitesses (i) d'aminoacylation, (ii) de déacylations spontanée et enzymatique et (iii) de la réaction inverse (Bonnet & Ebel, 1972). La constante de Michaelis-Menten $K_M$ reflète l'inverse de l'affinité de l'aaRS pour l'ARNt alors que la constante catalytique $k_{cat}$ est un paramètre décrivant directement l'activité de l'enzyme et sa vitesse de réaction (Fersht, 1977). Le rapport $k_{cat}/K_M$ représente l'efficacité catalytique du système. Le rapport $k_{cat}/K_M$ (référence) / $k_{cat}/K_M$ (testé) est indicatif de la perte d'efficacité catalytique d'un ARNt par rapport à une référence. Les paramètres cinétiques sont déterminés à partir de la représentation en doubles inverses selon Lineweaver et Burk : $1/v = K_M/V_{max} \times 1/[S_O] - 1/V_{max}$ (où $V_{max}$ est la vitesse maximale de la réaction, v la vitesse initiale et $[S_0]$ la concentration initiale en ARNt). Pour certains substrats, les capacités d'aminoacylation sont extrêmement faibles, ce qui rend impossible la détermination directe des paramètres cinétiques. Dans ce cas, l'efficacité catalytique ($k_{cat}/K_M$) est estimée à partir de la représentation de v en fonction de $[S_0]$ à un temps long (permettant la réaction). La pente de la droite obtenue représente $Vmax/K_M$, c'est à dire $[E_0] \times k_{cat}/K_M$ (où $[E_0]$ est la concentration initiale en enzyme) (Schulman & Pelka, 1988).

Avant chaque réaction, l'ARNt est systématiquement renaturé dans l'eau, par un chauffage de 90 sec à 60°C, suivi d'un refroidissement lent de 6 min à température ambiante. La réaction d'aminoacylation est initiée par ajout de l'enzyme diluée. Des aliquotes sont ensuite prélevées à différents temps et déposées sur des rondelles de papier Wattman 3MM. Ces papiers sont immédiatement plongés pendant 10 min dans une solution d'acide trichloroacétique à 5 % froid (TCA, Roth), qui précipite les ARNt et en particulier l'aa-ARNt (l'acide aminé libre reste en solution). Les papiers sont lavés deux autres fois 10 min au TCA 5 %, puis déshydratés par rinçage à l'éthanol. Les rondelles séchées sont déposées dans des fioles de comptage contenant 2,5 ml de liquide de scintillation (Ecoscint 0$^{TM}$, National Diagnostics) et la radioactivité est quantifiée dans un compteur à scintillation (modèle LS6500, Beckman Coulter).

# *2. Conditons optimales d'aminoacylation*

## 2.1. Conditions optimales pour l'ArgRS de levure

Les conditions optimales sont les suivantes : 50 mM HEPES-NaOH pH 7,5, 30 mM KCl, 10 mM ATP, 15 mM $MgCl_2$, 2,5 mM glutathion, 50 µM [$^3$H]-arginine (74 GBq/mmol), une quantité adéquate en ARNt transcrit et en ArgRS (ces quantités varient en fonction des paramètres étudiés et du couple enzyme-substrat). Les tests sont réalisés dans un volume final de 100 µl et les aliquotes prélevées sont de 20 µl. Les concentrations en ArgRS et en transcrits ainsi que les temps d'incubation à 30°C varient selon le but du test. Le facteur d'extinction, correspondant à l'amortissement par le papier 3MM des rayonnements émis par le tritium de l'arginine radioactive, a été établi dans ce travail, et est de 2,3. Ainsi, l'activité spécifique de l'acide aminé fixé sur l'ARNt est 2,3 fois plus élevée que celle mesurée pour l'acide aminé libre.

### 2.2. Conditions optimales pour l'AspRS de levure

Les conditions optimales sont les suivantes : 50 mM HEPES-NaOH pH 7,5, 30 mM KCl, 5 mM ATP, 15 mM $MgCl_2$, 32 µM [$^3$H]-aspartate (208 GBq/mmol), et une quantité adéquate en ARNt transcrit et en AspRS (ces quantités varient en fonction des paramètres étudiés et du couple enzyme-substrat). Le volume final réactionnel est de 50 µl et le volume des aliquotes prélevées est de 10 µl. Le facteur d'extinction de l'aspartate tritié, établi précédemment au laboratoire, est de 3,6.

### 2.3. Conditions optimales pour l'AspRS mt humaine

Les conditions réactionnelles utilisées pour les tests effectués avec l'AspRS mt humaine sont celles mises au point précédemment par Bénédicte Sohm pour la LeuRS mt humaine (Sohm et al., 2003). Elles ont été validées pour l'AspRS mt humaine. Le milieu est le suivant : 50 mM Hepes-KOH pH 7,6, 25 mM KCl, 12 mM MgCl2, 2,5 mM ATP, 0,2 mg/ml BSA, 1 mM spermine, 32 µM [$^3$H]-aspartate (208 GBq/mmol), et une quantité adéquate d'ARNt transcrits et d'enzyme. Le volume final réactionnel est de 50 µl et le volume des aliquotes prélevé est de 10 µl.

## G. Etude de l'aminoacylation *in vivo* chez la levure

Les expériences *in vivo* réalisées dans la levure ont été menées avec Gilbert Eriani et Renaud Geslain (IBMC) et selon le protocole décrit (Geslain et al., 2003).

### *1. Construction du gène muté dans le vecteur pRS314*

Le plasmide pRS314 contenant le gène du variant $ARNt_4^{Arg/C38}$ (précédemment cloné par Renaud Geslain au laboratoire) a été converti en pRS314-$ARNt_4^{Arg/C38/G73}$ (contenant le

mutant $ARNt_4^{Arg/C38/G73}$) grâce à l'utilisation du kit « QuickChange™ Site-Directed Mutagenesis » (Stratagene) selon le protocole fourni et plus amplement décrit dans le paragraphe D.3.

## 2. Transformation de la levure, sélection blanc/rouge

50 ml de milieu liquide complet (YPD) est ensemencé par la souche YAL5 et mis en culture à 30°C jusqu'à $A_{700nm}$ de 0,7 à 0,9. Après centrifugation 2 min à 500 rpm, le culot de cellules est repris dans 1 ml LiAc 100 mM et placé 5 min à 30°C. Les cellules sont récupérées par centrifugation et reprises dans un volume de 350 µl contenant 34 % PEG, 100 mM LiAc, 100 µg d'ADN « carrier » simple brin dénaturé à 90°C, et 100 ng à 5 µg pRS314-$ARNt_4^{Arg/C38:G73}$. Les cellules sont incubées 20 min à 42°C, puis centrifugées et reprises dans 100 µl d'eau stérile. Des dilutions successives sont étalées sur milieu minimum sélectif (supplémenté par 20 µg/ml uracile, 30 µg/ml lysine, 60µg/ml leucine, et une concentration minimale de 2 µg/ml adénine). Après 72 h d'incubation à 30°C, les secteurs blancs des colonies sont isolés et à nouveau étalés sur milieu sélectif. Ces secteurs blancs correspondent aux cellules ayant perdu le plasmide sauveur pAL5 et ne possédant plus que le plasmide pRS314-$ARNt_4^{Arg/C38/G73}$. Ainsi le double mutant $ARNt_4^{Arg/C38/G73}$ est le seul $ARNt_4^{Arg}$ exprimé et est suffisant à la viabilité des cellules.

## 3. Analyse d'ARNt par northern blot en condition acide

### 3.1. Préparation des ARN cellulaires en conditions acides

250 ml de milieu sélectif sont ensemencés par la souche obtenue et cultivé à 30°C. Les cellules sont collectées en phase exponentielle ($A_{700nm}$=0,7) et en phase stationnaire ($A_{700nm}$=2). Les ARN sont extraits par mélange vigoureux des cellules culottées dans 15 ml de phénol acide (saturé par 0,3 M acétate de sodium et 10 mM $Na_2EDTA$, pH 5) et récupérés dans la phase aqueuse après centrifugation du mélange 10 min à 8 000 rpm. Cette étape est

réalisée 2 fois avant précipitation des acides nucléiques. Après centrifugation, le culot d'ARN est repris dans le tampon 0,3 M acétate de sodium, 10 mM $Na_2EDTA$, pH 5.

## 3.2. Séparation des ARN sur gel acide

Les ARN sont séparés par électrophorèse sur un gel de polyacrylamide de 1 mm d'épaisseur (6,5 % polyacrylamide, 8 M urée, 0,1 M acétate de sodium pH 5). L'échantillon déposé est composé de 30 µg d'ARN dans un volume de 15 µl additionné de 5 µl de bleu de dépôt (20% saccharose, 7 M urée et 0,05% de bleu de bromophénol et de xylène-cyanol). Les poches de charge ont une largeur de 1 cm. Un témoin déacylé, au préalable traité 30 min à 37°C dans un tampon alcalin Tris-HCl 1M pH8,8, est déposé en parallèle. La migration est réalisée à 18W en chambre froide pendant environ 15h, jusqu'à ce que le bleu de bromophénol ait migré au 3 quart de la plaque.

## 3.3. Transfert sur membrane de nylon

Un fragment de gel centré sur le bleu de xylène-cyanol est découpé et placé en sandwich entre une membrane en nylon de type Hybond-XL (Amersham Pharmacia Biosciences) et d'une feuille de plastique. Trois couches additionnelles de papier Wattman 3MM sont déposées sur la membrane de nylon. NB : si le gel de séparation n'est pas acide, la membrane et le gel doivent être équilibrés 10 min dans un tampon phosphate 25 mM pH 6,5. Les ARN sont ensuite transférés sous vide et fixés sur la membrane par chauffage à 80°C pendant 2 h. Immédiatement après transfert, le gel et la membrane sont séparés dans un bain d'eau déminéralisée. A ce stade, la membrane peut être stockée ou directement préhybridée.

## 3.4. Marquage des sondes en 5'

150 pmoles d'oligonucléotides sont marqués 2 h à 37°C dans un volume final de 30 µl contenant : 5 U T4 PNK, 1x tampon T4 PNK commercial, et 100 µCi [γ32P] ATP. Une fois l'enzyme inactivée par 10 min d'incubation à 65°C, l'ATP radioactif libre est éliminé du

milieu par gel filtration sur Sephadex G25 medium dans une seringue de 1 ml. La colonne est au préalable débarrassée de l'eau qu'elle contient par 5 min de centrifugation à 280 g. Le milieu de marquage est ensuite chargé dans la seringue et élué par une étape de centrifugation identique. Les oligonucléotides marqués exclus du gel sont élués alors que l'ATP libre reste inclus dans la matrice.

## 3.5. Hybridation des sondes radioactives et révélation

L'étape de préhybridation de la membrane se fait pendant un minimum de 2 h dans une ampoule contenant 50 ml d'une solution 5x SSPE[2], 0,5 % (p/v) SDS et 1 % (v/v) Denhardt. La température de préhybridation est ajustée à la température de fusion des oligonucléotides à hybrider, c'est à dire ~60°C. L'hybridation se fait à la même température, pendant 12 h et dans 15 ml de solution de préhybridation et de 15 µl de sonde radioactive. Pour chaque analyse, nous avons systématiquement préparé 3 types de sondes différentes : 2 sondes témoins, l'une dirigée contre l'ARN ribosomal 5S (120 nucléotides) dont l'expression est constante et qui sert à normaliser les signaux mesurés dans les pistes, l'autre contre l'$ARNt_4^{Leu}$ (84 nucléotides) qui constitue un contrôle de la qualité de la préparation acide. La troisième sonde est celle dirigée contre l'$ARNt_4^{Arg}$ (77 nucléotides). Les deux populations d'ARN (ARNr et ARNt) ont des tailles différentes et sont bien échelonnés après migration. La sonde de l'ARNr 5S est un oligonucléotide de 30 mers qui s'hybride sur sa cible de la position 60 à 90. Les sondes des ARNt sont des oligonucléotides de 20 mers s'appariant à l'extrêmité 3' de l'ARNt.

Après hybridation, la membrane est lavée 2 fois 15 min à 25°C dans une solution 2x SSPE et 0,1 % (p/v) SDS. Le signal est observé par autoradiographie de la membrane. Si la membrane doit être réhybridée, il est impératif qu'elle soit maintenue humide dans un sac en plastique scellé ; un traitement de 10 min à 90°C dans une solution 1 % SDS permet de déshybrider totalement les sondes radioactives de la membrane. Dans le cas, où la membrane n'est pas réutilisée, elle peut être séchée 15 min au four à 70°C.

[2] Une solution 20x SSPE contient 175,3 g NaCl, 27,6 g $NaH_2PO_4 \cdot H_2O$ et 7,4 g EDTA pour 1 l. Le pH est ajusté à 7,4 avec de la soude.

# H. Production d'anticorps et test Elisa

Nous avons produit des anticorps dirigés contre l'AspRS mt humaine obtenue d'après la méthode décrite en E.2. Le principe est de purifier la protéine sur gel afin de l'injecter au lapin. Plusieurs injections sont réalisées afin d'obtenir la réponse immunitaire souhaitée et les sérums de lapin sont récupérés 8 jours après chaque injection. La présence d'anticorps anti-AspRS est testée par ELISA (« Enzyme-Linked Immunosorbent Assay »).

## *1. Préparation de l'antigène*

L'AspRS obtenue selon la méthode décrite précédemment, subit une nouvelle étape de purification sur gel dénaturant de protéines. Après coloration 15 min au bleu de Coomassie et décoloration, la bande correspondant à l'AspRS est découpée. A ce stade, elle peut être conservée à -20°C. Une quantité de gel équivalent à 50 mg de protéine est broyé dans 500 µl d'une solution physiologique contenant 0,9 % NaCl. Un volume d'adjuvant de Freund (Sigma) complet (première injection) ou incomplet est ajouté. L'adjuvant de Freund est composé de débris de bactéries qui vont déclencher ou stimuler la réaction immunitaire. La préparation obtenue peut ainsi être directement injecté par voie intramusculaire au lapin.
Si la solution mère de protéines n'est pas suffisamment concentrée, l'échantillon peut être concentré par précipitation à l'acétone, à -20°C pendant la nuit. Après lavge et centrifugation, le culot de protéines est directement repris dans le tampon de dépôt.

## *2. Déroulement des injections*

Au total 7 injections séparées de 3 semaines ont été réalisées. A partir de la troisième injection, le sérum du lapin a été récupéré 8 jours après chaque injection. Au total, 6 sérums ont ainsi été récupérés et conservés à –20°C après ajout d'un volume de glycérol 100 %.

## 3. Test Elisa

Les tests ELISA ont eu 2 objectifs : (i) vérifier la présence d'anticorps anti-AspRS mt dans les différents sérums, et (ii) tester la présence d'anticorps dirigés contre la queue poly-His présentes à l'extrémité C-terminale de l'AspRS injectée.

Le test ELISA mis en œuvre a été réalisé en 5 étapes majeures : (i) adsorption sur la plaque de l'antigène (AspRS) à différentes dilutions (ou à concentration optimale si déterminée), (ii) saturation des sites aspécifiques, (iii) incubation avec les sérums contenant potentiellement les anticorps spécifiques (appelés anticorps primaire ou AcI), à différentes dilutions, (iv) incubation avec un anticorps secondaire (AcII) dirigé contre la région invariable de l'AcI (ici AcII de chèvre anti-IgG de lapin, Bio-rad) couplé à la peroxydase puis (v) révélation de l'activité enzymatique. Le produit libéré par la réaction donne une coloration jaune dont l'absorbance est mesurée à 450 nm.

**Adsorption de l'antigène sur la plaque.** 50µl d'antigène (par puits de la plaque ELISA), à différentes concentrations dans le tampon d'adsorption 15 mM $Na_2CO_3$, 55 mM $NaHCO_3$ et 0,02 % (v/v ?) azide, sont incubés la nuit à 4°C ou 1 h à 37°C. Le tampon d'adsorption peut être remplacé par du PBS[3] 1x ou de l'eau ultrapure. Les différentes concentrations testées d'antigène sont 1 µg/ml, 100 ng/ml, 10 ng/ml et 1 ng/ml. La concentration optimale est 1 µg/ml.

Notons que l'efficacité d'adsorption de la plaque ELISA a un rôle important et qu'elle est fonction de l'antigène testé. En effet, nous avons testé deux types de plaques (Maxisorp de Dutscher et les plaques Falcon référencées 353912) et l'adsorption n'est pas la même sur les deux plaques. Elle est nettement meilleure pour les plaques fournies par Falcon.

**Lavages (facultatifs).** L'étape d'adsorption peut être suivie des différents lavages successifs suivants : 3 fois avec le tampon 100 mM $Na_2HPO_4/KH_2PO_4$ pH 7, 500 mM NaCl, 0,05 % (v/v) Tween20, suivis de 3 fois avec de l'eau distillée.

[3] Une solution PBS 1x de 1 l contient 8 g NaCl, 0,2 g KCl, 1,44 g $Na_2HPO_4$ et 0,24 g $KH_2PO_4$. Le pH de la solution est ajusté à 7,4 avec une solution d'HCl.

**Saturation des sites aspécifiques.** A chaque puits de la plaque sont ajoutés 100 µl 1x PBS, 0,05 % (v/v) Tween20 et 3 % (p/v) lait en poudre. Ce tampon est nommé tampon de saturation. La plaque est incubée la nuit à 4°C ou 1 h à 37°C.

**Incubation de l'anticorps primaire.** L'anticorps primaire (pur ou contenu dans le sérum) est incubé dans 50 µl de tampon de saturation par puits, la nuit à 4°C ou 1 h à 37°C. De façon générale, le test ELISA permet de tester différentes dilutions d'anticorps. On réalise ainsi des dilutions en cascade directement sur la plaque ELISA, colonne par colonne. Les dilutions de sérum testées vont de 1/20 à 1/10 240. NB : toutes les dilutions qui ne sont pas réalisées directement sur la plaque doivent être effectuées dans des tubes à hémolyse en verre de façon à éviter l'adsorption des anticorps sur les parois du tube.

**Lavages.** Trois lavages successifs sont réalisés avec le tampon 1x PBS, 0,05 % (v/v) Tween20.

**Incubation de l'anticorps secondaire.** L'anticorps secondaire, dirigé contre l'anticorps primaire, est incubé dans 50 µl de tampon 1x PBS, 0,05 % (v/v) Tween20, par puits. Ici, nous avons utilisés des anticorps secondaires de chèvre dirigés contre la partie invariable des IgG de lapin, dilué 5000 fois par rapport à la concentration du stock commercial (Bio-rad), et couplé à la péroxidase. L'incubation est effectuée 4 h à 4°C ou 1 h à 37°C.

**Lavages.** Cette étape est la plus importante (après l'étape de saturation). Les puits sont lavés 3 fois avec le tampon 1x PBS, 0,05 % (v/v) Tween20, suivis de 3 lavages avec le tampon PBS.

**Incubation du substrat.** A chaque puits lavé sont ajoutés 50 µl de tampon substrat. Le tampon substrat est obtenu par le mélange de 9 ml tampon 70 mM Na2HPO4, 30 mM acide citrique, pH 5,0, 1 ml tétraméthyl benzidine (TMB[4]) et 4 µl $H_2O_2$ 30 %. La réaction se fait quelques minutes à température ambiante, puis est arrêtée par l'ajout de 50 µl HCl 1 N par puits.

[4] 100 ml de TMB contient 300 mg 3,3',5,5'-tétraméthyl benzidine, 72 ml DMSO 100 %, 18 ml glycérol 100 % et 10 ml tampon 70 mM $Na_2HPO_4$, 30 mM acide citrique, pH 5,0.

Au total, six combinaisons possibles de blancs ont été testées ($Ag^-/AcI^-/AcII^-$, $Ag^-/AcI^-$, $Ag^-/AcII^-$, $Ag^-$, $AcI^-$, $AcII^-$) et donnent toutes une valeur d'absorbance similaire.

# I. Essais de cristallisation

## 1. Généralités

Si la cristallographie des rayons X est devenue un outil de choix de la biologie structurale moderne, l'obtention d'un cristal de bonne qualité de la molécule étudiée en reste le facteur limitant. Un monocristal est un solide au sein duquel une molécule (ou un complexe supramoléculaire) est disposée de façon périodique dans les trois dimensions de l'espace. Plus le cristal est ordonné, meilleure sera la qualité de la diffraction d'un faisceau de rayons X et plus précis le modèle de structure de la molécule formant ce cristal, calculé à partir des données de diffraction.

La cristallisation est un phénomène soumis à un grand nombre de paramètres physico-chimiques et biologiques tels que la température, le pH, la nature de l'agent cristallisant et du tampon, ou la pureté des produits utilisés. La cristallogenèse des macromolécules biologiques essaie de comprendre le rôle de chacun de ces paramètres et se fixe pour but d'établir des règles afin de rendre la démarche scientifique plus rationnelle (Ducruix & Giegé, 1999).

## 2. Principe général de la cristallisation

La cristallisation consiste à faire passer les molécules d'un état de soluté à un état solide cristallin comme décrit dans le diagramme de solubilité (Figure 30).

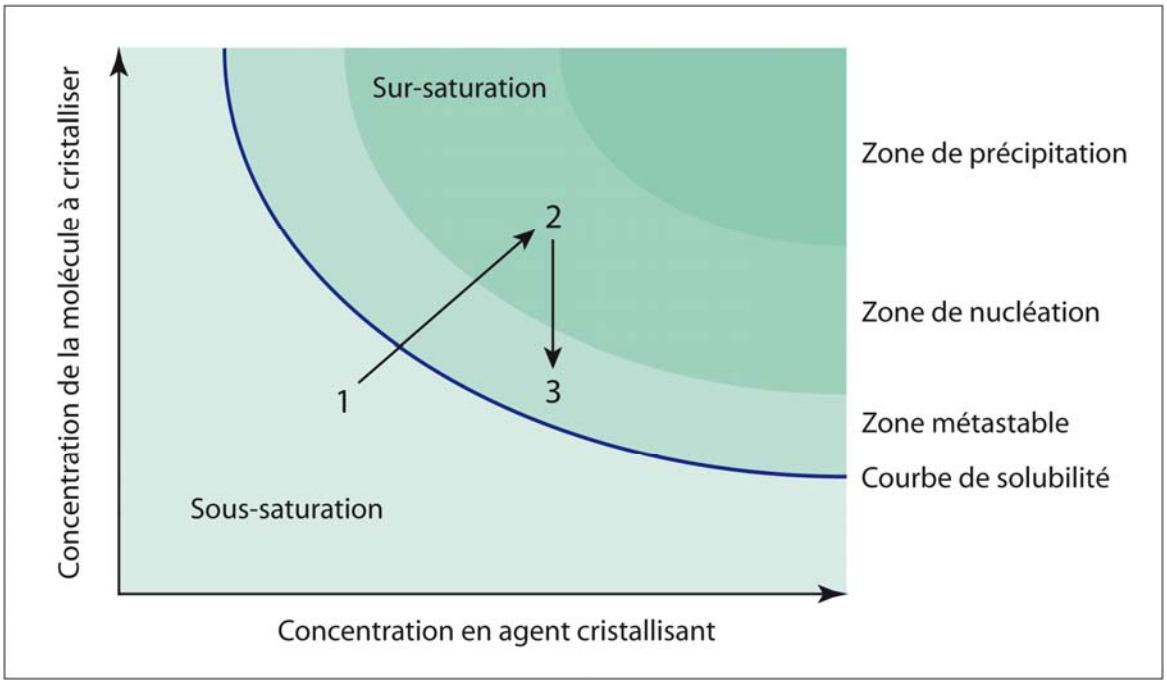



**Figure 30. Diagramme de solubilité théorique.** D'après (Riès-Kautt & Ducruix, 1992). Le diagramme décrit le comportement de la molécule à cristalliser en fonction de sa propre concentration (en ordonnée) et de celle de l'agent cristallisant (en abscisse).

Pour qu'une molécule s'exclue de la solution et forme un cristal, il est nécessaire qu'elle soit menée au delà de sa limite de solubilité (zone de sursaturation). Lorsqu'elle se sera enfoncée suffisamment loin dans cette zone « hors équilibre thermodynamique » ou zone métastable, elle aura acquis l'énergie d'activation nécessaire à la formation d'un noyau stable (amas ordonné de quelques molécules) à partir duquel le cristal pourra pousser (passage de 1 à 2 dans la Figure 30). Cette zone où la sursaturation élevée est compatible avec l'apparition d'un noyau stable est appelée zone de nucléation. A partir du noyau, le cristal va croître, nourri par les molécules en solution jusqu'à ce qu'il rejoigne un état d'équilibre, entre solide et solution, au niveau de la courbe de solubilité (passage de 2 à 3 dans la Figure 30). Si la molécule s'enfonce trop profondément dans la zone sursaturée, une nucléation brutale produira un précipité microcristallin. En revanche, si les conditions sont d'emblée défavorables à l'établissement d'un état sursaturé et donc à la nucléation, la molécule précipitera de façon désordonnée dès qu'elle aura franchi la courbe de solubilité. Il est donc

important de bien définir les conditions optimales (pH, température, agent cristallisant…) conduisant à la formation de cristaux de bonne qualité. Différentes techniques de cristallisation existent et sont couramment utilisées, nous avons choisi la technique de diffusion de vapeur, facile à mettre en oeuvre.

## 3. Principe de la cristallisation par diffusion de vapeur

Le principe de la cristallisation par diffusion de vapeur est le suivant : dans une enceinte close, une goutte de solution (200 nl à 50 µl) contenant la molécule à cristalliser ainsi que le tampon et l'agent cristallisant est mise à équilibrer contre un réservoir (volume de 70 µl à 1,5 ml) contenant le tampon et l'agent cristallisant à une concentration plus élevée que dans la goutte. L'équilibration des concentrations se fait par évaporation des composés volatiles (l'eau, par exemple) de la goutte vers le réservoir, jusqu'à ce que les tensions de vapeur au niveau des 2 compartiments soient identiques. Ainsi la goutte diminue de volume et ses composants se concentrent. Si les conditions sont favorables, la molécule est amenée à sursaturation (état hors équilibre) et cristallise.

Dans la pratique, les microméthodes, et aujourd'hui nanométhodes (grâce à la robotisation), de cristallogenèse permettent de travailler avec de très petits volumes (volume minimale de l'ordre de la nanogoutte) et de multiplier les essais sans consommer des quantités trop importantes de matériel (protéine et ARNt). Pour les essais de cristallisation, nous avons utilisé la technique de la goutte assise. Les gouttes de 6 µl sont préparées en mélangeant 3 µl de la solution contenant l'échantillon (protéine ou protéine/ARNt) à 3 µl de solution du réservoir apportant l'agent cristallisant. Ainsi, au départ les concentrations en agent cristallisant dans la goutte sont deux fois plus faibles que dans le réservoir. Du fait de l'équilibration des concentrations par diffusion de vapeur, la goutte se concentre deux fois en diminuant de volume.

L'approche généralement utilisée consite à tester un grand nombre de conditions au travers d'une matrice clairesemée, réunissant jusqu'à 96 combinaisons différentes de substances tampons, d'agents précipitants et d'additifs qui ont pour la plupart permis la cristallisation de molécules. Nous avons testé les matrices Wizard™ I et II (Emerald Biostructures) et une matrice de 24 conditions préparée par Bernard Lorber (IBMC), avec qui

les essais de cristallisation ont été conduits. La composition des différentes solutions de cristallisation des matrices est donnée dans les Tableaux 10, 11 et 12.

## 4. Préparation du matériel biologique

Les expériences de cristallisation nécessitent de grandes quantités de matériel. L'ARNt$^{Asp}$ mt humain transcrit *in vitro* et l'AspRS mt humaine ont été produits comme décrit précédemment mais dans des proportions plus importantes. Toutefois, suite à l'observation de la précipitation de l'enzyme lors de sa concentration, nous avons testé différentes conditions d'expression, de purification et de remplacement de tampon par dialyse. Tout ceci est détaillé dans la partie D du chapitre 2 entièrement dédiée aux essais de cristallisation.

| | Précipitant | Tampon (100 mM) | Sel (200mM) |
|---|---|---|---|
| 1 | 20% (w/v) PEG-8000 | CHES pH 9.5 | none |
| 2 | 10% (v/v) 2-propanol | HEPES pH 7.5 | NaCl |
| 3 | 15% (v/v) ethanol | CHES pH 9.5 | none |
| 4 | 35% (v/v) MPD | Imidazole pH 8.0 | MgCl2 |
| 5 | 30% (v/v) PEG-400 | CAPS pH 10.5 | none |
| 6 | 20% (w/v) PEG-3000 | Citrate pH 5.5 | none |
| 7 | 10% (w/v) PEG-8000 | MES pH 6.0 | Zn(OAc)2 |
| 8 | 2.0 M (NH4)2SO4 | Citrate pH 5.5 | none |
| 9 | 1.0 M (NH4)2HPO4 | Acetate pH 4.5 | none |
| 10 | 20% (w/v) PEG-2000 MME | Tris pH 7.0 | none |
| 11 | 20% (v/v) 1,4-butanediol | MES pH 6.0 | Li2SO4 |
| 12 | 20% (w/v) PEG-1000 | Imidazole pH 8.0 | Ca(OAc)2 |
| 13 | 1.26 M (NH4)2SO4 | Cacodylate pH 6.5 | none |
| 14 | 1.0 M sodium citrate | Cacodylate pH 6.5 | none |
| 15 | 10% (w/v) PEG-3000 | Imidazole pH 8.0 | Li2SO4 |
| 16 | 2.5 M NaCl | Na/K phosphate pH 6.2 | none |
| 17 | 30% (w/v) PEG-8000 | Acetate pH 4.5 | Li2SO4 |
| 18 | 1.0 M K/Na tartrate | Imidazole pH 8.0 | NaCl |
| 19 | 20% (w/v) PEG-1000 | Tris pH 7.0 | none |
| 20 | 0.4 M NaH2PO4/1.6 M K2HPO4 | Imidazole pH 8.0 | NaCl |
| 21 | 20% (w/v) PEG-8000 | HEPES pH 7.5 | none |
| 22 | 10% (v/v) 2-propanol | Tris pH 8.5 | none |
| 23 | 15% (v/v) ethanol | Imidazole pH 8.0 | MgCl2 |
| 24 | 35% (v/v) MPD | Tris pH 7.0 | NaCl |
| 25 | 30% (v/v) PEG-400 | Tris pH 8.5 | MgCl2 |
| 26 | 10% (w/v) PEG-3000 | CHES pH 9.5 | none |
| 27 | 1.2 M NaH2PO4/0.8 M K2HPO4 | CAPS pH 10.5 | Li2SO4 |
| 28 | 20% (w/v) PEG-3000 | HEPES pH 7.5 | NaCl |
| 29 | 10% (w/v) PEG-8000 | CHES pH 9.5 | NaCl |
| 30 | 1.26 M (NH4)2SO4 | Acetate pH 4.5 | NaCl |
| 31 | 20% (w/v) PEG-8000 | Phosphate-citrate pH 4.2 | NaCl |
| 32 | 10% (w/v) PEG-3000 | Na/K phosphate pH 6.2 | none |
| 33 | 2.0 M (NH4)2SO4 | CAPS pH 10.5 | Li2SO4 |
| 34 | 1.0 M (NH4)2HPO4 | Imidazole pH 8.0 | none |
| 35 | 20% (v/v) 1,4-butanediol | Acetate pH 4.5 | none |
| 36 | 1.0 M sodium citrate | Imidazole pH 8.0 | none |
| 37 | 2.5 M NaCl | Imidazole pH 8.0 | none |
| 38 | 1.0 M K/Na tartrate | CHES pH 9.5 | Li2SO4 |
| 39 | 20% (w/v) PEG-1000 | Phosphate-citrate pH 4.2 | Li2SO4 |
| 40 | 10% (v/v) 2-propanol | MES pH 6.0 | Ca(OAc)2 |
| 41 | 30% (w/v) PEG-3000 | CHES pH 9.5 | none |
| 42 | 15% (v/v) ethanol | Tris pH 7.0 | none |
| 43 | 35% (v/v) MPD | Na/K phosphate pH 6.2 | none |
| 44 | 30% (v/v) PEG-400 | Acetate pH 4.5 | Ca(OAc)2 |
| 45 | 20% (w/v) PEG-3000 | Acetate pH 4.5 | none |
| 46 | 10% (w/v) PEG-8000 | Imidazole pH 8.0 | Ca(OAc)2 |
| 47 | 1.26 M (NH4)2SO4 | Tris pH 8.5 | Li2SO4 |
| 48 | 20% (w/v) PEG-1000 | Acetate pH 4.5 | Zn(OAc)2 |

**Tableau 10. Composition des solutions de cristallisation de la matrice Wizard™ I** (Emerald Biostructures).

| | Précipitant | Tampon (100 mM) | Sel (200mM) |
|---|---|---|---|
| 1 | 10% (w/v) PEG-3000 | Acetate pH 4.5 | Zn(OAc)2 |
| 2 | 35% (v/v) MPD | MES pH 6.0 | Li2SO4 |
| 3 | 20% (w/v) PEG-8000 | Tris pH 8.5 | MgCl2 |
| 4 | 2.0 M (NH4)2SO4 | Cacodylate pH 6.5 | NaCl |
| 5 | 20% (v/v) 1,4-butanediol | HEPES pH 7.5 | NaCl |
| 6 | 10% (v/v) 2-propanol | Phosphate-citrate pH 4.2 | Li2SO4 |
| 7 | 30% (w/v) PEG-3000 | Tris pH 7.0 | NaCl |
| 8 | 10% (w/v) PEG-8000 | Na/K phosphate pH 6.2 | NaCl |
| 9 | 2.0 M (NH4)2SO4 | Phosphate-citrate pH 4.2 | none |
| 10 | 1.0 M (NH4)2HPO4 | Tris pH 8.5 | none |
| 11 | 10% (v/v) 2-propanol | Cacodylate pH 6.5 | Zn(OAc)2 |
| 12 | 30% (v/v) PEG-400 | Cacodylate pH 6.5 | Li2SO4 |
| 13 | 15% (v/v) ethanol | Citrate pH 5.5 | Li2SO4 |
| 14 | 20% (w/v) PEG-1000 | Na/K phosphate pH 6.2 | NaCl |
| 15 | 1.26 M (NH4)2SO4 | HEPES pH 7.5 | none |
| 16 | 1.0 M sodium citrate | CHES pH 9.5 | none |
| 17 | 2.5 M NaCl | Tris pH 7.0 | MgCl2 |
| 18 | 20% (w/v) PEG-3000 | Tris pH 7.0 | Ca(OAc)2 |
| 19 | 1.6 M NaH2PO4/0.4 M K2HPO4 | Phosphate-citrate pH 4.2 | none |
| 20 | 15% (v/v) ethanol | MES pH 6.0 | Zn(OAc)2 |
| 21 | 35% (v/v) MPD | Acetate pH 4.5 | none |
| 22 | 10% (v/v) 2-propanol | Imidazole pH 8.0 | none |
| 23 | 15% (v/v) ethanol | HEPES pH 7.5 | MgCl2 |
| 24 | 30% (w/v) PEG-8000 | Imidazole pH 8.0 | NaCl |
| 25 | 35% (v/v) MPD | HEPES pH 7.5 | NaCl |
| 26 | 30% (v/v) PEG-400 | CHES pH 9.5 | none |
| 27 | 10% (w/v) PEG-3000 | Cacodylate pH 6.5 | MgCl2 |
| 28 | 20% (w/v) PEG-8000 | MES pH 6.0 | Ca(OAc)2 |
| 29 | 1.26 M (NH4)2SO4 | CHES pH 9.5 | NaCl |
| 30 | 20% (v/v) 1,4-butanediol | Imidazole pH 8.0 | Zn(OAc)2 |
| 31 | 1.0 M sodium citrate | Tris pH 7.0 | NaCl |
| 32 | 20% (w/v) PEG-1000 | Tris pH 8.5 | none |
| 33 | 1.0 M (NH4)2HPO4 | Citrate pH 5.5 | NaCl |
| 34 | 10% (w/v) PEG-8000 | Imidazole pH 8.0 | none |
| 35 | 0.8 M NaH2PO4/1.2 M K2HPO4 | Acetate pH 4.5 | none |
| 36 | 10% (w/v) PEG-3000 | Phosphate-citrate pH 4.2 | NaCl |
| 37 | 1.0 M K/Na tartrate | Tris pH 7.0 | Li2SO4 |
| 38 | 2.5 M NaCl | Acetate pH 4.5 | Li2SO4 |
| 39 | 20% (w/v) PEG-8000 | CAPS pH 10.5 | NaCl |
| 40 | 20% (w/v) PEG-3000 | Imidazole pH 8.0 | Zn(OAc)2 |
| 41 | 2.0 M (NH4)2SO4 | Tris pH 7.0 | Li2SO4 |
| 42 | 30% (v/v) PEG-400 | HEPES pH 7.5 | NaCl |
| 43 | 10% (w/v) PEG-8000 | Tris pH 7.0 | MgCl2 |
| 44 | 20% (w/v) PEG-1000 | Cacodylate pH 6.5 | MgCl2 |
| 45 | 1.26 M (NH4)2SO4 | MES pH 6.0 | none |
| 46 | 1.0 M (NH4)2HPO4 | Imidazole pH 8.0 | NaCl |
| 47 | 2.5 M NaCl | Imidazole pH 8.0 | Zn(OAc)2 |
| 48 | 1.0 M K/Na tartrate | MES pH 6.0 | none |

**Tableau 11. Composition des solutions de cristallisation de la matrice Wizard™ II** (Emerald Biostructures).

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $Am_2SO_4$ | 1,8 M | PEG400 | 2% | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Acétate Na | pH 4,5 |
| 2 | $Am_2SO_4$ | 1,8 M | PEG400 | 2% | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Cacodylate Na | pH 6,5 |
| 3 | $Am_2SO_4$ | 1,8 M | PEG400 | 2% | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Tris HCl | pH 8,5 |
| 4 | $Am_2SO_4$ | 1,8 M | PEG400 | 2% | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Glycine HCl | pH 11 |
| 5 | PEG6000 | 12% | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Acétate Na | pH 4,5 |
| 6 | PEG6000 | 12% | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Cacodylate Na | pH 6,5 |
| 7 | PEG6000 | 12% | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Tris HCl | pH 8,5 |
| 8 | PEG6000 | 12% | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Glycine HCl | pH 11 |
| 9 | PEG8000 | 12% | NaCl | 1 M | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Acétate Na | pH 4,5 |
| 10 | PEG8000 | 12% | NaCl | 1 M | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Cacodylate Na | pH 6,5 |
| 11 | PEG8000 | 12% | NaCl | 1 M | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Tris HCl | pH 8,5 |
| 12 | PEG8000 | 12% | NaCl | 1 M | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Glycine HCl | pH 11 |
| 13 | MPD | 10% | Ethanol | 2% | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Acétate Na | pH 4,5 |
| 14 | MPD | 10% | Ethanol | 2% | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Cacodylate Na | pH 6,5 |
| 15 | MPD | 10% | Ethanol | 2% | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Cacodylate Na | pH 6,5 |
| 16 | MPD | 10% | Ethanol | 2% | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Glycine HCl | pH 11 |
| 17 | Jeffamin M-600 | 20% | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Cacodylate Na | pH 6,5 |
| 18 | PEGMME-2000 | 20% | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Cacodylate Na | pH 6,5 |
| 19 | PEG8000 | 12% | | | Glycérol | 15% | Tartrate Na | 0,1 M | Néant | pH 6 |
| 20 | Dioxane | 25% | | | Glycérol | 15% | Néant | | Néant | pH 8,5 |
| 21 | Malonate Na | 3,5 M | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | AcétateNa | pH 4,5 |
| 22 | Malonate Na | 3,5 M | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Tris HCl 0,36M | pH 3,5 |
| 23 | NaCl | 2,0 M | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Acétate Na | pH 4,5 |
| 24 | NaCl | 2,0 M | | | Glycérol | 15% | $LiSO_4$,$MgCl_2$ | 0,1M + 0,1M | Tris HCl | pH 8,5 |

**Tableau 13. Composition des solutions de cristallisation de la matrice 24 conditions** (préparées par Bernard Lorber).

# Références


**A**belson J, Trotta CR, Li H. 1998. tRNA splicing. *J Biol Chem 273*:12685-12688.

Adams KL, Palmer JD. 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol Phyl Evol 29*:380-395.

Ador L, Camasses A, Erbs P, Cavarelli J, Moras D, Gangloff J and Eriani G. 1999. Active site mapping of yeast aspartyl-tRNA synthetase by *in vivo* selection of enzyme mutations lethal for cell growth. *J Mol Biol 288*:231-242.

Agou F, Yang Y, Gesquière J-C, Waller J-P, Guittet E. 1995. Polyanion-induced α-helical structure of a synthetic 23-residue peptide representing the lysine-rich segment of the N-terminal extension of yeast cytoplasmic aspartyl-tRNA synthetase. *Biochemistry 34*:569-576.

Ahel I, Korencic D, Ibba M, Söll D. 2003. Trans-editing of mischarged tRNAs. *Proc Natl Acad Sci USA 100*:15422-15427.

Allen JF, Raven JA. 1996. Free-radical-induced mutation vs redox regulation: costs and benefits of genes in organelles. *J Mol Evolution 42*:482-492.

Almassy RJ, Janson CA, Hamlin R, Xuong NH, Eisenberg D. 1986. Novel subunit-subunit interactions in the structure of glutamine synthetase. *Nature 323*:304-309.

Ames BN, Tsang TH, Buck M, Christman MF. 1983. The leader mRNA of the histidine attenuator region resembles $tRNA^{His}$; possible general regulatory implications. *Proc Natl Acad Sci USA 80*:5240-5242.

An S, Musier-Forsyth K. 2004. Trans-editing of Cys-tRNAPro by Haemophilus influenzae YbaK protein. *J Biol Chem 279*:42359-42362.

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

Andrews R, Kubacka I, Chinnery P, Lightowlers R, Turnbull D, Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics 23*:143.

Arnez JG, Moras D. 1997. Structural and functional considerations of the aminoacylation reaction. *Trends Biochem Sci 22*:211-216.

Ataide SF, Jester BC, Devine KM, Ibba M. 2005. Stationary-phase expression and aminoacylation of a transfer-RNA-like small RNA. *EMBO Rep 6*:742-747.

Atkins JF, Gesteland RF. 1996. A case for *trans* translation. *Nature 379*:769-770.

Atkins JF, Gesteland RF. 2002. The 22nd amino acid. *Science 296*:1409-1410.

Attardi G, Montoya J. 1983. Analysis of human mitochondrial RNA. *Methods Enzymol 97*:435-469.

**B**aer RJ, Dubin DT. 1981. Methylated regions of hamster mitochondrial ribosomal RNA: structural and functional correlates. *Nucleic Acids Res 9*:323-337.

Barends S, Bink HH, Van den Worm SHE, Pleij CW, Kraal B. 2003. Entrapping ribosomes for viral translation: tRNA mimicry as a molecular trojan horse. *Cell 112*:123-129.

Bartel DP, Szostak JW. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science 261*:1411-1418.

Becker HD, Giegé R, Kern D. 1996. Identity of prokaryotic and eukaryotic $tRNA^{Asp}$ for aminoacylation by aspartyl-tRNA synthetase from *Thermus thermophilus. Biochemistry 35*:7447-7458.

Becker HD, Kern D. 1998. *Thermus thermophilus* - a link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc Natl Acad Sci USA 95*:12832-12837.

Becker HD, Reinbolt J, Kreutzer R, Giegé R, Kern D. 1997. Existence of two distinct aspartyl-tRNA synthetases in *Thermus thermophilus*. Structural and biochemical properties of the two enzymes. *Biochemistry 36*:8785-8797.

Beebe K, Ribas de Pouplana L, Schimmel P. 2003. Elucidation of tRNA-dependant editing by a class II tRNA synthetase and significance for cell viability. *EMBO J 22*:668-675.

Beuning PJ, Musier-Forsyth K. 1999. Transfer RNA recognition by aminoacyl-tRNA synthetases. *Biopolymers 52*:1-28.

Beuning PJ, Musier-Forsyth K. 2000. Hydrolytic editing by a class II aminoacyl-tRNA synthetase. *Proc Natl Acad Sci USA 97*:8916-8920.

Björk GR. 1995. Biosynthesis and function of modified nucleosides. In: Söll D, RajBhandary UL, eds. *tRNA: Structure, Biosynthesis, and Function.* Washington, D.C.: Am. Soc. Microbiol. Press. pp 165-205.

Blaise M, Becker HD, Keith G, Cambillau C, Lapointe J, Giegé R, Kern D. 2004. A minimalist glutamyl-tRNA synthetase dedicated to aminoacylation of the $tRNA^{Asp}$ QUC anticodon. *Nucleic Acid Res 32*:2768-2775.

Blight SK, Larue RC, Mahapatra A, Longstaff DG, Chang E, Zhao G, Kang PT, Green-Church KB, Chan MK, Krzycki JA. 2004. Direct charging of $tRNA_{CUA}$ with pyrrolysine *in vitro* and *in vivo*. *Nature 431*:333-335.

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

Bonnefond L, Giege R, Rudinger-Thirion J. 2005b. Evolution of the tRNA(Tyr)/TyrRS aminoacylation systems. *Biochimie 87*:873-883.

Bonnet J, Ebel J-P. 1972. Interpretation of incomplete reactions in tRNA aminoacylation. Aminoacylation of yeast $tRNA^{Val}$II with yeast valyl-tRNA synthetase. *Eur J Biochem 31*:335-344.

Börner GV, Mörl M, Janke A, Pääbo S. 1996. RNA editing changes the identity of a mitochondrial tRNA in marsupials. *EMBO J 15*:5949-5957.

Breitschopf K, Achsel T, Busch K, Gross HJ. 1995. Identity elements of human $tRNA^{Leu}$: Structural requirements for converting human $tRNA^{Ser}$ into a leucine acceptor *in vitro. Nucleic Acids Res 23*:3633-3637.

Briand C, Poterszman A, Eiler S, Webster G, Thierry J-C, Moras D. 2000. An intermediate step in the recognition of $tRNA^{Asp}$ by aspartyl-tRNA synthetase. *J Mol Biol 299*:1051-1060.

Brulé H, Holmes WM, Keith G, Giegé R, Florentz C. 1998. Effect of a mutation in the anticodon of human mitochondrial $tRNA^{Pro}$ on its post-transcriptional modification pattern. *Nucleic Acids Res 26*:537-543.

Buck CA, Nass MMK. 1969. Studies on mitochondrial tRNA from animal cells. I. A comparison of mitochondrial and cytoplasmic tRNA and aminoacyl-tRNA synthetases. *J Mol Biol 41*:67-82.

Buddha MR, Crane BR. 2005. Structure and activity of an aminoacyl-tRNA synthetase that charges tRNA with nitro-tryptophan. *Nat Struct Mol Biol 12*:274-275.

Bukau B, Horwich AL. 1998. The Hsp70 and Hsp60 chaperone machines. *Cell 92*:351-366.

Bullard J, Cai Y-C, Demeler B, Spremulli L. 1999. Expression and characterization of a human mitochondrial phenylalanyl-tRNA synthetase. *J Mol Biol 288*:567-577.

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

Bullock TL, Uter N, Nissan TA, Perona JJ. 2003. Amino acid discrimination by a class I aminoacyl-tRNA synthetase specified by negative determinants. *J Mol Biol 328*:395-408.

Burger G, Lang BF, Reith M, Gray MW. 1996. Genes encoding the same three subunits of respiratory complex II are present in the mitochondrial DNA of two phylogenetically distant eukaryotes. *Proc Natl Acad Sci U S A 93*:2328-2332.

Burke SA, Lo SL, Krzycki JA. 1998. Clustered genes encoding from methyltransferases of methanogenesis from monomethylamine. *Journal of Bacteriology 180*:3432-3440.

Caillet J, Nogueira T, Masquida B, Winter F, Graffe M, Dock-Brégeon AC, Torres-Larios A, Sankaranarayanan R, Westhof E, Ehresmann B, Ehresmann C, Romby P, Springer M. 2003. The modular structure of *Escherichia coli* threonyl-tRNA synthetase as both an enzyme and a regulator of gene expression. *Mol Microbiol 47*:961-974.

Carter Jr CW. 1993. Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. *Ann Rev Biochem 62*:715-748.

Carter Jr CW, Duax WL. 2002. Did tRNA synthetase classes arise on opposite strands of the same gene? *Mol Cell 10*:705-708.

Cavarelli J, Delagoutte B, Eriani G, Gangloff J, Moras D. 1998. L-Arginine recognition by yeast arginyl-tRNA synthetase. *EMBO 17*:5438-5448.

Cavarelli J, Rees B, Eriani G, Ruff M, Boeglin M, Gangloff J, Thierry J-C, Moras D. 1994. The active site of yeast aspartyl-tRNA synthetase: Structural and functional aspects of the aminoacylation reaction. *EMBO J 13*:327-337.

Cavarelli J, Rees B, Ruff M, Thierry J-C, Moras D. 1993. Yeast $tRNA^{Asp}$ recognition by its cognate class II aminoacyl-tRNA synthetase. *Nature 362*:181-184.

Charron C, Roy H, Blaise M, Giegé R, Kern D. 2003. Non-discriminating and discriminating aspartyl-tRNA synthetases differ in the anticodon-binding domain. *EMBO J 22*:1632-1643.

Cheong H-K, Park JY, Kim EH, Lee C, Kim S, Kim YG, Choi B-S, Cheong C. 2003. Structure of the N-terminal extension of human aspartyl-tRNa synthetase: implication for its biological function. *Intern J Biochem & Cell Biol 35*:1548-1557.

Chihade JW, Brown JR, Schimmel P, Ribas de Pouplana L. 2000. Origin of mitochondria in relation to evolutionary history of eukaryotic alanyl-tRNA synthetase. *Proc Natl Acad Sci USA 97*:12153-12157.

Chihade JW, Hayashibara K, Shiba K, Schimmel P. 1998. Strong selective pressure to use G:U to mark an RNA acceptor stem for alanine. *Biochemistry 37*:9193-9202.

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

Chimnaronk S, Jeppesen MG, Shimada N, Suzuki T, Nyborg J, Watanabe K. 2004. Crystallization and preliminart X-ray diffractioon study of mammalian mitochondrial seryl-tRNA synthetase. *Acta Cryst D 60*:1319-1322.

Choi H, Kay G, Schneider JA, Otten S, McClain WH. 2003. Recognition of acceptor-stem structure of $tRNA^{Asp}$ by Escherichia coli aspartyl-tRNA synthetase. *RNA 9*:386-393.

Chomyn A, Enriquez JA, Micol V, Fernandez-Silva P, Attardi G. 2000. The mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episode syndrome-associated human mitochondrial $tRNA^{Leu(UUR)}$ mutation causes aminoacylation

deficiency and concomitant reduced association of mRNA with ribosomes. *J Biol Chem 275*:19198-19209.
Clayton DA. 1982. Replication of animal mitochondrial DNA. *Cell 28*:693-705.
Cookson BD. 1998. The emergence of mupirocin resistance: a challenge to infection control and antibiotic prescribing practice. *J Antimicrob Chemother 41*:11-18.
Corti O, Hampe C, Koutnikova H, Darios F, Jacquier S, Prigent A, Robinson JC, Pradier L, Ruberg M, Mirande M, Hirsch E, Rooney T, Fournier A, Brice A. 2003. The p38 subunit of the aminoacyl-tRNA synthetase complex is a Parkin substrate: linking protein biosynthesis and neurodegeneration. *Hum Mol Genet 12*:1427-1437.
Crick FHC. 1955. On degenerate templates and the adaptor hypothesis: A note for the RNA tie club; cité par M. B. Hoagland (1960). In: Chargaff E, Davidson JN, eds. *The Nucleic Acids*. New York: Academic Press. pp 349.
Curnow AW, Ibba M, Söll D. 1996. tRNA-dependent asparagine formation. *Nature 382*:589-590.
Cusack S. 1997. Aminoacyl-tRNA synthetases. *Curr Opin Struct Biol 7*:881-889.
Cusack S, Berthet-Colominas C, Härtlein M, Nassar N, Leberman R. 1990. A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.5 A. *Nature 347*:249-255.

**D**'Errico I, Gadaleta G, Saccone C. 2004. Pseudogenes in metazoa: origin and features. *Brief Funct Genomic Proteomic 3*:157-167.
Dalluge JJ, Hashizume T, Sopchik AE, McCloskey JA, Davis DR. 1996. Conformational flexibility in RNA: the role of dihydrouridine. *Nucleic Acids Res 24*:1073-1079.
Dardel F, Panvert M, Fayat G. 1990. Transcription and regulation of expression of the *Escherichia coli* methionyl-tRNA synthetase gene. *Mol Gen Genet 223*:121-133.
de Duve C. 1988. The second genetic code. *Nature 333*:117-118.
Degoul F, Brulé H, Cepanec C, Helm M, Marsac C, Leroux J-P, Giegé R, Florentz C. 1998. Isoleucylation properties of native human mitochondrial $tRNA^{Ile}$ and $tRNA^{Ile}$ transcripts. Implications for cardiomyopathy-related point mutations (4269, 4317) in the $tRNA^{Ile}$ gene. *Hum Mol Gen 7*:347-354.
Deinert K, Fasiolo F, Hurt EC, Simos G. 2001. Arc1p organizes the yeast aminoacyl-tRNA synthetase complex and stabilizes its interaction with the cognate tRNAs. *J Biol Chem 276*:6000-6008.
Delagoutte B, Moras D, Cavarelli J. 2000. tRNA aminoacylation by arginyl-tRNA synthetase: induced conformations during substrates binding. *EMBO J 19*:5599-5610.
DeLano WL. 2002. The PyMOL molecular graphics system. San Carlos, CA, USA: Delano Scientific (http://www.pymol.org).
Delarue M, Poterszman A, Nikonov S, Garber M, Moras D, Thierry J-C. 1994. Crystal structure of a prokaryotic aspartyl-tRNA synthetase. *EMBO J 13*:3219-3229.
Deutscher MP. 1995. tRNA processing nucleases. In: Söll D, RajBhandary UL, eds. *tRNA: Structure, Biosynthesis, and Function.* Washington D.C.: ASM Press. pp 51-65.
Di Giulio M. 1999. The non-monophyletic origin of the tRNA molecule. *J Theor Biol 197*:403-414.
Di Giulio M. 2004. The origin of the tRNA molecule: implications for the origin of protein synthesis. *J Theor Biol 226*:89-93.
DiMauro S, Moraes CT. 1993. Mitochondrial encephalomyopathies. *Arch Neurol 50*:1197-1208.
Dimauro S, Schon E. 2001. Mitochondrial DNA Mutations in Human Disease. *American Journal of Medical Genetics 106*:18-26.

Dirheimer G, Keith G, Dumas P, Westhof E. 1995. Primary, secondary and tertiary structures of tRNAs. In: Söll D, RajBhandary UL, eds. *tRNA: Structure, Biosynthesis, and Function.* Washington, DC: Am. Soc. Microbiol. Press. pp 93-126.

Dock-Bregeon A-C, Westhof E, Giegé R, Moras D. 1989. Solution structure of a tRNA with a large variable region: Yeast $tRNA^{Ser}$. *J Mol Biol 206*:707-722.

Dock-Bregeon AC, Rees B, Torres-Larios A, Bey G, Caillet J, Moras D. 2004. Achieving error-free translation; the mechanism of proofreading of threonyl-tRNA synthetase at atomic resolution. *Mol Cell 16*:375-386.

Dong G, Nowakowsky J, Hoffman DW. 2002. Structure of small protein B: the protein component of the tmRNA-SmpB system for ribosome rescue. *EMBO J 12*:1845-1854.

Doolittle WF. 2000. The nature of the universal ancestor and the evolution of the proteome. *Current Opinion in Structural Biology 10*:355-358.

Doolittle WF, Brown JR. 1994. Tempo, mode, the progenote, and the universal root. *Proc Natl Acad Sci U S A 91*:6721-6728.

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

Dubois DY, Blaise M, Becker HD, Campanacci V, Keith G, Giegé R, Cambillau C, Lapointe J, Kern D. 2004. An aminoacyl-tRNA synthetase-like protein encoded by the *Escherichia coli* yadB gene glutamylates specifically tRNAAsp. *Proc Natl Acad Sci U S A 101*:7530-7535.

Ducruix A, Giegé R. 1999. *Crystallization of Nucleic Acids and Proteins: A Practical Approach (second edition).* Oxford: IRL Press.

**E**bel J-P, Giegé R, Bonnet J, Kern D, Befort N, Bollack C, Fasiolo F, Gangloff J, Dirheimer G. 1973. Factors determining the specificity of the tRNA aminoacylation reaction. *Biochimie 55*:547-557.

Edwards H, Schimmel P. 1987. An *E. coli* aminoacyl-tRNA synthetase can substitute for yeast mitochondrial enzyme function *in vivo*. *Cell 51*:643-649.

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

Enriquez JA, Attardi G. 1996. Evidence for aminoacylation-induced conformational changes in human mitochondrial tRNAs. *Proc Natl Acad Sci USA 93*:8300-8305.

Enriquez JA, Chomyn A, Attardi G. 1995. MtDNA mutation in MERRF syndrome causes defective aminoacylation of $tRNA^{Lys}$ and premature translation termination. *Nature Genetics 10*:47-55.

Entelis NS, Kolesnikova OA, Dogan S, Martin RP, Tarassov IA. 2001. 5S rRNA and tRNA import into human mitochondria: comparison of in vitro requirements. *J Biol Chem 276*:45642-45653.

Eriani G, Delarue M, Poch O, Gangloff J, Moras D. 1990. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature 347*:203-206.

Eriani G, Gangloff J. 1999. Yeast aspartyl-tRNA synthetase residues interacting with $tRNA^{Asp}$ identity bases connectively contribute to $tRNA^{Asp}$ binding in the ground and transition state complex and discriminate against non-cognate tRNAs. *J Mol Biol 291*:761-773.

**F**echter P, Rudinger J, Giegé R, Théobald-Dietrich A. 1998. Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: production and activity. *FEBS Lett 436*:99-103.

Fechter P, Rudinger-Thirion J, Florentz C, Giegé R. 2001. Novel features in the tRNA-like world of plant viral RNAs. *CMLS, Cell Mol Life Sci 58*:1547-1561.

Fechter P, Rudinger-Thirion J, Théobald-Dietrich A, Giegé R. 2000. Identity of tRNA for yeast tyrosyl-tRNA synthetase: Tyrosylation is more sensitive to identity nucleotides than to structural features. *Biochemistry 39*:1725-1733.

Fedor MJ, Williamson JR. 2005. The catalytic diversity of RNAs. *Nat Rev Mol Cell Biol 6*:399-412.

Fersht A. 1977. *Enzyme, Structure and Mechanism.* Reading and San Francisco: W. H. Freeman and Company.

Fett R, Knippers R. 1991. The primary structure of human glutaminyl-tRNA synthetase. *J Biol Chem 266*:1448-1455.

Figler RA, Omote H, Nakamoto RK, Al-Shawi MK. 2000. Use of chemical chaperones in the yeast Saccharomyces cerevisiae to enhance heterologous membrane protein expression: high-yield expression and purification of human P-glycoprotein. *Arch Biochem Biophys 376*:34-46.

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

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

Forterre P, Philippe H. 1999. The last universal common ancestor (LUCA), simple or complex? *Biol Bull 196*:373-375; discussion 375-377.

Francklyn C. 2003. tRNA synthetase paralogs: evolutionary links in the transition from tRNA-dependent amino acid biosynthesis to de novo biosynthesis. *Proc Natl Acad Sci U S A 100*:9650-9652.

Francklyn C, Perona JJ, Pütz J, Hou Y-M. 2002. Aminoacyl-tRNA synthetases: Versatile players in the changing theater of translation. *RNA 8*:1363-1372.

Francklyn C, Schimmel P. 1989. Aminoacylation of RNA minihelices with alanine. *Nature 337*:478-481.

Frugier M, Florentz C, Giegé R. 1994a. Efficient aminoacylation of resected RNA helices by class II aspartyl-tRNA synthetase dependent on a single nucleotide. *EMBO J 13*:2218-2226.

Frugier M, Florentz C, Schimmel P, Giegé R. 1993. Triple aminoacylation specificity of a chimerized transfer RNA. *Biochemistry 32*:14053-14061.

Frugier M, Helm M, Felden B, Giegé R, Florentz C. 1998. Sequences outside recognition sets are not neutral for tRNA aminoacylation: evidence for non-permissive combinations of nucleotides in the acceptor stem of yeast $tRNA^{Phe}$. *J Biol Chem 273*:11605-11610.

Frugier M, Moulinier L, Giegé R. 2000. A domain in the N-terminal extension of class IIb eukaryotic aminoacyl-tRNA synthetases is important for tRNA binding. *EMBO J 19*:2371-2380.

Frugier M, Ryckelynck M, Giege R. 2005. tRNA-balanced expression of a eukaryal aminoacyl-tRNA synthetase by an mRNA-mediated pathway. *EMBO Rep 6*:860-865.

Frugier M, Söll D, Giegé R, Florentz C. 1994b. Identity switches between tRNAs aminoacylated by class I glutaminyl- and class II aspartyl-tRNA synthetase. *Biochemistry 33*:9912-9921.

Gabriel K, McClain WH. 1999. A set of plasmids constitutively producing different RNA levels in *Escherichia coli*. *J Mol Biol 290*:385-390.

Gangloff J, Ebel J-P, Dirheimer G. 1973. Isolation of a complex between yeast arginyl-tRNA synthetase and yeast $tRNA^{Asp}$, and mischarging of $tRNA^{Asp}$ with arginine. *International Research Communications System 12*:8.

Garcia A, Giegé R. 1992. Footprinting evidences for close contacts of the yeast $tRNA^{Asp}$ anticodon region with aspartyl-tRNA synthetase. *Biochem Biophys Res Comm 186*:956-962.

Geslain R, Martin F, Camasses A, Eriani G. 2003. A yeast knockout strain to discriminate between active and inactive tRNA molecules. *Nucleic Acids Res 31*:4729-4737.

Geslain R, Martin F, Delagoutte B, Cavarelli J, Gangloff J, Eriani G. 2000. *In vivo* selection of lethal mutation reveals two functional domains in arginyl-tRNA synthetase. *RNA 6*:434-448.

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

Giegé R, Frugier M. 2003. Transfer RNA structure and identity. In: Lapointe J, Brakier-Gringas L, eds. *Translation mechanisms*. Georgetown, TX: Landes Sciences. pp 1-24.

Giegé R, Rees B. 2005. Aspartyl-tRNA synthetases. In: Ibba M, Francklyn C, Cusack S, eds. *Aminoacyl-tRNA Synthetases.* Georgetown, TX: Landes Bioscience. pp 210-226.

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

Gilbert W. 1986. The RNA world. *Nature 319*:618.

Gilbert W, de Souza SJ, Long M. 1997. Origin of genes. *Proc Natl Acad Sci U S A 94*:7698-7703.

Goldgur Y, Mosyak L, Reshetnikova L, Ankilova V, Lavrik O, Khodyreva S, Safro M. 1997. The crystal structure of phenylalanyl-tRNA synthetase from *Thermus thermophilus* complexed with cognate $tRNA^{Phe}$. *Structure 5*:59-68.

Grosshans H, Hurt E, Simos G. 2000. An aminoacylation-dependent nuclear tRNA export pathway in yeast. *Genes Dev 14*:830-840.

Gu W, Jackman JE, Lohan AJ, Gray MW, Phizicky EM. 2003. tRNAHis maturation: an essential yeast protein catalyzes addition of a guanine nucleotide to the 5' end of tRNAHis. *Genes Dev 17*:2889-2901.

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

Guigou L, Shalak V, Mirande M. 2004. The tRNA-interacting factor p43 associates with mammalian arginyl-tRNA synthetase but does not modify its tRNA aminoacylation properties. *Biochemistry 43*:4592-4600.

Guo F, Gabor J, Cen S, Hu K, Mouland AJ, Kleiman L. 2005. Inhibition of cellular HIV-1 protease activity by lysyl-tRNA synthetase. *J Biol Chem 280*:26018-26023.

**H**amann CS, Hou Y-M. 1997. An RNA structural determinant for tRNA recognition. *Biochemistry 36*:7967-7972.

Hammen PK, Weiner H. 1998. Mitochondrial leader sequences: structural similarities and sequence differences. *The journal of experimental zoology 282*:280-283.

Hao H, Moraes CT. 1997. A disease-associated G5703A mutation in human mitochondrial DNA causes a conformational change and a marked decrease in steady-state levels of mitochondrial $tRNA^{Asn}$. *Mol Cel Biol 17*:6831-6837.

Harris ME, Christian EL. 2003. Recent insights into the structure and function of the ribonucleoprotein enzyme ribonuclease P. *Curr Opin Struct Biol 13*:325-333.

Hartman H. 1995. Speculations on the origin of the genetic code. *J Mol Evol 40*:541-544.

Hasegawa T, Himeno H, Ishikura H, Shimizu M. 1989. Discriminator base of $tRNA^{Asp}$ is involved in amino acid acceptor activity. *Biochem Biophys Res Comm 163*:1534-1538.

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

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

Helm M, Giegé R, Florentz C. 1999. A Watson-Crick base-pair disrupting methyl group ($m^1A9$) is sufficient for cloverleaf folding of human mitochondrial $tRNA^{Lys}$. *Biochemistry 38*:13338-13346.

Himeno H, Hasegawa T, Ueda T, Watanabe K, Miura K, Shimizu M. 1989. Role of the extra G-C pair at the end of the acceptor stem of $tRNA^{His}$ in aminoacylation. *Nucleic Acids Res 17*:7855-7863.

Hipps D, Shiba K, Henderson B, Schimmel P. 1995. Operational RNA code for amino acids: species-specific aminoacylation of minihelices switched by a single nucleotide. *Proc Natl Acad Sci USA 92*:5550-5552.

Hoagland MB, Stephenson ML, Scott JF, Hecht LI, Zamecnik PC. 1958. A soluble ribonucleic acid intermediate in protein synthesis. *J Biol Chem 231*:241-257.

Hohsaka T, Sisido M. 2002. Incorporation of non-natural amino acids into proteins. *Curr Opin Chem Biol 6*:809-815.

Holley RW, Apgar J, Everett GA, Madison JT, Marquise M, Merril SH, Penswick JR, Zamir R. 1965. Structure of a ribonucleic acid. *Science 147*:1462-1465.

Holt IJ, Lorimer HE, Jacobs HT. 2000. Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell 100*:515-524.

Hopper AK, Phizicky EM. 2003. tRNA transfers to the limelight. *Genes Dev 17*:162-180.

Hou Y-M, Schimmel P. 1988. A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature 333*:140-145.

Hou Y-M, Westhof E, Giegé R. 1993. An unusual RNA tertiary interaction has a role for the specific aminoacylation of a transfer RNA. *Proc Natl Acad Sci USA 90*:6776-6780.

Hou YM, Schimmel P. 1989. Evidence that a major determinant for the identity of a transfer RNA is conserved in evolution. *Biochemistry 28*:6800-6804.

Houman F, Rho SB, Zhang J, Shen X, Wang C-C, Schimmel P, Martinis SA. 2000. A prokaryote and human tRNA synthetase provide an essential RNA splicing function in yeast mitochondria. *Proc Natl Acad Sci USA 97*:13743-13748.

Hubert N, Walczak R, Sturchler C, Myslinski E, Schuster C, Westhof E, Carbon P, Krol A. 1996. RNAs mediating cotranslational insertion of selenocysteine in eukaryotic selenoproteins. *Biochimie 78*:590-596.

Ibba M, Morgan S, Curnow AW, Pridmore DR, Vothknecht UC, Gardner W, Lin W, Woese CR, Söll D. 1997. A euryarchaeal lysyl-tRNA synthetase: resemblance to class I synthetases. *Science 278*:1119-1122.

Ibba M, Söll D. 2000. Aminoacyl-tRNA synthesis. *Annu Rev Biochem 69*:617-650.

Ikeuchi Y SA, Ote T, Kato J, Sekine Y, Suzuki T. 2005. Molecular mechanism of lysidine synthesis that determines tRNA identity and codon recognition. *Mol Cell 19*:235-246.

Ingman M, Kaessmann H, Pääbo S, Gyllenstein U. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature 408*:708-713.

Ishitani R, Nureki O, Nameki N, Okada N, Nishimura S, Yokoyama S. 2003. Alternative tertiary structure of tRNA for recognition by a post-transcriptional modification enzyme. *Cell 113*:383-394.

**J**ahn D, Verkamp E, Söll D. 1992. Glutamyl-transfer RNA - A precursor of heme and chlorophyll biosynthesis. *Trends in Biochemical Sciences 17*:215-218.

Jakubowski H, Goldman E. 1992. Editing of errors in selection of amino acids for protein synthesis. *Microbiol Rev 56*:412-429.

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

Jentsch S. 1996. When proteins receive deadly messages at birth. *Science 271*:955-956.

Jiang M, Mak J, Ladha A, Cohen E, Klein M, Rovinski B, Kleiman L. 1993. Identification of tRNAs incorporated into wild-type and mutant human immunodeficiency virus type 1. *J Virol 67*:3246-3253.

Jovine L, Djordjevic S, Rhodes D. 2000. The crystal structure of yeast phenylalanine tRNA at 2.0Å resolution: cleavage by $Mg^{2+}$ in 15-year old crystals. *J Mol Biol 301*:401-414.

Jørgensen R, Søgarrd MM, Rossing AB, Martensen PM, Justesen J. 2000. Identification and characterization of human mitochondrial tryptophanyl-tRNA synthetase. *J Biol Chem 275*:16820-16826.

**K**aiser E, Hu B, Becher S, Eberhard D, Schray B, Baack M, Hameister H, Knippers R. 1994. The human EPRS locus (formerly the QARS locus): a gene encoding a class I and a class II aminoacyl-tRNA synthetase. *Genomics 19*:280-290.

Keiler KC, Shapiro L, Williams KP. 2000. tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: A two-piece tmRNA functions in *Caulobacter. Proc Natl Acad Sci USA 97*:7778-7783.

Keiler KC, Waller RH, Sauer RT. 1996. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science 271*:990-993.

Kelley S, Steinberg S, Schimmel P. 2000. Functional defects of pathogenic human mitochondrial tRNAs related to structural fragility. *Nature Struc Biol 7*:862-865.

Kelley SO, Steinberg SV, Schimmel P. 2001. Fragile T-stem in disease-associated human mitochondrial tRNA sensitizes structure to local and distant mutations. *J Biol Chem 276*:10607-10611.

Kim JY, Kang YS, Lee JW, Kim HJ, Ahn YH, Park H, Ko YG, Kim S. 2002. p38 is essential for the assembly and stability of macromolecular tRNA synthetase complex: implications for its physiological significance. *Proc Natl Acad Sci U S A 99*:7912-7916.

Kim S, Lee C, Choi EC, Choi SY. 2003. Aminoacyl-tRNA synthetases and their inhibitors as a novel family of antibiotics. *Appl Microbiol Biotechnol 61*:278-288.

Kim SH, Suddath FL, Quigley GJ, McPherson A, Sussman JL, Wang AHJ, Seeman NC, Rich A. 1974. Three dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science 185*:435-440.

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

Kirino Y, Yasukawa T, Ohta S, Akira S, Ishihara K, Watanabe K, Suzuki T. 2004. Codon-specific translational defect caused by a wobble modification deficiency in mutant tRNA from a human mitochondrial disease. *Proc Natl Acad Sci U S A 101*:15070-11507 15075.

Klipcan L, Safro M. 2004. Amino acid biogenesis, evolution of the genetic code and aminoacyl-tRNA synthetases. *Journal of Theoretical Biology 228*:389-396.

Kogelnik AM, Lott MT, Brown MD, Navathe SB, Wallace DC. 1998. MITOMAP: a human mitochondrial genome database -1998 update. *Nucleic Acids Res 26*:112-115.

Köhrer C SE, RajBhandary UL. 2004. Complete set of orthogonal 21st aminoacyl-tRNA synthetase-amber, ochre and opal suppressor tRNA pairs: concomitant suppression of three different termination codons in an mRNA in mammalian cells. *Nucleic Acids Res 32*:6200-6211.

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

Krings M, Stone A, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S. 1997. Neandertal DNA sequences and the origin of modern humans. *Cell 90*:19-30.

Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TC. 1982. Self-splicing of RNA: Autoexcision and autocyclizationof ribosomal RNA intervening sequence of *Tetrahymena*. *Cell 31*:147-157.

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

Labouesse M. 1990. The yeast mitochondrial leucyl-tRNA synthetase is a splicing factor for the excision of several group I introns. *Mol Gen Genet 224*:209-221.

Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature 387*:493-497.

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

Larsson N-G, Clayton DA. 1995. Molecular genetic aspects of human mitochondrial disorders. *Annu Rev Genetics 29*:151-178.

Le Grice SF. 2003. "In the beginning": initiation of minus strand DNA synthesis in retroviruses and LTR-containing retrotransposons. *Biochemistry 42*:14349-14355.

Leandro P, Lechner MC, Tavares de Almeida I, Konecki D. 2001. Glycerol increases the yield and activity of human phenylalanine hydroxylase mutant enzymes produced in a prokaryotic expression system. *Mol Genet Metab 73*:173-178.

Lee N, Bessho Y, Wei K, Szostak JW, Suga H. 2000. Ribozyme-catalyzed tRNA aminoacylation. *Nature Struct Biol 7*:28-33.

Lee S, Cho B, Park S, Kim S. 2004. Aminoacyl-tRNA synthetase complexes: beyond translation. *J Cell Science 117*:3725-3734.

Leinfelder W, Zehelein E, Mandrand-Berthelot M-A, Böck A. 1988. Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature 331*:723-725.

Lenhard B, Orellana O, Ibba M, Weygand-Durasevic I. 1999. tRNA recognition and evolution of determinants in seryl-tRNA synthesis. *Nucleic Acids Res 27*:721-729.

Levinger L, Jacobs O, James M. 2001. In vitro 3' end endonucleolytic processing defect in a human mitochondrial $tRNA^{Ser(UCN)}$ precursor with the U7445C substitution, which causes non-syndromic deafness. *Nucleic Acids Res 29*:4334-4340.

Levinger L, Mörl M, Florentz C. 2004. Mitochondrial tRNA 3' end metabolism and human disease. *Nucleic Acids Research 32*:5430-5441.

Liu M, Chu W-C, Liu JC-H, Horowitz J. 1997. Role of acceptor stem conformation in $tRNA^{Val}$ recognition by its cognate synthetase. *Nucleic Acids Res 25*:4883-4890.

Liu M, Spremulli LL. 2000. Interaction of mammalian mitochondrial ribosomes with the inner membrane. *J Biol Chem 275*:29400-29406.

Lovato MA, Chihade JW, Schimmel P. 2001. Translocation within the acceptor helix of a major tRNA identity determinant. *EMBO J 20*:4846-4853.

Lund E, Dahlberg JE. 1998. Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science 282*:2082-2085.

Lynch DC, Attardi G. 1976. Amino acid specificity of the transfer RNA species coded for by HeLa cell mitochondrial DNA. *J Mol Biol 102*:125-141.

Lynch M. 2000. The evolutionary fate and consequences of duplicate genes. *Science 290*:1151-1155.

**M**aizels N, Weiner AL. 1994. Phylogeny from function: Evidence from the molecular fossil record that tRNA originated in replication, not translation. *Proc Natl Acad Sci USA 91*:6729-6734.

Maniataki E, Mourelatos Z. 2005. Human mitochondrial tRNAMet is exported to the cytoplasm and associates with the Argonaute 2 protein. *RNA 11*:849-852.

Marquet R, Isel C, Ehresmann C, Ehresmann B. 1995. tRNAs as primer of reverse transcriptases. *Biochimie 77*:113-124.

Martin W. 1999. Mosaic bacterial chromosomes: a challenge en route to a tree of genomes. *Bioessays 21*:99-104.

Martinis SA, Plateau P, Cavarelli J, Florentz C. 1999. Aminoacyl-tRNA synthetases: a new image for a classical family. *Biochimie 81*:683-700.

Martinis SA, Schimmel P. 1995. Small RNA oligonucleotide substrates for specific aminoacylations. In: Söll D, RajBhandary UL, eds. *tRNA: Structure, Biosynthesis, and Function.* Washington, DC: Am. Soc. Microbiol. Press. pp 349-370.

Masucci JP, Schon EA. 1996. tRNA processing in human mitochondrial disorders. *Mol Biol Rep 22*:187-193.

McClain WH. 1995. The tRNA identity problem: past, present and future. In: Söll D, RajBhandary UL, eds. *tRNA: Structure, Biosynthesis, and Function*. Washington DC: Am. Soc. Microbiol. Press. pp 335-347.

McClain WH, Foss K. 1988. Changing the acceptor identity of a transfer RNA by altering nucleotides in a "variable pocket". *Science 241*:1804-1807.

McClain WH, Gabriel K. 2001. Construction of an Escherichia coli knockout strain for functional analysis of tRNA(Asp). *J Mol Biol 310*:537-542.

McCloskey JA, Crain PF. 1998. The RNA modification database-1998. *Nucleic Acids Res 26*:196-197.

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

Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA template. *Nucleic Acids Res 15*:8783-8798.

Mirande M, Lazard M, Martinez R, Latreille M-T. 1992. Engineering mammalian aspartyl-tRNA synthetase to probe structural features mediating its association with the multisynthetase complex. *Eur J Biochem 203*:459-466.

Moore B, Persson BC, Nelson CC, Gesteland RF, Atkins JF. 2000. Quadruplet codons: implications for code expansion and the specification of translation step size. *J Mol Biol 298*:195-209.

Mörl M, Dörner M, Pääbo S. 1995. C to U editing and modifications during the maturation of the mitochondrial $tRNA^{Asp}$ in marsupials. *Nucleic Acids Res 23*:3380-3384.

Mörl M, Marchfelder A. 2001. The final cut. the importance of tRNA 3'-processing. *EMBO Rep 21*:17-20.

Moulinier L, Eiler S, Eriani G, Gangloff J, Thierry J-C, Gabriel K, McClain WH, Moras D. 2001. The structure of an AspRS-tRNA$^{Asp}$ complex reveals a tRNA-dependant control mechanism. *EMBO J 20*:5290-5301.

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

Muramatsu T, Nishikawa K, Nemoto F, Kuchino Y, Nishimura S, Miyazawa T, Yokoyama S. 1988. Codon and amino acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature 336*:179-181.

Murzin AG. 1993. OB (oligonucleotide/oligosaccharide binding)-binding fold: common structural and functional solution for non-homologous sequences. *EMBO J 12*:861-867.

**N**agel GM, Doolittle RF. 1991. Evolution and relatedness in two aminoacyl-tRNA synthetase families. *Proc Natl Acad Sci USA 88*:8121-8125.

Nakatsu T, Kato H, Oda J. 1998. Crystal structure of asparagine synthetase reveals a close evolutionary relationship to class II aminoacyl-tRNA synthetase. *Nature Struct Biol 5*:15-19.

Nameki N. 1995. Identity elements of tRNA$^{Thr}$ towards *Saccharomyces cerevisiae* threonyl-tRNA synthetase. *Nucleic Acids Res 23*:2831-2836.

Nameki N, Tamura K, Himeno H, Asahara H, Hasegawa T, Shimuzu M. 1992. *E. coli* tRNA$^{Asp}$ recognition mechanism differing from that of the yeast system. *Biochem Biophys Res Comm 189*:856-862.

Ng JD, Sauter C, Lorber B, Kirkland N, Arnez J, Giegé R. 2002. Comparative analysis of space-grown and earth-grown crystals of an aminoacyl-tRNA synthetase: space-grown crystals are more useful for structural determination. *Acta Cryst D58*:645-652.

Noller HF, Hoffarth V, Zimniak L. 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science 236*:1416-1419.

Norcum MT, Warrington JA. 2000. The cytokine portion of p43 occupies a central position within the eukaryotic multisynthetase complex. *J Biol Chem 275*:17921-17924.

Normanly J, Kleina LG, Masson JM, Abelson J, Miller JH. 1990. Construction of *Escherichia coli* amber suppressor tRNA genes. III. Determination of tRNA specificity. *J Mol Biol 213*:719-726.

Normanly J, Masson JM, Kleina LG, Abelson J, Miller JH. 1986. Construction of two *Escherichia coli* amber suppressor genes: tRNA$^{Phe/CUA}$ and tRNA$^{Cys/CUA}$. *Proc Natl Acad Sci USA 83*:6548-6552.

Nunnari J, Marshall WF, Straight A, Murray A, Sedat JW, Walter P. 1997. Mitochondrial transmission during mating in Saccharomyces cerevisiae is determined by mitochondrial fusion and fission and the intramitochondrial segregation of mitochondrial DNA. *Mol Biol Cell 8*:1233-1242.

Nureki O, Vassylyev DG, Tateno M, Shimida A, Nakama T, Fukai S, Konno M, Hendrickson T, Schimmel P, Yokoyama S. 1998. Enzyme structure with two catalytic sites for double-sieve selection of substrate. *Science 280*:578-582.

**O**'Brien TW. 2003. Properties of human mitochondrial ribosomes. *IUBMB Life 55*:505-513.

O'Hanlon TP, Miller FW. 2002. Genomic organization, transcriptional mapping, and evolutionary implications of the human bi-directional histidyl-tRNA synthetase locus (HARS/HARSL). *Biochem Biophys Res Commun 294*:609-614.

O'Hanlon TP, Raben N, Miller FW. 1995. A novel gene oriented in a head-to-head configuration with the human histidyl-tRNA synthetase (HRS) gene encodes an mRNA that predicts a polypeptide homologous to HRS. *Biochem Biophys Res Commun 210*:556-566.

Ofengand J, Bakin A. 1997. Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts. *J Mol Biol 266*:246-268.

**P**almer JD. 1997. Organelle genomes: going, going, gone! *Science 275*:790-791.

Park BJ, Kang JW, Lee SW, Choi SJ, Shin YK, Ahn YH, Choi YH, Choi D, Lee KS, Kim S. 2005a. The haploinsufficient tumor suppressor p18 upregulates p53 via interactions with ATM/ATR. *Cell 120*:209-221.

Park SG, Ewalt KL, Kim S. 2005b. Functional expansion of aminoacyl-tRNA synthetases and their interacting factors: new perspectives on housekeepers. *Trends Biochem Sci 30*:569-574.

Park SG, Jung KH, Lee JS, Jo YJ, Motegi H, Kim S, Shiba K. 1999. Precursor of pro-apoptotic cytokine modulates aminoacylation activity of tRNA synthetase. *J Biol Chem 274*:16673-16676.

Park SG, Kim HJ, Min YH, Choi EC, Shin YK, Park BJ, Lee SW, Kim S. 2005c. Human lysyl-tRNA synthetase is secreted to trigger proinflammatory response. *Proc Natl Acad Sci U S A 102*:6356-6361.

Paukstelis PJ, Coon R, Madabusi L, Nowakowski J, Monzingo A, Robertus J, Lambowitz AM. 2005. A tyrosyl-tRNA synthetase adapted to function in group I intron splicing by acquiring a new RNA binding surface. *Molecular Cell 17*:417-428.

Paul L, Ferguson DJ, Krzycki JA. 2000. The trimethylamine methyltransferase gene and multiple dimethylamine methyltransferase genes of *Methanosarcina barkeri* contain in-frame and read-through amber codons. *Journal of Bacteriology 182*:2520-2529.

Perona JJ, Rould MA, Steitz TA. 1993. Structural basis for transfer RNA aminoacylation by glutaminyl-tRNA synthetase. *Biochemistry 32*:8758-8771.

Perret V, Florentz C, Puglisi JD, Giegé R. 1992. Effect of conformational features on the aminoacylation of tRNAs and consequences on the permutation of tRNA specificities. *J Mol Biol 226*:323-333.

Perret V, Garcia A, Grosjean H, Ebel J-P, Florentz C, Giegé R. 1990. Relaxation of transfer RNA specificity by removal of modified nucleotides. *Nature 344*:787-789.

Pesole G, Gissi C, De Chirico A, Saccone C. 1999. Nucleotide substitution rate of mammalian mitochondrial genomes. *J Mol Evol 48*:427-434.

Pfanner N. 2000. Protein sorting: recognizing mitochondrial presequences. *Current Biology 10*:412-415.

Placido A, Gagliardi D, Gallerani R, Grienenberger JM, Marechal-Drouard L. 2005. Fate of a Larch Unedited tRNA Precursor Expressed in Potato Mitochondria. *J Biol Chem 280*:33573-33579.

Polycarpo C, Ambrogelly A, Bérubé A, Winbush SM, McCloskey JA, Crain PF, Wood JL, Söll D. 2004. An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. *PNAS 101*:12450-12454.

Poterszman A, Delarue M, Thierry J-C, Moras D. 1994. Synthesis and recognition of aspartyl-adenylate by *Thermus thermophilus* aspartyl-tRNA synthetase. *J Mol Biol 244*:158-167.

Puglisi JD, Pütz J, Florentz C, Giegé R. 1993. Influence of tRNA tertiary structure and stability on aminoacylation by yeast aspartyl-tRNA synthetase. *Nucleic Acids Res 21*:41-49.

Pütz J, Florentz C, Benseler F, Giegé R. 1994. A single methyl group prevents the mischarging of a tRNA. *Nature Struct Biol 1*:580-582.

Pütz J, Puglisi JD, Florentz C, Giegé R. 1991. Identity elements for specific aminoacylation of yeast $tRNA^{Asp}$ by cognate aspartyl-tRNA synthetase. *Science 252*:1696-1699.

Pütz J, Puglisi JD, Florentz C, Giegé R. 1993. Additive, cooperative and anti-cooperative effects between identity nucleotides of a tRNA. *EMBO J 12*:2949-2957.

Putzer H, Grunberg-Manago M, Springer M. 1995. Bacterial Aminoacyl-tRNA synthetases: Genes and regulation of expression. In: Söll D, RajBhandary UL, eds. *tRNA: Structure, Biosynthesis, and Function.* Washington, DC: Am. Soc. Microbiol. Press. pp 293-333.

**Q**uevillon S, Agou F, Robinson J-C, Mirande M. 1997. The p43 component of the mammalian multi-synthetase complex is likely to be the precursor of the endothelial monocyte-activating polypeptide II cytokine. *Journal of Biological Chemistry 272*:32573-32579.

Quevillon S, Mirande M. 1996. The p18 component of the multisynthetase complex shares a protein motif with the beta and gamma subunits of eukaryotic elongation factor 1. *FEBS Lett 395*:63-67.

**R**andau L, Munch R, Hohn MJ, Jahn D, Soll D. 2005. Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5'- and 3'-halves. *Nature 433*:537-541.

Reed VS, Wastney ME, Yang DC. 1994. Mechanisms of the transfer of aminoacyl-tRNA from aminoacyl-tRNA synthetase to the elongation factor 1 alpha. *J Biol Chem 269*:32932-32936.

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

Rho SB, Lincecum TLJ, Martinis SA. 2002. An inserted region of leucyl-tRNA synthetase plays a critical role in group I intron splicing. *EMBO J 21*:6874-6881.

Rho SB, Martinis SA. 2000. The bI4 group I intron binds directly to both its protein splicing partners, a tRNA synthetase and maturase, to facilitate RNA splicing activity. *RNA 6*:1882-1894.

Ribas de Pouplana L, Schimmel P. 2001a. Aminoacyl-tRNA synthetases: potential markers of genetic code development. *Trends Biochem Sci 26*:591-596.

Ribas de Pouplana L, Schimmel P. 2001b. Operational RNA code for amino acids in relation to genetic code in evolution. *J Biol Chem 276*:6881-6884.

Ribas de Pouplana L, Schimmel P. 2001c. Two classes of tRNA synthetases suggested by sterically compatible dockings on tRNA acceptor stem. *Cell 104*:191-193.

Ribas de Pouplana L, Turner R, Steer B, Schimmel P. 1998. Genetic code origins: tRNAs older than their synthetases? *Proc Natl Acad Sci USA 95*:11295-11300.

Rich A, RajBhandary UL. 1976. Transfer RNA: molecular structure, sequence, and properties. *Annu Rev Biochem 45*:805-860.

Riès-Kautt M, Ducruix A. 1992. Phase diagrams. In: Ducruix A, Giegé R, eds. *Crystallization of Nucleic Acids and Proteins: A Practical Approach.* Oxford: IRL Press. pp 195-218.

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

Ripmaster T, Shiba K, Schimmel P. 1995. Wide cross-species aminoacyl-tRNA synthetase replacement *in vivo*: Yeast cytoplasmic alanine enzyme replaced by human polymyositis serum antigen. *Proc Natl Acad Sci U S A 92*:4932-4936.

Rizzuto R, Pinton P, Carrington W, Fay FS, Fogarty KE, Lifshitz LM, Tuft RA, Pozzan T. 1998. Close contacts with the endoplasmic reticulum as determinants of mitochondrial Ca2+ responses. *Science 280*:1763-1766.

Robertus JD, Ladner JE, Finch JT, Rhodes D, Brown RS, Clark BFC, Klug A. 1974. Structure of yeast phenylalanine tRNA at 3Å resolution. *Nature 250*:546-551.

Rodin SN, Ohno S. 1995. Two types of aminoacyl-tRNA synthetases could be originally encoded by complementary strands of the same nucleic acid. *Orig Life Evol Biosph 25*:565-589.

Romby P, Caillet J, Ebel C, Sacerdot C, Graffe M, Eyermann F, Brunel C, Moine H, Ehresmann C, Ehressman B, Springer M. 1996. The expression of *E. coli* threonyl-tRNA synthetase is regulated at the translational level by symmetrical operator-repressor interactions. *EMBO J 15*:5976-5987.

Rossmann MG, Moras D, Olsen KW. 1974. Chemical and biological evolution of a nucleotide-binding protein. *Nature 250*:194-199.

Rould MA, Perona JJ, Söll D, Steitz TA. 1989. Structure of *E. coli* glutaminyl-tRNA synthetase complexed with $tRNA^{Gln}$ and ATP at 2.8Å resolution. *Science 246*:1135-1142.

Roy H, Becker HD, Reinbolt J, Kern D. 2003. When contemporary aminoacyl-tRNA synthetases invent their cognate amino acid metabolism. *Proc Natl Acad Sci U S A 100*:9837-9842.

Rudinger J, Blechschmidt B, Ribeiro S, Sprinzl M. 1994. Minimalist aminoacylated RNAs as efficient substrates for elongation factor Tu. *Biochemistry 33*:5682-5688.

Rudinger J, Hillenbrandt R, Sprinzl M, Giegé R. 1996. Antideterminants present in $minihelix^{Sec}$ hinder its recognition by prokaryotic elongation factor Tu. *EMBO J 15*:650-657.

Rudinger J, Puglisi JD, Pütz J, Schatz D, Eckstein F, Florentz C, Giegé R. 1992. Determinant nucleotides of yeast $tRNA^{Asp}$ interact directly with aspartyl-tRNA synthetase. *Proc Natl Acad Sci USA 89*:5882-5886.

Ruff M, Krishnaswamy S, Boeglin M, Poterszman A, Mitschler A, Podjarny A, Rees B, Thierry JC, Moras D. 1991. Class II aminoacyl transfer RNA synthetases: Crystal structure of yeast aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$. *Science 252*:1682-1689.

Ryckelynck M, Giege R, Frugier M. 2005. tRNAs and tRNA mimics as cornerstones of aminoacyl-tRNA synthetase regulations. *Biochimie 87*:835-845.

Ryckelynck M, Giegé R, Frugier M. 2003. Yeast $tRNA^{Asp}$ charging accuracy is threatened by the N-terminal extension of aspartyl-tRNA synthetase. *J Biol Chem 278*:9683-9690.

**S**accone C, De Giorgi C, Gissi C, Pesole G, Reyes A. 1999. Evolutionary genomics in metazoa: the mitochondrial DNA as a model system. *Gene 238*:195-209.

Saito H, Watanabe K, Suga H. 2001. Concurrent molecular recognition of the amino acid and tRNA by a ribozyme. *RNA 7*:1867-1878.

Saks ME, Sampson JR. 1996. Variant minihelix RNAs reveal sequence-specific recognition of the helical $tRNA^{Ser}$ acceptor stem by *E. coli* seryl-tRNA synthetase. *EMBO J 15*:2843-2849.

Saks ME, Sampson JR, Abelson J. 1998. Evolution of a transfer RNA gene through a point mutation in the anticodon. *Science 279*:1665-1670.

Salazar JC, Ambrogelly A, Crain PF, McCloskey JA, Söll D. 2004. A truncated aminoacyl-tRNA synthetase modifies RNA. *PNAS 101*:7536-7541.

Sampath P, Mazumder B, Seshadri V, Gerber CA, Chavatte L, Kinter M, Ting SM, Digman JD, Kim S, Driscoll DM, Fox PL. 2004. Noncanonical function of glutamyl-prolyl-tRNA synthetase: gene-specific silencing of translation. *Cell 119*:195-208.

Sampson JR, Uhlenbeck OC. 1988. Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed *in vitro. Proc Natl Acad Sci USA 85*:1033-1037.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA 74*:5463-5467.

Sanni A, Walter P, Boulanger Y, Ebel J-P, Fasiolo F. 1991. Evolution of aminoacyl-tRNA synthetase quaternary structure and activity: *Saccharomyces cerevisiae* mitochondrial phenylalanyl-tRNA synthetase. *Proc Natl Acad Sci USA 88*:8387-8391.

Santos MAS, Moura G, Massey SE, Tuite MF. 2004. Driving change: the evolution of alternative genetic codes. *Trends in Genetics 20*:95-101.

Sarkar S, Azad AK, Hopper AK. 1999. Nuclear tRNA aminoacylation and its role in nuclear export of endogenous tRNAs in *Saccharomyces cerevisiae. Proc Natl Acad Sci USA 96*:14366-14371.

Sauerwald A, Zhu W, Major TA, Roy H, Palioura S, Jahn D, Whitman WB, Yates JR, Ibba M, Söll D. 2005. RNA-dependent cysteine biosynthesis in Archaea. *Science 307*:1969-1972.

Sauter C, Lorber B, Cavarelli J, Moras D, Giegé R. 2000. The free yeast aspartyl-tRNA synthetase differs from the $tRNA^{Asp}$-complexed enzyme by structural changes in the catalytic site, hinge region, and anticodon binding domain. *J Mol Biol 299*:1313-1324.

Sauter C, Lorber B, Kern D, Cavarelli J, Moras D, Giegé R. 1999. Crystallogenesis studies on aspartyl-tRNA synthetase: use of phase diagram to improve crystal quality. *Acta Cryst D55*:149-156.

Scheffler I. 1999. *Mitochondria*. New-York, Chichester, Weinheim, Brisbane, Singapore, Toronto: John Wiley & Sons, INC., Publication.

Schiffer S, Rosch S, Marchfelder A. 2002. Assigning a function to a conserved group of proteins: the tRNA 3'-processing enzymes. *EMBO J 21*:2769-2777.

Schimmel P. 1995. An operational RNA code for amino acids and variations in critical nucleotide sequences in evolution. *J Mol Evol 40*:531-536.

Schimmel P, Giegé R, Moras D, Yokoyama S. 1993. An operational RNA code for amino acids and possible relationship to genetic code. *Proc Natl Acad Sci USA 90*:8763-8768.

Schimmel P, Kelley SO. 2000. Exiting an RNA world. *Nature Struct Biol 7*:5-7.

Schimmel P, Ribas de Pouplana L. 2000. Footprints of aminoacyl-tRNA synthetases are everywhere. *Trends Biochem Sci 25*:207-209.

Schimmel P, Tamura K. 2003. tRNA structure goes from L to lambda. *Cell 113*:276-278.

Schmitt E, Moulinier L, Fujiwara S, Imanaka T, Thierry J-C, Moras D. 1998. Crystal structure of aspartyl-tRNA synthetase from *Pyrococcus kodakaraensis* KOD: archeon specificity and catalytic mechanism of adenylate formation. *EMBO J 17*:5227-5237.

Schmitz FJ, Lindenlauf E, Hofmann B, Fluit AC, Verhoef J, Heinz HP, Jones ME. 1998. The prevalence of low- and high-level mupirocin resistance in staphylococci from 19 European hospitals. *J Antimicrob Chemother 42*:489-495.

Schön A, Kannangara G, Gough S, Söll D. 1988. Protein biosynthesis in organelles requires misaminoacylation of tRNA. *Nature 331*:187-190.

Schön A, Söll D. 1988. tRNA specificity of a mischarging aminoacyl-tRNA synthetase: Glutamyl-tRNA synthetase from barley chloroplasts. *FEBS Lett 228*:241-244.

Schon EA, Bonilla E, DiMauro S. 1997. Mitochondrial DNA mutations and pathogenesis. *J Bioenerg Biomemb 29*:131-149.

Schulman LH, Pelka H. 1988. Anticodon switching changes the identity of methionine and valine transfer RNAs. *Science 242*:765-768.

Seneca S, Goemans N, Van Coster R, Givron P, Reybrouck T, Sciot R, Meulemans A, Smet J, Van Hove JL. 2005. A mitochondrial tRNA aspartate mutation causing isolated mitochondrial myopathy. *Am J Med Genet A 137*:170-175.

Shaheen HH, Hopper AK. 2005. Retrograde movement of tRNAs from the cytoplasm to the nucleus in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A 102*:11290-11295.

Shalak V, Kaminska M, Mitnacht-Kraus R, Vandenabeele P, Clauss M, Mirande M. 2001. The EMAPII cytokine is released from the mammalian multisynthetase complex after cleavage of its p43/proEMAPII component. *J Biol Chem 276*:23769-23776.

Sherman JM, Rogers K, Rogers MJ, Söll D. 1992a. Synthetase competition and tRNA context determine the *in vivo* identity of tRNA discriminator mutants. *J Mol Biol 228*:1055-1062.

Sherman JM, Rogers MJ, Söll D. 1992b. Competition of aminoacyl-tRNA synthetases for tRNA ensures the accuracy of aminoacylation. *Nucleic Acids Res 20*:2847-2852.

Shiba K, Motegi H, Schimmel P. 1997. Maintaining genetic code through adaptations of tRNA synthetases to taxonomic domains. *Trends Biochem Sci 22*:453-457.

Shiba K, Schimmel P, Motegi H, Noda T. 1994a. Human glycyl-tRNA synthetase. Wide divergence of primary structure from bacterial counterpart and species-specific aminoacylation. *J Biol Chem 269*:30049-30055.

Shiba K, Suzuki N, Shigesada K, Namba Y, Schimmel P, Noda T. 1994b. Human cytoplasmic isoleucyl-tRNA synthetase: Selective divergence of the anticodon-binding domain and acquisition of a new structural unit. *Proc Natl Acad Sci USA 91*:7435-7439.

Shimada N, Matsuzaki K, Suzuki T, Suzuki T, Watanabe K. 2002. Quality control of translation through the kinetic discrimination of tRNAs in the network of aminoacyl-tRNA synthetases. *Nucleic Acids Res Suppl*:79-80.

Shtilbans A, El-Schahawi M, Malkin E, Shanske S, Musumeci O, DiMauro S. 1999. A novel mutation in the mitochondrial DNA transfer ribonucleic acidAsp gene in a child with myoclonic epilepsy and psychomotor regression. *J Child Neurol 14*:610-613.

Sikorski RS, Hieter P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics 122*:19-27.

Simos G, Sauer A, Fasiolo F, Hurt EC. 1998. A conserved domain within Arc1p delivers tRNA to aminoacyl-tRNA synthetases. *Mol Cell 1*:235-242.

Sisido M, Ninomiya K, Ohtsuki T, Hohsaka T. 2005. Four-base codon/anticodon strategy and non-enzymatic aminoacylation for protein engineering with non-natural amino acids. *Methods 36*:270-278.

Sissler M, Delorme C, Bond J, Ehrlich SD, Renault P, Francklyn C. 1999. An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis. *Proc Natl Acad Sci USA 96*:8985-8990.

Sissler M, Eriani G, Martin F, Giegé R, Florentz C. 1997. Mirror image alternative interaction patterns of the same tRNA with either class I arginyl-tRNA synthetase or class II aspartyl-tRNA synthetase. *Nucleic Acids Res 25*:4899-4906.

Sissler M, Giegé R, Florentz C. 1996. Arginine aminoacylation identity is context-dependent and ensured by alternate recognition sets in the anticodon loop of accepting tRNA transcripts. *EMBO J 15*:5069-5076.

Sissler M, Helm M, Frugier M, Giegé R, Florentz C. 2004. Aminoacylation properties of pathology-related variants of human mitochondrial $tRNA^{Lys}$ variants. *RNA 10*:841-853.

Sissler M, Pütz J, Fasiolo F, Florentz C. 2005. Mitochondrial aminoacyl-tRNA synthetases. In: Ibba M, Francklyn C, Cusack S, eds. *Aminoacyl-tRNA synthetases*. Georgetown, TX: Landes Biosciences. pp 271-284.

Sohm B, Frugier M, Brulé H, Olszak K, Przykorska A, Florentz C. 2003. Towards understanding human mitochondrial leucine aminoacylation identity. *J Mol Biol 328*:995-1010.

Spencer A, Heck A, Takeuchi N, Watanabe K, Spremulli L. 2004. Characterization of the human mitochondrial methionyl-tRNA synthetase. *Biochemistry 43*:9743-9754.

Springer M, Graffe M, Dondon J, Grunberg-Manago M. 1989. tRNA-like structures and gene regulation at the translational level. A case of molecular mimicry in *E. coli*. *EMBO J 8*:2417-2424.

Sprinzl M, Vassilenko KS. 2005. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res 33*:D139-D140.

Sutovsky P, Moreno RD, Ramalho-Santos J, Dominko T, Simerly C, Schatten G. 1999. Ubiquitin tag for sperm mitochondria. *Nature 402*:371-372.

**T**akano A, Endo T, Yoshihisa T. 2005. tRNA actively shuttles between the nucleus and cytosol in yeast. *Science 309*:140-142.

Tamura K, Himeno H, Asahara H, Hasegawa T, Shimizu M. 1992. *In vitro* study of *E. coli* $tRNA^{Arg}$ and $tRNA^{Lys}$ identity. *Nucleic Acids Res 20*:2335-2339.

Tarassov IA, Martin RP. 1996. Mechanisms of tRNA import into yeast mitochondria: An overview. *Biochimie 78*:502-510.

Théobald-Dietrich A, Frugier M, Giegé R, Rudinger-Thirion J. 2004. Atypical archaeal tRNA pyrrolysine transcript behaves towards EF-Tu as a typical elongator tRNA. *Nucleic Acids Res 32*:1091-1096.

Tinkle Peterson E, Pan T, Coleman J, Uhlenbeck OC. 1994. *In vitro* selection of small RNAs that bind to *Escherichia coli* phenylalanyl-tRNA synthetase. *J Mol Biol 242*:186-192.

Tolkunova E, Park H, Xia J, King MP, Davidson E. 2000. The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual splicing of the primary transcript. *J Biol Chem 275*:35063-35069.

Tsui WC, Fersht AR. 1981. Probing the principles of amino acid selection using the alanyl-tRNA synthetase from Escherichia coli. *Nucleic Acids Res 9*:4627-4637.

**U**eda T, Yotsumoto Y, Ikeda K, Watanabe K. 1992. The T-loop region of animal mitochondrial $tRNA^{Ser(AGY)}$ is a main recognition site for homologous seryl-tRNA synthetase. *Nucleic Acids Res 20*:2217-2222.

Uy R, Wold F. 1977. Post-translational covalent modification of proteins. *Science 198*:890-896.

**V**erkamp E, Kumar AM, Lloyd A, Martins O, Stange-Thomann N, Söll D. 1995. Glutamyl-tRNA as an intermediate in glutamate conversions. In: Söll D, RajBhandary U, eds. *tRNA : Structure, Biosynthesis and Function*. Washington, DC: ASM Press. pp 545-550.

Vörtler S, Pütz J, Giegé R. 2001. Manipulation of tRNA properties by structure-based and combinatorial *in vitro* approaches. *Prog Nucleic Acid Res Mol Biol 70*:291-334.

**W**akasugi K, Quinn CL, Tao N, Schimmel P. 1998. Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant. *EMBO J 17*:297-305.

Wang L, Brock A, Herderich B, Schultz PG. 2001. Expanding the genetic code of *Escherichia coli*. *Science 292*:498-500.

Weiner AM. 2004. tRNA maturation: RNA polymerization without a nucleic acid template. *Curr Biol 14*:R883-885.

Widmann J, Giulio MD, Yarus M, Knight R. 2005. tRNA Creation by Hairpin Duplication. *J Mol Evol 61*:524-530.

Woese CR. 1990. Evolutionary questions: the "progenote". *Science 247*:789.

Woese CR. 1998. The universal ancestor. *PNAS 95*:6854-6859.

Woese CR, Olsen GJ, Ibba M, Söll D. 2000. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol and Mol Biol Reviews 64*:202-236.

Wolf YI, Aravind L, Grishin NV, Koonin EV. 1999. Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res 9*:689-710.

Wolfe CL, Warrington JA, Treadwell L, Norcum MT. 2005. A three-dimensional working model of the multienzyme complex of aminoacyl-tRNA synthetases based on electron microscopic placements of tRNA and proteins. *J Biol Chem*. in press.

Wolfson A, Knight R. 2005. Occurrence of the aminoacyl-tRNA synthetases in high-molecular weight complexes correlates with the size of substrate amino acids. *FEBS Lett 579*:3467-3472.

Wolfson AD, Khvorova AM, Sauter C, Florentz C, Giegé R. 1999. Mimics of yeast $tRNA^{Asp}$ and their recognition by aspartyl-tRNA synthetase. *Biochemistry 38*:11926-11932.

Wong FC, Beuning PJ, Nagan M, Shiba K, Musier-Forsyth K. 2002. Functional role of the prokaryotic proline-tRNA synthetase insertion domain in amino acid editing. *Biochemistry 41*:7108-7115.

**X**iao S, Day-Storms JJ, Srisawat C, Fierke CA, Engelke DR. 2005. Characterization of conserved sequence elements in eukaryotic RNase P RNA reveals roles in holoenzyme assembly and tRNA processing. *Rna 11*:885-896.

Xie J, Schultz PG. 2005. An expanding genetic code. *Methods 36*:227-238.

Xue H, Tong KL, Marck CG, H., Wong JT. 2003. Transfer RNA paralogs: evidence for genetic code-amino acid biosynthesis coevolution and an archaeal root of life. *Gene 310*:59-66.

**Y**anagisawa T, Lee JT, Wu HC, Kawakami M. 1994. Relationship of protein structure of isoleucyl-tRNA synthetase with pseudomonic acid resistance of *Escherichia coli*. A

proposed mode of action of pseudomonic acid as an inhibitor of isoleucyl-tRNA synthetase. *J Biol Chem 269*:24304-24309.

Yang XL, Schimmel P, Ewalt KL. 2004. Relationship of two human tRNA synthetases used in cell signaling. *Trends Biochem Sci 29*:250-256.

Yanisch-Perron C, Vieira J, Messing J. 1985. Improved M13 phage cloning vectors and host strain: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene 33*:103-119.

Yaremchuk A, Kriklivyi I, Tukalo M, Cusack S. 2002. Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. *EMBO J 21*:3829-3840.

Yasukawa T, Hino N, Suzuki T, Watanabe K, Ueda T, Ohta S. 2000a. A pathogenic point mutation reduces stability of mitochondrial mutant $tRNA^{Ile}$. *Nucleic Acids Res 28*:3779-3784.

Yasukawa T, Suzuki T, Ishii N, Ueda T, Ohta S, Watanabe K. 2000b. Defect in modification at the anticodon wobble nucleotide of mitochondrial $tRNA^{Lys}$ with the MERRF encephalomyopathy pathogenic mutation. *FEBS Lett 467*:175-178.

Yasukawa T, Suzuki T, Suzuki T, Ueda T, Ohta S, Watanabe K. 2000c. Modification defect at anticodon wobble nucleotide of mitochondrial $tRNAs^{Leu(UUR)}$ with pathogenic mutations of mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes. *J Biol Chem 275*:4251-4257.

Yokogawa T, Shimada N, Takeuchi N, Benkowski L, Suzuki T, Omori A, Ueda T, Nishikawa K, Spremulli LL, Watanabe K. 2000. Characterization and tRNA recognition of mammalian mitochondrial seryl-tRNA synthetase. *J Biol Chem 275*:19913-19920.

Yoshihisa T, Yunoki-Esaki K, Ohshima C, Tanaka N, Endo T. 2003. Possibility of cytoplasmic pre-tRNA splicing: the yeast tRNA splicing endonuclease mainly localizes on the mitochondria. *Mol Biol Cell 14*:3266-3279.

Yun HJ, Lee SW, Yoon GM, Kim SY, Choi S, Lee YS, Choi EC, Kim S. 2003. Prevalence and mechanisms of low- and high-level mupirocin resistance in staphylococci isolated from a Korean hospital. *J Antimicrob Chemother 51*:619-623.

**Z**hang B, Cech TC. 1997. Peptide bond formation by *in vitro* selected ribozymes. *Nature 390*:96-100.

## Summary

My work has focused on the specific recognition of transfer RNAs (tRNAs) by aminoacyl-tRNA synthetases (aaRSs), an obligate prerequisite for translation fidelity. I have taken advantage of molecular biology strategies, based on in vitro transcribed tRNAs and cloned enzymes, to explore the structure/function relationships of yeast and human mitochondrial (mt) aminoacylation systems using large mutagenic analyses. Structural and functional aspects were further tackled by crystallization assays and in vivo approaches, respectively.

So far, it was believed that recognition and aminoacylation rules of isoacceptor tRNAs from a given organism are identical. Investigation of the family of arginine isoaccepting tRNAs in yeast and its peculiar relationship with $tRNA^{Asp}$ lead me to the following discoveries: (i) isoacceptors are aminoacylated with different efficiencies (~20 fold range) and are protected from mischarging by idiosyncratic antideterminants, (ii) isoacceptor $tRNA_4^{Arg}$ is a remnant aspartate acceptor since only two point mutations were sufficient to convert its specificity - this is a direct example of genesis of molecular diversity from a common ancestor. Aminoacylation systems of mammalian mitochondria remain under-explored despite their tRNAs, coded by mt genome, are structurally "bizarre" and involved in severe disorders. Our efforts lead to the assignment of 10 missing nuclear genes coding for human mt aaRSs, which turned out to be encoded by a different set of genes than the one for cytosolic aaRSs. Detailed analysis of the aspartylation system, chosen as a model mt system, revealed (i) less stringent identity of a mt tRNA than of classical tRNAs, (ii) a subtle and focused adaptation of the bacterial-type nuclear-encoded mt AspRS. This illustrates co-evolutionary processes of the human mt and nuclear genomes. Further, I have uncovered the signals hindering a mt $tRNA^{Asp}$ to be a substrate for a non-mt aaRS. Strikingly, it is not the global structural degeneracy of the tRNA which hinders the most cross-aminoacylation, but a single base-pair in the D-stem.

**Key words:** aminoacylation, tRNA, aminoacyl-tRNA synthetase, tRNA identity elements, human mitochondrial systems, isoacceptors, evolution.

## Résumé

Le travail de cette thèse s'inscrit dans le cadre de l'étude des règles qui régissent la spécificité d'aminoacylation des ARN de transfert (ARNt) par les aminoacyl-ARNt synthétases (aaRS). La précision de cette réaction est cruciale puisqu'elle détermine la fidélité de la traduction de l'information génétique et la synthèse de protéines fonctionnelles. J'ai tiré profit des stratégies de biologie moléculaire, basées sur la transcription *in vitro* des ARNt, la production d'enzymes clonées, et la mutagénèse, afin d'explorer les relations structure/fonction des systèmes d'aminoacylation de levure et de la mitochondrie humaine. Les aspects fonctionnels et structuraux ont été davantage explorés par des essais de cristallisation et des approches *in vivo*.

Jusqu'à présent, il était admis que les règles de reconnaissance et d'aminoacylation d'ARNt isoaccepteurs pour un système donné devaient être identiques. L'analyse d'une famille d'ARNt isoaccepteurs de l'arginine de levure et de sa relation particulière avec l'ARNt$^{Asp}$ nous ont permis d'établir que : (i) les isoaccepteurs sont arginylés avec des efficacités différentes (un facteur 20 les sépare) et sont protégés de la misaminoacylation par des antidéterminants idiosyncrasiques, (ii) l'isoaccepteur ARNt$_4^{Arg}$ possède des propriétés d'aspartylation, vestiges de son histoire évolutive, puisque seulement deux mutations sont suffisantes pour convertir sa spécificité – c'est un exemple de génération de la diversité moléculaire par duplication de gènes. Les systèmes d'aminoacylation mt de mammifères restent peu étudiés, et ce malgré la « bizarrerie » structurale et l'implication dans des pathologies sévères de leurs ARNt, codés par le génome mt. Nos efforts ont permis l'assignement des 10 gènes nucléaires manquants codant pour les aaRS mt humaines. Ceux-ci sont portés par un jeu de gènes différents de celui codant pour les sysnthétases cytoplasmiques. L'analyse détaillée du système d'aspartylation, choisi comme système modèle a révélé (i) une identité de l'ARNt mt moins stringente que celle des ARNt classiques, (ii) une adaptation subtile et ciblée de l'aaRS mt, codée par le génome nucléaire et de type bactérien. Ceci illustre un processus de co-évolution entre les génomes mt et nucléaire humain. De plus, j'ai déterminé les signaux qui protègent l'ARNt$^{Asp}$ mt d'être un substrat des aaRS non mt. De manière surprenante, ce n'est pas la dégénérescence structurale globale de l'ARNt qui empêche le plus cette aminoacylation croisée mais une simple paire de bases du bras D.

**Mots clés :** aminoacylation, ARNt, aminoacyl-ARNt synthétase, éléments d'identité de l'ARNt, systèmes d'aminoacylation mitochondriaux humains, isoaccepteurs, évolution.