



**HAL**  
open science

# Rôle des modifications des ARN dans la fidélité de la traduction chez les eucaryotes

Charlène Valadon

► **To cite this version:**

Charlène Valadon. Rôle des modifications des ARN dans la fidélité de la traduction chez les eucaryotes. Biochimie, Biologie Moléculaire. Université Paris-Saclay, 2022. Français. NNT : 2022UPASL004 . tel-04861208

**HAL Id: tel-04861208**

**<https://theses.hal.science/tel-04861208v1>**

Submitted on 2 Jan 2025

**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.

# Rôle des modifications des ARN dans la fidélité de la traduction chez les eucaryotes

*Role of RNAs' modifications in eukaryotic translation fidelity*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 577, Structure et Dynamique des  
Systèmes Vivants (SDSV)

Spécialité de doctorat : Biologie moléculaire et cellulaire

Graduate School : Life Science and Health. Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **I2BC (Université Paris-Saclay, CEA, CNRS)**,  
sous la direction d'**Olivier NAMY**, Directeur de recherche, la co-direction de **Karen  
PERRONET**, Chargée de Recherche

Thèse soutenue à Paris-Saclay, le 13 janvier 2022, par

**Charlène VALADON**

## Composition du Jury

<b>Herman VAN TILBEURGH</b> Professeur, UPSaclay	Président
<b>Clément CARRE</b> Maitre de conférences, HDR, SorbonneU	Rapporteur & Examineur
<b>Alexandre DAVID</b> Chargé de recherche, HDR, Inserm, UMontpellier	Rapporteur & Examineur
<b>Valérie HEURGUE HAMARD</b> Directeur de recherche, CNRS, SorbonneU	Examinatrice
<b>Olivier NAMY</b> Directeur de recherche, CNRS, UPSaclay	Directeur de thèse



Afin de débiter mon manuscrit de thèse je souhaite avant tout remercier l'ensemble des personnes qui m'ont entourée durant mon doctorat.

Olivier,

Durant ma soutenance tu as fait un très beau discours, avec notamment un passage sur ma capacité à garder le cap dans les épreuves. Nous pouvons dire que j'ai eu un très bon professeur puisque tu es la personne la plus optimiste que je n'ai jamais rencontrée. Ta passion pour la recherche a été une énorme source d'inspiration depuis le premier jour, j'ai toujours adoré ces heures de discussions autour d'une question qui nous était passée par la tête pour x ou y raisons. Tu as toujours été à l'écoute et m'as formée pour devenir une doctorante accomplie, que ce soit dans la démarche et la logique scientifique, la communication autour de mon travail à l'oral ou à l'écrit – ce qui t'auras valu l'apparition d'un certain nombre de cheveux blancs – et l'indépendance. Je te remercie pour ta bienveillance et tes conseils, ta confiance et les moyens que tu m'as donnés pour concrétiser l'ensemble de mes projets. Je te remercie également de m'avoir fait découvrir la patate cuite au barbecue, certains diront que c'est un détail alors qu'il s'agissait de la dernière forme de patate que je n'avais encore jamais goûtée, et dieu sait que j'aime les patates !

Karen,

Telle que Dark Vader tu m'as poussée dans le côté obscur de la recherche, la physique ! J'ai pris beaucoup de plaisir à travailler à tes côtés, et de ces longues journées de mises au point dans le noir je ne retiendrai que le positif : l'entraide, l'écoute, la confiance et la persévérance que nous avons partagées. J'ai aimé être formée sur un sujet aussi éloigné de mes connaissances initiales, et cela a impliqué de ton côté que tu encadres une doctorante avec des bagages très différents des tiens. Alors merci d'avoir tenu ton rôle de co-encadrante, d'avoir toujours fait preuve de patience, tant pour me former que pour m'écouter parler de biologie.

Laure,

C'est compliqué de mettre des mots de remerciements pour représenter ce que tu as pu m'apporter durant ces années. Tout d'abord encadrante de stage, tu es rapidement devenue mon binôme de laboratoire. Nous avons traversé ensemble de nombreuses phases de mises au point, avec des surprises comme l'observation de cultures cellulaires fluorescentes ou les bouchages des microfiltres pour la purification de la GST qui m'auront donné l'occasion de goûter des extraits de cellules. Tu as toujours fait preuve d'écoute, d'attention et de compréhension qui t'ont rapidement donné une casquette de maman de labo à mes yeux. Alors voilà, merci de m'avoir accompagnée, soutenue et aidé à grandir, merci pour tout.

Agnès,

Mi-dragon, mi-mère poule, j'ai pu découvrir au fur et à mesure l'ensemble des facettes de ta personnalité qui sort de l'ordinaire. Dans un premier temps impressionnée, j'ai par la suite extrêmement apprécié ton ouverture à partager autour de nos résultats à tout moment de la journée. Avec toi, pas besoin d'attendre les réunions de labo pour avoir une oreille attentive et attentionnée et je souhaiterais t'en remercier. Tu as également réussi un exploit qui est de me convertir au tri des déchets, et crois moi qu'à présent à chaque fois que je m'approche d'une poubelle j'ai une mini Agnès sur l'épaule me poussant à choisir le bon collecteur !

Isabelle,

Une expression populaire dit que lorsque l'on a une question « Google est ton ami ». Visiblement la personne ayant inventé cette expression n'a jamais mis un pied dans l'équipe GST. Une question sur un protocole ? Demandes à Isa. Une question sur d'anciens résultats ? Va voir Isa. Une question sur le stock ou la commande d'un produit ? Attends une petite heure, Isa est au tennis ! Merci pour toute l'aide précieuse que tu as pu m'apporter tout au long de mon séjour dans l'équipe, avec par exemple la production de mes banques pour le ribo seq. Merci également d'avoir été présente dans des moments aussi importants que la fin du premier confinement ou encore lors d'une crise existentielle sur le devenir de bambous déposés en bord de route dans la fac. Merci de m'avoir confié tes bébés plantes, même si honnêtement à les croire elles ne t'en sont pas reconnaissantes !

Céline,

Céline la mystérieuse ! Je pense que tu es la personne de l'équipe avec laquelle j'ai eu la chance d'échanger le plus tardivement. Souvent à courir de partout, tu as pourtant pris le temps de te poser pour échanger autour de mes résultats dès que j'en avais besoin. Je voudrais te remercier de m'avoir aiguillée à de nombreuses reprises dans les mises au point de protocoles. Tu n'es jamais en manque d'inspiration pour trouver de nouvelles pistes lorsque l'on pense avoir tout testé, et je pense que je peux remercier également tes célèbres galères expérimentales pour ça !

Sabrina,

Je pense que tu n'as pas besoin que j'étales ici les raisons pour lesquelles j'aimerais te remercier. Tu as été mon duo de bureau, de manips et de café. En général les doctorants se relaient dans le temps au sein d'une équipe alors que j'ai eu la chance de passer l'ensemble de mon doctorat avec toi. Nos moments rien qu'à nous étaient un incontournable de ma journée. Il paraît qu'à Lyon il y a de bons restaurants et de

belles randos, j'espère que ce sera un argument assez fort pour te voir débarquer sur ta moto et passer de nouveaux moments à tes côtés :-p

Stéphane,

Stéphane le bricoleur, le pâtissier, le dessinateur, le résolveur d'énigmes, le sportif, le levuriste, le bio-info... Merci d'avoir été aussi efficace dans l'analyse de mon ribo-seq qui est arrivé comme un cheveu sur la soupe à moins de deux semaines du dépôt de mon manuscrit. Merci également pour tous ces moments off à parler de tout et de rien autour d'une brioche, d'une boîte de sardines, ou d'une mousse au chocolat keto. En revanche je ne te remercie pas de m'avoir fait connaître ton tiramisu, il faut être un monstre pour me le faire goûter sans jamais me le refaire à nouveau !

Hugo,

Avec le confinement et la rédaction de mon manuscrit nous n'avons pas eu l'occasion de partager beaucoup de moments. Pourtant dès que j'en ai eu besoin tu as répondu présent et toujours avec le sourire. Merci de m'avoir aidée dans mes stats, et surtout d'avoir tout pris en charge pour la diffusion de ma soutenance de thèse. J'espère que tu n'auras pas trop de séquelles suite à cette aide, et que l'image de ma mère criant en robe de chambre sur ton écran ne te hantera pas jusqu'à la fin de ta vie !

Nelly,

Je pense sincèrement que tu es la petite pépite cachée de l'équipe GST. Toujours présente malgré tout ce qui peut t'arriver, c'est tous les membres de l'équipe, voir du couloir, voir du bâtiment qui ont toujours pu compter sur ta volonté à ce que tout soit parfait pour que l'on puisse travailler dans les meilleures conditions. Merci pour tout ce que tu fais, et surtout merci pour toutes ces conversations longues de sens qui nous ont permis de construire une belle amitié.

Pour compléter ces remerciements, j'aimerais remercier sincèrement David Cornu de m'avoir accueillie dans son laboratoire, de m'avoir formée sur la spectrométrie de masse et accompagnée dans l'ensemble de mes analyses, toujours dans la bonne humeur. Je remercie également Mickaël Bourge pour le tri de mes cellules par cytométrie, mais surtout de m'avoir permis d'être présente pendant ces tris tout en répondant à mes nombreuses questions sans jamais sourcilier. J'adresse un grand merci à l'ensemble des coachs de l'équipe iGEM go Paris Saclay 2021 pour le soutien et les échanges scientifiques tout au long de cette année riche en émotions.

Pour finir, comment ne pas remercier mes proches qui ont été des soutiens irremplaçables depuis le début comme mes parents ou depuis les bancs de l'école comme ma sœur de cœur Caroline. Un énorme et dernier merci à mon homme Kévin, pour toute cette joie au quotidien, merci d'avoir été un pilier sans faille pendant les

tempêtes, d'avoir supporter toutes mes crises existentielles et de toujours m'encourager dans la réalisation de mes rêves.

On dit que lorsque l'on croit assez dans ces rêves, ils se réalisent. J'y ai pourtant cru très fort ce qui ne m'empêchait pas de trébucher. La volonté c'est important, mais des mains tendues m'aidant à monter les différentes marches, je dirai que c'était ça la clé de la réussite. Alors merci à tous de m'avoir accompagnée, soutenue et aidée à grandir - mentalement, car visiblement pour la taille personne n'avait de solution – pendant toutes ces années. Des bisous !

## Table des matières

I.	La traduction chez les eucaryotes .....	1
1.	Les acteurs de la traduction .....	1
A.	Le ribosome .....	1
B.	L'ARNm.....	5
C.	L'ARNt.....	7
D.	Le code génétique.....	9
2.	Les étapes de la traduction .....	11
A.	L'initiation .....	13
B.	L'élongation .....	21
C.	La terminaison.....	23
D.	Le recyclage .....	27
II.	La fidélité de la traduction chez les eucaryotes .....	31
1.	Les mécanismes de contrôle-qualité de la traduction chez les eucaryotes.....	31
A.	Dégradation cytoplasmique des ARNm.....	31
B.	Les voies de surveillance de la traduction .....	33
2.	Le rôle du ribosome dans le maintien de la fidélité.....	37
A.	Composition variable du ribosome eucaryote.....	39
B.	Données structurales sur le site de décodage .....	41
3.	Le recodage.....	45
A.	Le décalage du cadre de lecture .....	45
B.	La translecture.....	49
III.	Les modifications d'ARN dans la fidélité de la traduction.....	52
1.	Rôle des modifications d'ARN dans la fidélité d'incorporation des acides aminés	52
A.	Control of Translation Fidelity by Modifications to Cytosolic tRNAs.....	55
B.	Role of rRNA Modifications in Translation Fidelity .....	59
C.	mRNA Modifications Influence the Reading of the Genetic Code .....	60
D.	Manipulation of RNA Modifications to Treat Human Diseases .....	62
2.	Les modifications de l'anticodon de l'ARNt <sub>Tyr</sub> humain.....	65

A.	Intérêt thérapeutique des modifications de la boucle anticodon des ARNt dans le cadre des maladies à codon stop prématuré .....	65
B.	Galactosyl-queuosine-34 .....	69
C.	Pseudouridine-35 de l'ARNt <sub>Tyr</sub> .....	71
3.	2'-O-méthylation de l'ARNr humain .....	73
A.	2'-O-méthylation de l'ARNr : structure, fonctions et formation .....	73
B.	Variabilité dans le cadre de cancers .....	77
IV.	Projets et objectifs de thèse .....	82
V.	Résultats .....	84
1.	Les modifications de la boucle anticodon des ARNt .....	84
	Résumé de l'article "Stop codon readthrough in human: natural suppressors and impact of tRNA anticodon loop modifications" .....	84
	Stop codon readthrough in human: natural suppressors and impact of tRNA anticodon loop modifications .....	86
	Abstract .....	86
	Introduction .....	86
	Materials and Methods .....	88
	Results .....	91
	Discussion .....	100
2.	Résultats supplémentaires sur les modifications de la boucle anticodon des ARNt	103
A.	Mise au point de l'analyse des acides aminés incorporés par translecture	103
B.	Impact des concentrations en queuine dans le milieu de culture sur la translecture .....	109
C.	Obtention de la lignée $\Delta$ PUS7 .....	109
D.	Analyse fonctionnelle des gènes différentiellement exprimés en lignée $\Delta$ PUS7 .....	111
E.	Obtention des lignées $\Delta$ PUS7L .....	115
F.	Analyse de la translecture dans les lignées $\Delta$ PUS7L et $\Delta$ PUS7 $\Delta$ PUS7L .....	117
3.	Modification de l'ARNr .....	117
A.	Reconstitution du système de traduction <i>in vitro</i> .....	119

B.	Analyse cinétique de l'IRES EMCV par smTIRF.....	127
VI.	Discussion générale .....	133
VII.	Protocoles.....	137
1.	Transcription in vitro.....	137
2.	Transfection du système CRISPR Cas9.....	137
3.	Transfection des plasmides des systèmes rapporteurs.....	138
4.	Purification GST .....	138
5.	Western Blot .....	139
6.	Extraction d'ARN totaux .....	139
7.	Extraction d'ADN génomique.....	140
VIII.	Références bibliographiques .....	141
IX.	Annexes.....	175
	Annexe 1 : références de la revue .....	175
	Données supplémentaires de l'article de résultats.....	183

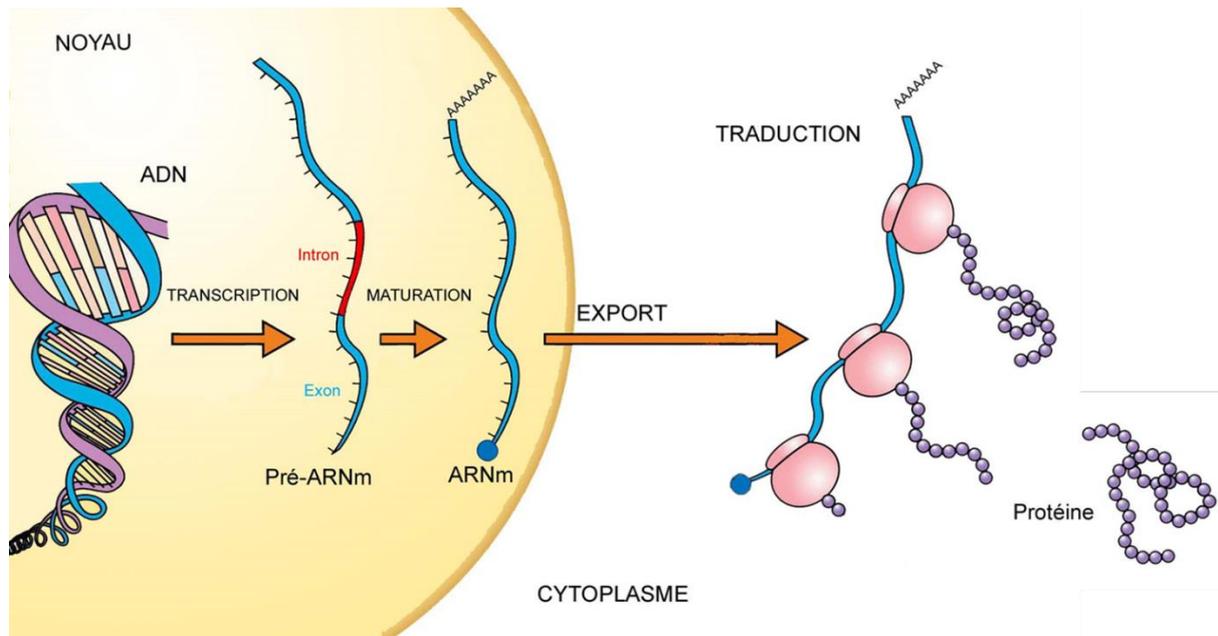


Figure 1 : Le dogme central de la biologie moléculaire. Les gènes présents sur l'ADN sont transcrits en pré-ARNm. Le pré-ARNm est composé d'exons (bleu) et d'introns (rouge). Au cours de la maturation les introns sont éliminés par épissage et les extrémités 5' et 3' sont protégées contre la digestion par les exonucléases par l'ajout d'une coiffe 7-méthylguanosine triphosphate et d'une queue poly-adénines respectivement. Ces molécules d'ARNm matures sont exportées vers le cytoplasme où les ribosomes le traduisent en une chaîne polypeptidique qui se replie en une protéine.

Chaque cellule d'organismes uni ou pluricellulaires possède un ensemble de caractéristiques issues de l'expression de son génome. Ainsi, de manière simpliste, les gènes sont transcrits en ARN messagers avant d'être traduits en protéines (Figure 1). En 2018, l'équipe de Grant W. Brown indiquait la présence de 42 millions de protéines par cellule eucaryote [1]. Ces données soulignent l'importance pour la cellule de la traduction des ARNm en protéines. Ce processus cellulaire est au cœur de mon sujet de thèse qui porte sur le rôle des modifications d'ARN dans la fidélité de la traduction eucaryote. Je vais débiter mon manuscrit par une description de la traduction eucaryote, d'une part via les différents acteurs qui interviennent, puis via les différentes étapes qui la composent. Je continuerai avec un chapitre portant sur la fidélité de la traduction, tant sur les mécanismes cellulaires de maintien de cette fidélité, que sur les erreurs et les dérives de traduction qu'il est possible de rencontrer. Pour finir, j'aborderai dans cette introduction l'importance des modifications d'ARN dans la fidélité de la traduction mais également de l'importance de leur étude pour le traitement de pathologies humaines.

## **I. La traduction chez les eucaryotes**

La traduction est un processus cellulaire indispensable à l'expression génique. Au cours de la traduction, le ribosome décode l'ARN messager (ARNm) et permet ainsi de produire de la protéine à partir des ARN de transfert (ARNt) et du code génétique. Les différents acteurs ainsi que les étapes de la traduction sont décrits dans ce chapitre.

### **1. Les acteurs de la traduction**

Avant d'entrer dans les détails mécanistiques de la traduction, je vais commencer par présenter ses acteurs moléculaires principaux : le ribosome, l'ARNm, l'ARNt. Je finirai par la description des règles de décodage qui la régissent : le code génétique.

#### **A. Le ribosome**

C'est au milieu des années 1950 que George Palade, prix nobel en 1974, décrit pour la première fois la présence d'un « petit composant particulaire » dans le cytoplasme [2]. En effet, il observe par microscopie électronique la présence de « granulés » associés au réticulum endoplasmique qu'il nomma microsomes. Quelques années après, ces microsomes, composés d'ARN et protéines, sont décrits comme étant le site cellulaire de la synthèse protéique. Ils sont alors nommés ribosomes par la communauté scientifique. Il faudra attendre le début des années 2000 pour avoir une première structure complète du ribosome procaryote [3], et 2011 pour avoir celle du ribosome eucaryote [4]. Le ribosome est composé de deux sous unités, la petite sous unité 30S et la grande sous unité 50S (procaryotes), la petite sous unité 40S et la grande sous unité 60S (eucaryotes) pour leur coefficient de sédimentation (unités Svedberg), formant le ribosome 70S et 80S respectivement une fois assemblées.

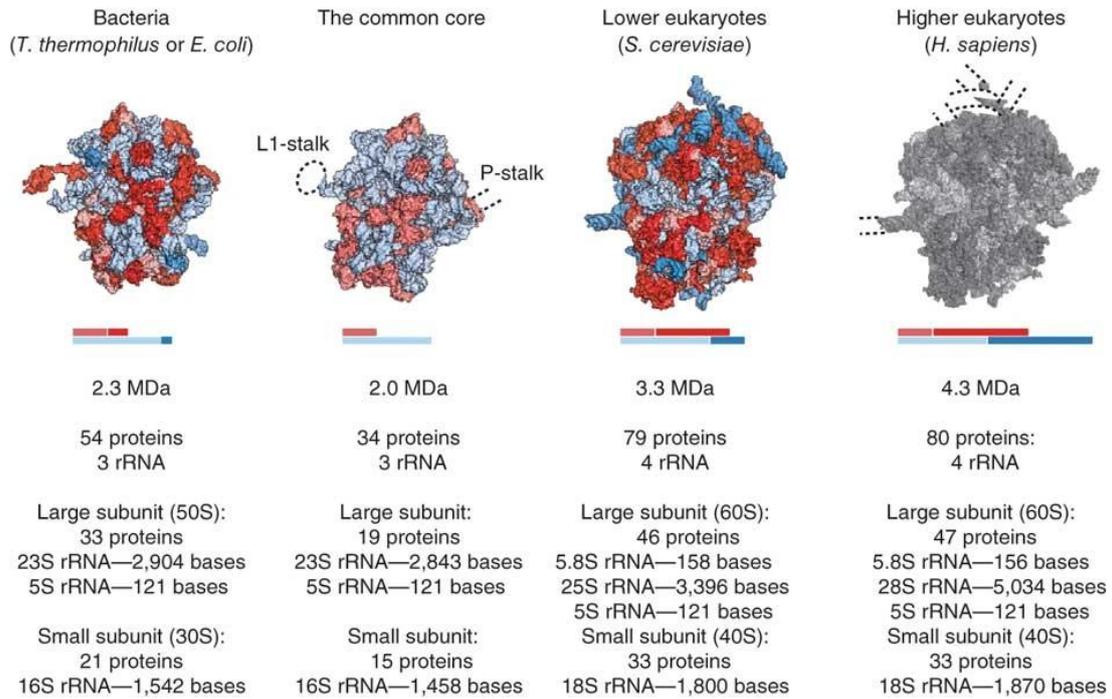


Figure 2 : Comparaison de la composition et de la structure des ribosomes procaryotes et eucaryotes. Les ribosomes bactériens et eucaryotes partagent un noyau conservé composé d'ARN (bleu clair) et de protéines (rouge clair). Ils contiennent également des protéines ou extensions protéiques (rouge foncé), et des extensions de l'ARNr (bleu foncé) qui leur sont propres. Les ARNr 5.8S et 25S-28S eucaryotes sont tous deux homologues de l'ARNr 23S procaryote. Les lignes pointillées autour du noyau commun indiquent les positions des tiges flexibles des ribosomes. La structure présentée du 80S des eucaryotes dits supérieurs est en fait une copie de celles des eucaryotes dits inférieurs, avec des lignes pointillées indiquant les positions des segments d'expansion de l'ARNr spécifique de l'Homme. Issue de [5].

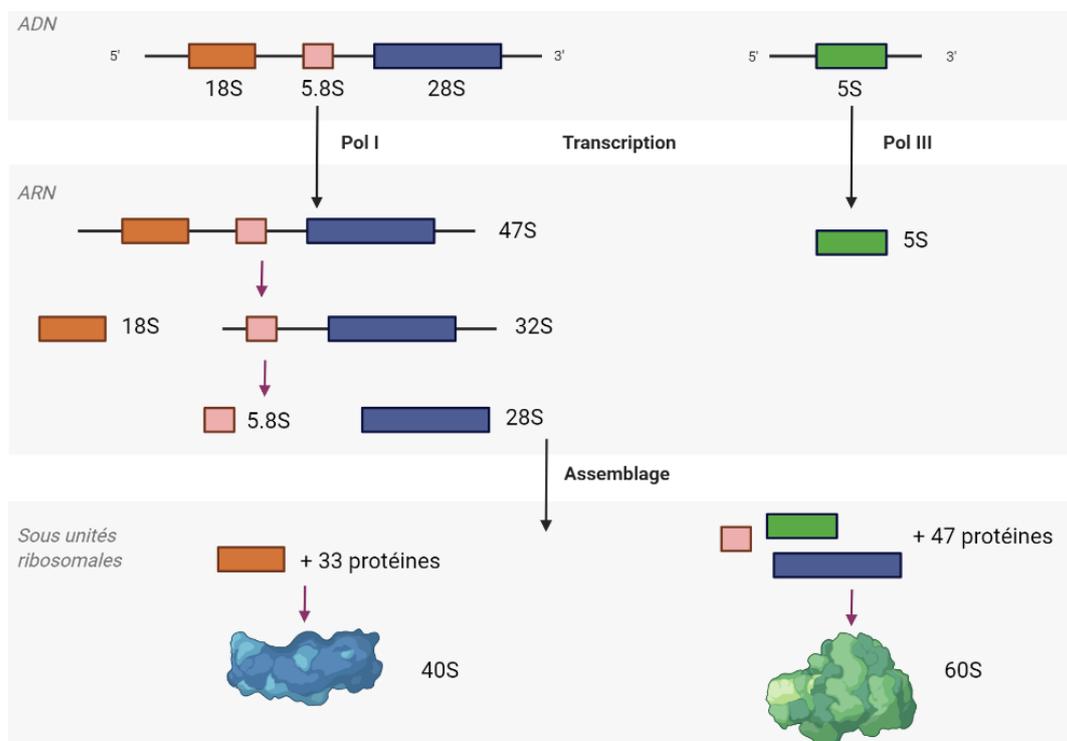


Figure 3 : Voies schématiques de la production des sous unités ribosomiques dans le modèle humain. L'ARN 47S est transcrit par l'ARN polymérase I avant d'être clivé en trois ARNr : 5.8S, 18S et 28S. Le 5S est transcrit par l'ARN polymérase III. Le 18S est ensuite assemblé à 33 protéines ribosomales pour former la petite sous unité 40S. Les ARNr 5S, 5.8S et 28S sont assemblés à 47 protéines ribosomales pour former la grande sous unité 60S.

Chacune de ces sous unités est un assemblage complexe d'ARN et protéines qui leur confèrent des sites d'activité qui leur sont propres.

### **a. Les protéines ribosomiques**

Le ribosome eucaryote, décrit la première fois chez la levure *S. cerevisiae* [4], est constitué d'environ 80 protéines (Figure 2), ce nombre étant relativement stable, même s'il existe des variations selon les organismes [6]. Des études ont mis en évidence la conservation de 34 protéines ribosomiques (15 dans la petite sous unité et 19 dans la grande sous unité) entre les ribosomes procaryotes et eucaryotes. En revanche, le ribosome eucaryote possède 46 protéines qui lui sont propres, 18 dans la sous unité 40S et 28 dans la sous unité 60S. De nombreuses protéines ribosomiques ont été associées à des fonctions extra-ribosomiques, notamment dans la tumorigenèse et la signalisation immunitaire [7].

Au sein du ribosome, les protéines ont un rôle dans la protection de l'ARNr, l'assemblage du ribosome (P1/P2), la structure du centre peptidyl transférase (uL3) et du centre de décodage (uL11 - uS12), l'appariement de ligands (uL29) et dans la sortie de l'ARNm (uL1) [8–12]. Cependant jusqu'à présent elles ne semblent pas participer à l'activité catalytique du ribosome, rôle endossé par l'ARN ribosomique (ARNr) exclusivement.

### **b. Les ARN ribosomiques**

Les ARN ribosomiques représentent environ 80% de l'ARN total eucaryote [13]. Le ribosome eucaryote est composé de 4 ARN, le 18S (40S) et les 5S, 5.8S et 28S (60S) (Figure 3). Le 18S, 5.8S et 28S sont issus du clivage d'un même ARN, le 47S, transcrit par la polymérase I ; le 5S est transcrit par la polymérase III. Les ARN ribosomiques sont maturés au cours d'un long processus de clivages, repliement, modifications chimiques, avant d'être assemblés aux protéines ribosomiques. Une fois maturés, ils sont de tailles comparables entre la levure et l'humain, sauf pour le 28S (Homme) 1,5 fois plus long que le 25S (levure)[14].

### **c. Les centres d'activité ribosomiques**

En 2009, Ramakrishnan, Steitz et Yonath recevaient le prix Nobel de chimie pour leurs travaux de cristallographie aux rayons X qui ont permis de déterminer la structure du ribosome procaryote à l'échelle atomique ; il faudra finalement attendre 2011 pour avoir enfin celle du ribosome eucaryote [4,15]. Le ribosome est donc composé d'une petite et d'une grande sous unités (Figure 4). Trois sites sont présents à l'interface entre les deux sous-unités dans lesquels l'ARNm, l'ARNt et le peptide en cours de synthèse interagissent : les sites A, P et E. Le site Aminoacyl-(A) est le site d'entrée de l'ARNt aminoacylé. Le site Peptidyl (P) est le site de liaison entre l'acide aminé chargé sur l'ARNt et le peptide en cours de synthèse. Pour finir, le site Exit (E) est le site dans lequel est présent l'ARNt désacétylé avant sa sortie du ribosome.

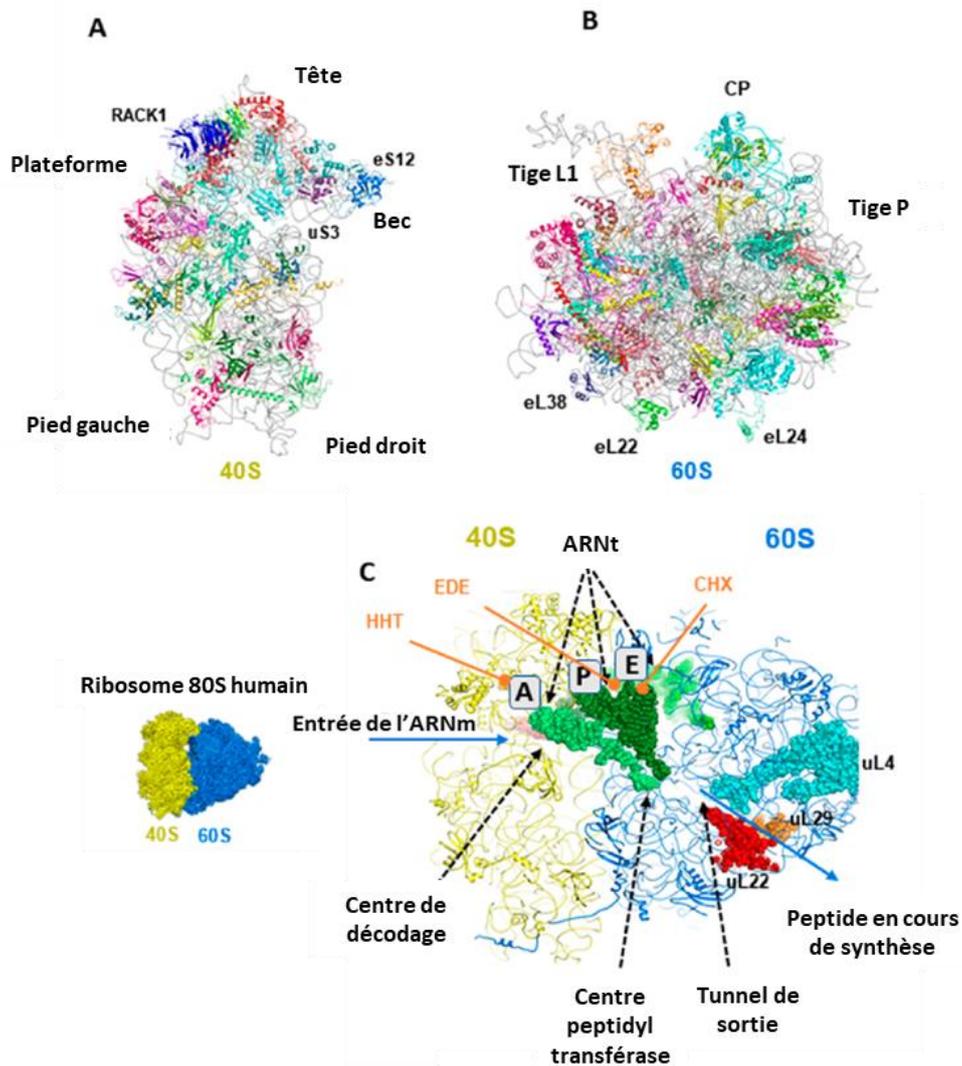


Figure 4 : Analyse structurale du ribosome cytosolique 80S humain. (A,B) Aperçu de la structure des sous-unités ribosomales 40S et 60S du ribosome humain (modèle atomique dérivé de la structure cryo-EM du ribosome humain 80S). (C) Sites fonctionnels typiques sur le ribosome, qui sont directement pertinents pour la liaison des inhibiteurs. Adapté de [16]

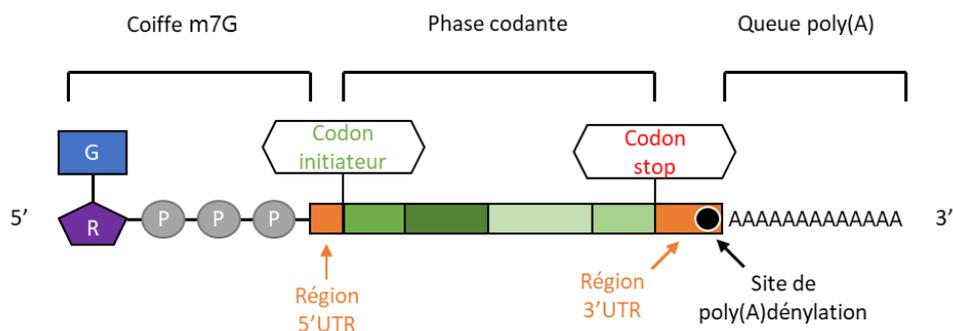


Figure 5 : Structure schématique d'un ARNm. L'ARNm mature comprend : une coiffe à l'extrémité 5', une région non-codante en 5' (5'UTR), une région codante commençant par un codon d'initiation et se finissant au premier codon de terminaison rencontré, une région non-codante en 3' (3'UTR) et enfin une queue poly(A).

De façon individuelle, la petite sous unité 40S est structurellement composée d'une tête, d'une plateforme et d'un corps (Figure 4). La tête possède un bec et le corps une épaule. L'interface entre la tête et le corps compose un sillon dans lequel se loge l'ARN messager en cours de traduction[17–19]. Le sillon entre le bec et l'épaule comporte quant à lui les protéines à activité hélicase permettant l'ouverture des éventuelles structures secondaires présentes sur l'ARNm[20,21].

La grande sous unité 60S contient quant à elle le tunnel de sortie du peptide en cours de synthèse ainsi que le centre peptidyl-transférase [4,22,23] (Figure 4). Le centre peptidyl-transférase catalyse la liaison peptidique entre la chaîne en cours de synthèse portée par l'ARNt au site P et l'acide aminé porté par l'ARNt au site A. Des analyses poussées par rayons X du centre peptidyl-transférase ont permis de révéler qu'il est exclusivement constitué d'ARNr [17], suggérant que le ribosome soit en réalité un « ribosyme » qui serait issu d'un monde purement ARN [22,24].

Le ribosome ainsi structuré permet de traduire les ARNm en protéines. Les ARNm sont ainsi les deuxièmes acteurs de la traduction dont je vais discuter ici.

## **B. L'ARNm**

L'ARNm est un ARN dit codant, c'est-à-dire qu'il est transcrit par l'ARN polymérase II avant d'être traduit afin de produire la protéine correspondante. Il est estimé que le génome humain compte 22 000 gènes codants. Ces gènes sont transcrits en 140 000 ARN différents [25]. Ensuite, environ 65% de ces ARN sont réellement traduits en protéines. Avec une moyenne finale de 4 transcrits traduits par gène codant la transcription et la maturation des ARNm permettent de contribuer à la diversité des protéines cellulaires.

Les ARNm sont maturés tout au long de la transcription afin d'obtenir un ARNm stable et fonctionnel (Figure 5). La maturation est une succession d'étapes qui débute par l'ajout d'une coiffe 7-méthylguanosine triphosphate (m7G) en 5'. Cette coiffe permet de protéger l'ARNm de la dégradation, mais également de recruter le ribosome lors de l'initiation traductionnelle. Le gène en cours de transcription comporte des séquences non codantes, les introns, qui sont excisées de l'ARNm tandis que les régions codantes, les exons, sont réunis par ligation : il s'agit de l'épissage. Le génome humain contient en moyenne 8,8 exons et 7,8 introns par gène, avec une taille moyenne respective de 170 et 5419 nucléotides [26]. Ainsi, les introns représentent plus de 90% de la taille totale d'un ARNm avant épissage. En fin de transcription, c'est l'extrémité 3' de l'ARN qui est modifiée par l'ajout d'une queue poly-adénines (poly(A)). La poly(A)dénylation est réalisée en deux étapes [27]. Elle débute par le clivage de l'ARN, dont le site est principalement défini par le motif consensus hexamérique AAUAAA. Chez l'homme ce motif est présent sur la grande majorité des sites poly(A) et est positionné à environ 21 nucléotides en amont du site de clivage [28].

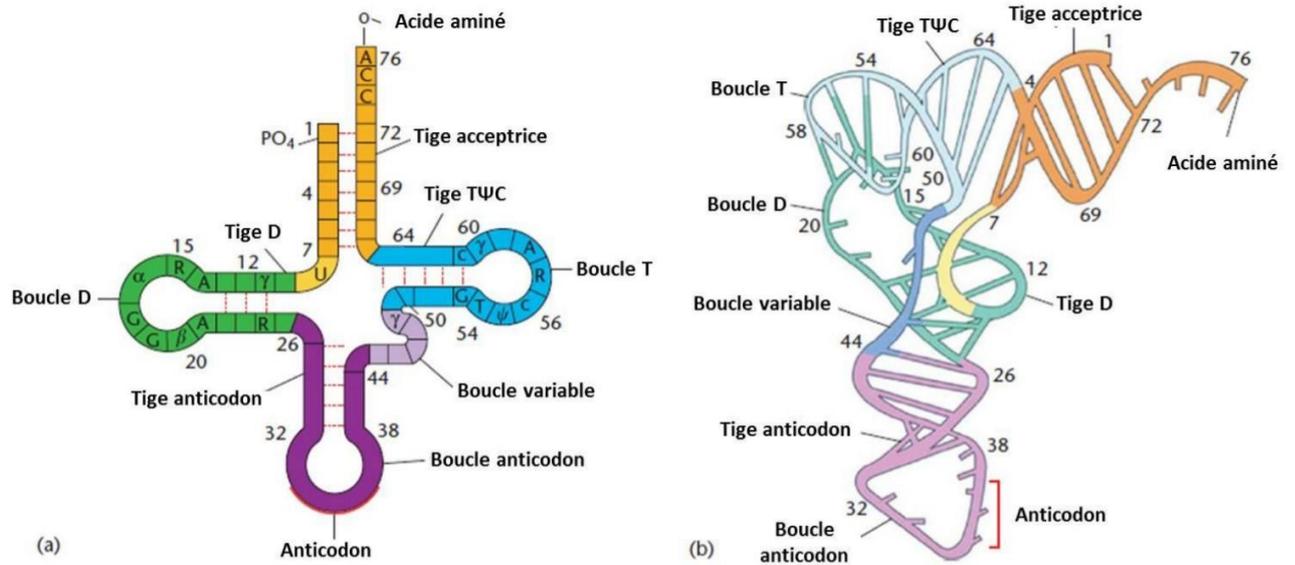


Figure 6 : Structure secondaire et tertiaire de l'ARNt (a) Diagramme de la structure secondaire d'un ARNt schématiquement représentée en forme de trèfle. Les bases universellement conservées sont indiquées par le numéro de leur position. (b) Structure tridimensionnelle en forme de L adoptée par l'ARNt lorsque la tige D se replie sur la tige TΨC. Adapté de la thèse d'O. Bugaud.

Le complexe Cleavage and Poly(A)denylation Factor (CPSF) Poly(A) effectue le clivage endonucléotidique avant l'ajout de la queue poly(A) sur le fragment 5' par la poly(A) polymérase. La protéine poly(A) binding protein nuclear 1 (PABPN1) se lie à la queue poly(A) en cours de synthèse et perturbe l'interaction entre CPSF et la poly(A) polymérase lorsque la queue est d'environ 250 nucléotides, permettant ainsi de contrôler la taille finale de la queue poly(A). Cette étape finale est nécessaire à l'export de l'ARNm mature du noyau vers le cytoplasme et participe à la stabilité de l'ARN ainsi qu'à l'efficacité de traduction [29].

Pour conclure, l'ARNm mature est donc composé de 5' à 3' d'une coiffe m7G, d'une séquence non codante appelée 5'UTR, de la région codante (CDS) qui débute au codon initiateur et termine au codon stop, d'une séquence non codante appelée 3'UTR et de la queue poly(A). Il se trouve que la frontière entre codant et non codant est fine et que de la traduction peut avoir lieu en 5' et 3' UTRs, ce qui sera abordé dans le prochain chapitre. Afin d'être traduit en protéine, l'ARNm doit être lu par le ribosome qui teste l'appariement entre les codons (triplets de nucléotides) du transcrite et les anticodons des ARNt correspondants. Ainsi les ARNt sont les prochains acteurs de la traduction mis en lumière ci-dessous.

### C. L'ARNt

L'ARNt est la seconde classe d'ARN les plus présents dans les cellules eucaryotes en représentant 15% de l'ARN total [13]. Ce n'est pas moins de 500 gènes dans le génome humain qui sont transcrits par l'ARN polymérase III afin de produire les ARNt cellulaires[30]. Décrits pour la première fois en 1965 par R. Holley [31,32], ce sont des ARN porteurs d'un acide aminé qui est ajouté à l'extrémité 3' par des aminoacyl-ARNt synthases hautement spécifiques. Ainsi, ils ont pour fonction d'être les adaptateurs entre les codons et les acides aminés.

La majorité des ARNt cytosoliques font entre 75 et 95 nucléotides dont la structure 3D forme un L inversé [33] (Figure 6). La représentation 2D de cette structure s'apparente à un trèfle dont les trois feuilles sont des tiges boucles (boucle D, boucle anticodon et boucle TΨC) et la tige d'appariement entre les séquences 5' et 3' de l'ARNt (bras accepteur). D'un point de vue fonctionnel, la boucle D, nommée ainsi pour la présence de la base modifiée dihydrouridine, est impliquée dans la reconnaissance de l'ARNt par l' aminoacyl-ARNt synthase [34]. La boucle anticodon contient les trois nucléotides (36, 35, 34) concernés dans l'appariement et la reconnaissance du codon de l'ARNm [35]. La boucle TΨC (pour la présence de thymidine, pseudouridine et cytidine) est documentée pour son rôle dans la reconnaissance par le ribosome ainsi que dans le repliement 3D de l'ARNt [36]. Le bras accepteur, formé par l'appariement des extrémités 5' et 3' de l'ARNt additionné d'un triplet de nucléotides CCA non appariés en 3', est le groupement impliqué dans le chargement de l'acide aminé [37].

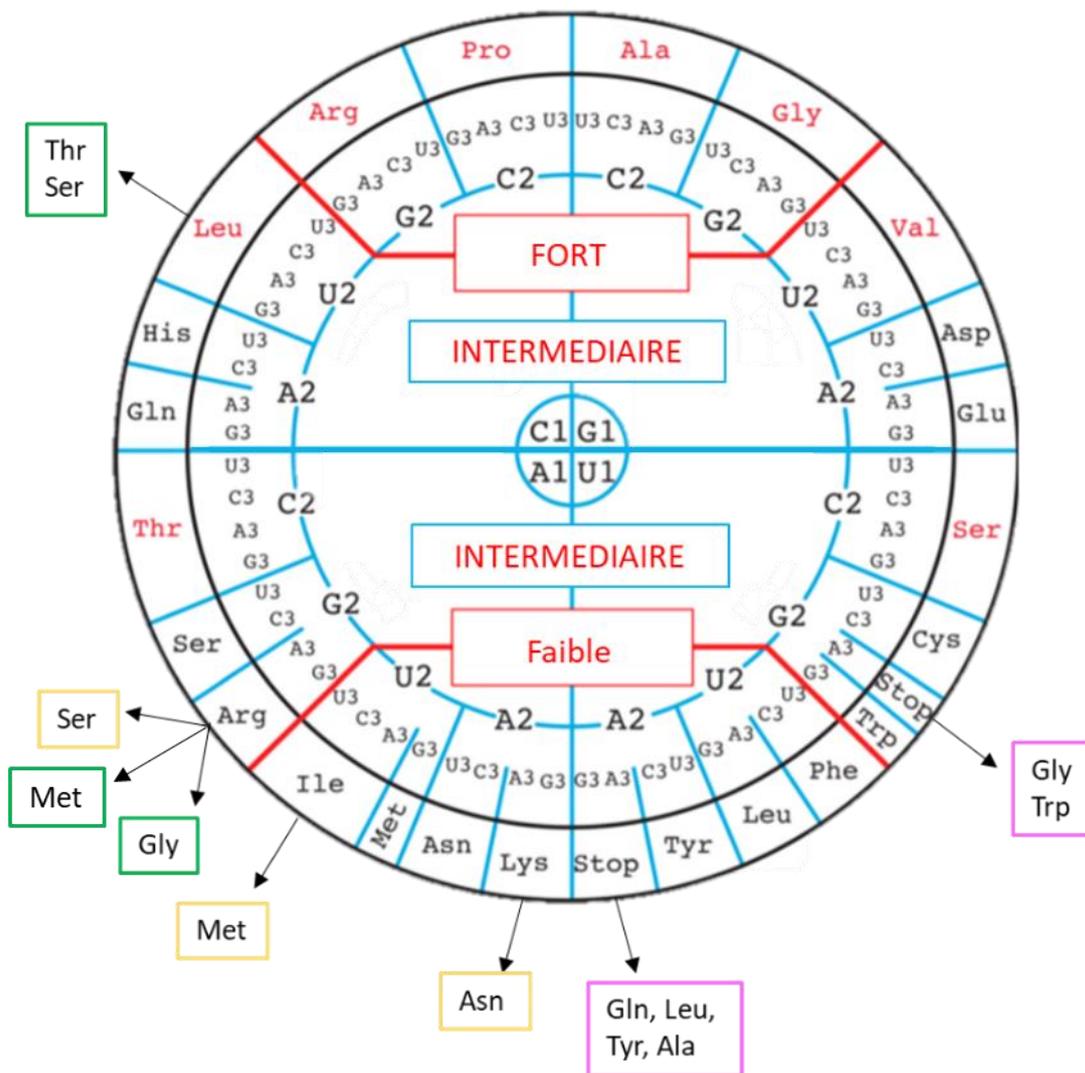


Figure 7 : Représentation énergétique du code génétique standard et du code génétique alternatif. Chaque base est numérotée en fonction de sa position sur le codon. Les codons sont classés dans trois groupes « fort », « intermédiaire » et « faible » en fonction de leur énergie de Turner. Des lignes rouges épaisses séparent les trois principales régions. Les acides aminés du code génétique standard sont indiqués dans le dernier cercle. Les acides aminés réassignés rapportés de codes génétiques alternatifs sont indiqués en dehors du cercle et classés par couleur en fonction de l'origine de la réassignation. Vert : mutation de la séquence d'ARNt et erreur de chargement par l'aminoacyl-ARNt synthase. Jaune : mutation d'une enzyme de modification de l'anticodon de l'ARNt. Rose : mutation de la séquence d'ARNt et mutation d'un facteur de terminaison. Adapté à partir de [38] et [39].

Chaque ARNt possède également une boucle variable de 3 à 21 nucléotides qui favorise la reconnaissance par l'aminoacyl-ARNt synthase mais également la stabilité de l'ARN [36].

Chaque ARNt est nommé en fonction de l'acide aminé qu'il porte. L'aminoacylation est l'étape clé de l'identification de l'ARNt, notamment qu'elle est irréversible. Alors que plus de 260 ARNt cytosoliques constituent le pool d'ARNt, seulement 21 acides aminés sont connus pour être chargés sur les ARNt humains [40]. Ainsi, chaque acide aminé peut être porté par un grand nombre d'ARNt différents qui sont classés soit comme isoaccepteurs (même acide aminé mais séquences différentes) ou isodécodateurs (même acide aminé avec le même anticodon)[41]. Malgré cette abondance d'ARNt par rapport au nombre d'acides aminés, 15 codons dans le transcriptome humain ne possèdent pas d'ARNt avec un anticodon complémentaire [42]. La complexité de la reconnaissance codon-anticodon, ou bien code génétique, est ainsi le quatrième facteur clé de la traduction.

#### **D. Le code génétique**

Durant la traduction, les quatre bases de l'ARNm (Adénine, Uracile, Guanine et Cytosine) sont lues sous forme de triplets (codons) par le ribosome. Chaque codon correspond à l'anticodon d'un ARNt et donc à un acide aminé spécifique : c'est le code génétique (Figure 7). Mis en lumière par Francis Crick et ses collaborateurs en 1961 [43], et complété par Nirenberg en 1965 [44], le code génétique regroupe l'ensemble des règles de correspondance entre codon et acide aminé : chaque ARNm est lu à partir d'un point de départ spécifique dit « codon initiateur », généralement un codon AUG ; les codons ne se chevauchent pas sans pour autant être distingués par des séparateurs ; le code génétique est « dégénéré », c'est-à-dire que plusieurs codons codent pour le même acide aminé, avec 64 codons correspondant aux 20 acides aminés standards. Cette dernière règle exclu les trois codons stop (ou non-sens) UAA (ocre), UAG (ambre) et UGA (opale) qui ne correspondent à aucun acide aminé, indiquant ainsi la terminaison de la traduction. Ceci n'est cependant pas universel et nous verrons par la suite que les codons stop peuvent être réassignés à un acide aminé chez certains organismes, voir organites comme la mitochondrie. Il existe également deux acides aminés, la pyrrolysine et la sélénocystéine, dont les ARNt possèdent un anticodon complémentaire respectivement aux codons UAG et UGA, et qui sont connus pour être incorporés en fonction de la présence de séquences précises au niveau du codon stop. Nous aborderons leur cas dans le chapitre suivant.



Malgré la dégénérescence du code génétique, 15 codons, pourtant traduits en acides aminés, ne possèdent pas d'anticodon complémentaire [42]. On parle alors de la flexibilité du code génétique : un mis-appariement entre la base 34 de l'ARNt et donc de la position 3 du codon peut être toléré et entraîner l'incorporation d'un acide aminé au niveau du codon en question[45]. Il s'agit de la position « wobble » des ARNt. Les ARNt sont alors classés selon le nombre d'appariements que leur anticodon effectue avec le codon : un ARNt cognat effectue 3 appariements sur 3, un ARNt proche-cognat effectue 2 appariements sur 3, et un ARNt non-cognat effectue moins de 2 appariements sur 3.

Longtemps considéré comme universel, le code génétique a ensuite été reconsidéré avec la découverte d'un code génétique alternatif [46] (Figure 7). En effet, au sein même de la cellule les mitochondries utilisent un code génétique alternatif, dans lequel certains codons sont réassignés [47]. On retrouve également une dérive du code génétique cytoplasmique chez certaines espèces [46]. On peut prendre comme exemples les espèces du genre *Candida*, dont le codon CUG, correspond à la Sérine au lieu de la Leucine classique ; chez les ciliés, le codon UGA classiquement non-sens est réassigné en Cystéine (espèce *Euplotes*) ou en Tryptophane (espèce *Blepharisma*) tandis que les codons UAA et UAG sont réassignés en Glutamine (espèces *Paramecium* et *Tetrahymena*).

La conservation du code génétique et l'apparition du code génétique alternatif seraient la conséquence d'une différence dans les forces d'interaction entre les bases du codon et de l'anticodon [38]. En effet, les réassignations de codons les plus fréquentes correspondent aux codons les plus riches en A/U, qui sont les bases avec la plus faible force d'appariement (force de Turner). Au contraire, les bases G/C, qui sont celles avec la plus grande force d'appariement, sont très peu concernées par la réassignation : peu d'exemples de codons commençant par un C<sub>1</sub> sont connus pour être réassignés, et aucun exemple de codon réassigné commençant par un G<sub>1</sub> a été mis en évidence. Les principales causes de réassignations sont la mutation de la séquence de l'ARNt, la mutation de l'aminoacyl-ARNt synthase, la mutation d'une enzyme de modifications de l'ARNt et la mutation d'un facteur de terminaison.

## **2. Les étapes de la traduction**

La traduction est un processus qui se déroule en quatre étapes principales : l'initiation, l'élongation, la terminaison et le recyclage du ribosome. Nous allons ainsi retracer d'un point de vue moléculaire l'ensemble de ces étapes clés de la production protéique en débutant par l'initiation.

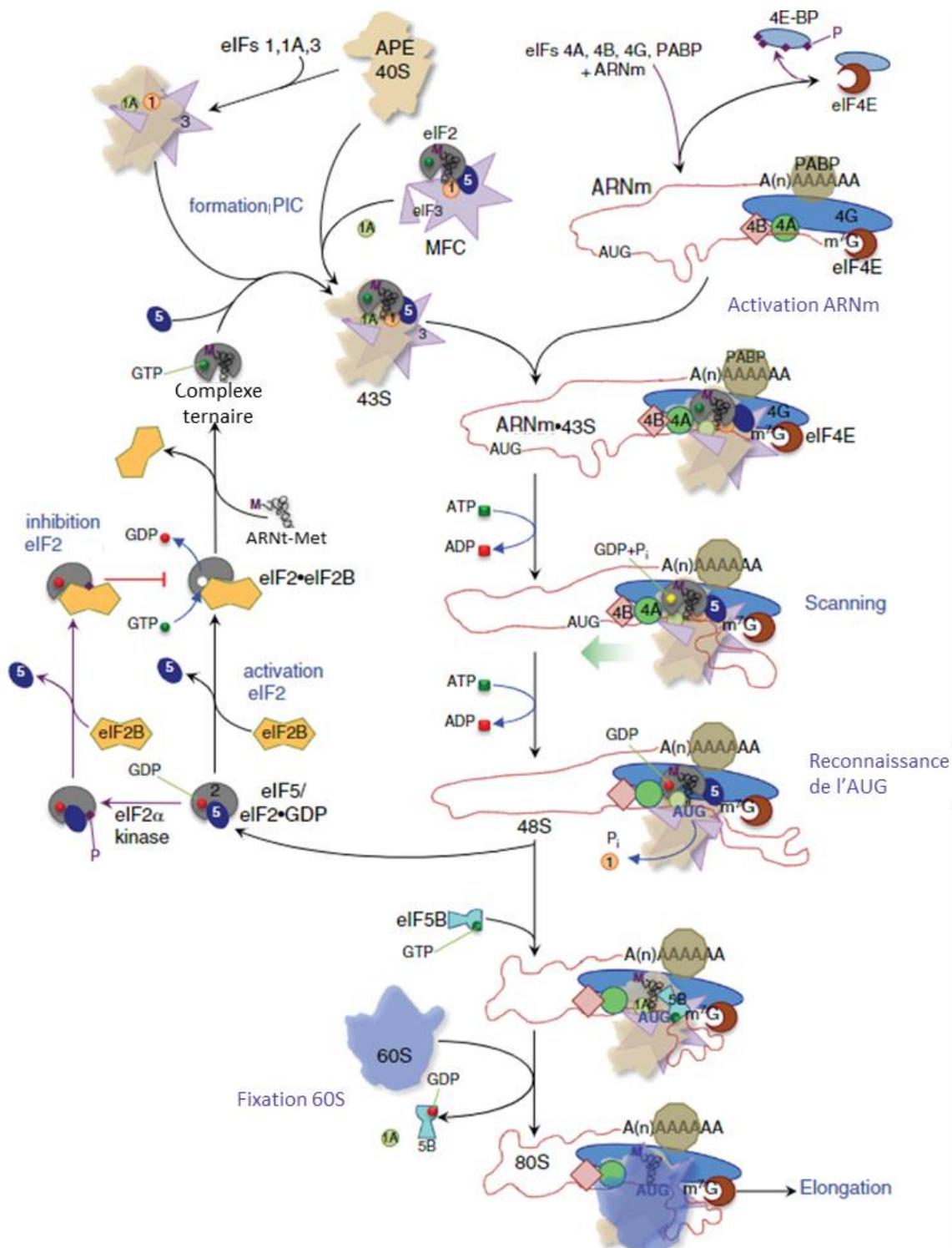


Figure 8 : La voie générale d'initiation de la traduction chez les eucaryotes. La voie de recrutement de l'ARNt initiateur au codon AUG de l'ARNm dans le contexte d'un ribosome 80S (en bas à droite) est représentée comme une série d'étapes majeures, étiquetées en bleu, reliées par des flèches noires. Les complexes et facteurs d'initiation sont étiquetés avec du texte noir et les réactions d'hydrolyse des nucléotides/libération du phosphate inorganique sont représentées par des flèches bleues. La large flèche verte indique la direction du scanning vers le codon AUG. Les réactions de régulation conduisant à l'inhibition de eIF2 et eIF4E sont représentées par des flèches rouges et prunes. Adapté de [48].

## A. L'initiation

L'initiation est la première étape de la traduction, qui est par ailleurs la plus régulée des quatre. Les cellules eucaryotes ont mis en place un système complexe faisant intervenir de nombreux facteurs afin de recruter le ribosome précisément au niveau du codon initiateur, première garantie du maintien du cadre de lecture de l'ARNm. Cette initiation, généralement dépendante de la coiffe m<sup>7</sup>G, peut présenter différentes variantes allant jusqu'à l'absence d'utilisation d'une coiffe m<sup>7</sup>G remplacée par le recrutement du ribosome via un site d'entrée interne : l'Internal Ribosome Entry Site (IRES).

### a. L'initiation dépendante de la coiffe m<sup>7</sup>G

Afin de garantir la fixation de la petite sous unité sur la coiffe m<sup>7</sup>G, l'initiation débute par la formation d'un complexe ternaire (TC) composé de l'ARNt-Met initiateur et d'eIF2 associé à un GTP (Figure 8). La protéine G eIF2 tient son affinité à l'ARNt<sub>Met</sub> de la présence du GTP [49]. La sous unité 40S libre est alors liée aux facteurs eIF1, 1A, 3 et 5, et le TC : le complexe 43S de pré-initiation (PIC) est formé. De nos jours, deux voies d'assemblage du PIC sont connues : eIF1, 1A et 3 peuvent être issus du recyclage d'un 80S et s'associer au 40S avec de recruter eIF5 et le TC [50]; il est également possible que le TC et eIF1, 3 et 5 forment un complexe indépendant (MFC) avant de se lier à la 40S et recruter eIF1A [51,52]. La présence d'eIF1 et 1A, au niveau des sites P et A respectivement, permet d'ouvrir la sous unité 40S entre la « tête » et le « corps » afin de faciliter le passage de l'ARNm et ainsi le scanning : la petite sous unité est en conformation « ouverte ».

En parallèle de la formation du PIC libre, la coiffe m<sup>7</sup>G de l'ARNm est reconnue par le complexe eIF4F. Ce dernier est composé d'eIF4E 4G et 4A. eIF4E est impliqué dans la fixation sur l'ARNm d'eIF4F. eIF4G a pour rôle d'interagir avec les protéines PABP liées à la queue poly(A), ce qui permet d'obtenir une boucle « fermée » de l'ARNm par rapprochement du 5' avec le 3' [53,54]. Un des rôles associés à cette boucle serait de favoriser la liaison avec le PIC [55,56]. Le dernier rôle de ce complexe eIF4F est l'ouverture des structures secondaires de l'ARNm par l'ARN hélicase eIF4A, permettant ainsi la fixation du PIC sur l'ARNm. Ce dernier se fixe alors sur la coiffe m<sup>7</sup>G, formant ainsi le 48S, et le scanning peut débuter. Une étude récente chez l'Homme a permis de mettre en évidence que le PIC reste associé à la coiffe m<sup>7</sup>G durant la phase de scanning [57]. Cela a pour conséquence que l'initiation, contrairement à l'élongation, ne peut être réalisée que par un seul ribosome à la fois sur un même transcrit, ce qui implique que la longueur du 5' UTR soit un facteur limitant de l'efficacité d'initiation de la traduction.



Le PIC comportant l'ARN-Met initiateur va être capable d'identifier le codon initiateur AUG lorsqu'il est présent au niveau du site P de la sous unité 40S. Il est également connu que cette reconnaissance dépend d'autres facteurs liés à l'ARNm : la présence d'une séquence consensus au niveau du codon AUG (5' -(A/G)NNAUGG-3') et l'absence de structures secondaires à proximité. La reconnaissance du codon AUG provoque l'hydrolyse du GTP associé à eIF2 par eIF5. Cela entraîne la dissociation d'eIF1 et ainsi le passage de la conformation « ouverte » à « fermée » du PIC. Le PIC est ainsi stabilisé sur l'ARNm et eIF2-GDP et eIF5 se dissocient du complexe, le scanning est terminé [58]. En parallèle le complexe protéique SBDS-eIF1 dissocie le facteur eIF6 de la 60S qui l'empêchait de s'associer à la sous unité 40S [59,60]. eIF5B, associé à la sous unité 60S, collabore alors avec eIF1A afin de former le ribosome 80S. Une fois les deux sous unités assemblées, eIF4A est relargué, permettant au ribosome de débiter la phase d'élongation.

Alors que le groupe N7-méthyle de la coiffe joue un rôle majeur dans l'initiation de la traduction, de nouvelles études mettent en évidence une diversité de compositions dans les coiffes m<sup>7</sup>G cellulaires [61]. Les coiffes m<sup>7</sup>G sont soumises à d'autres modifications chimiques réalisées par des méthyltransférases. Les méthylations sont déposées en position 2'-O des nucléotides présents en position +1 ou +2 de la coiffe m<sup>7</sup>G, conduisant respectivement aux formes « coiffe 1 » et « coiffe 2 ». Lorsque le nucléotide initiateur de la transcription est une adénosine, sa position N<sup>6</sup> est couramment modifiée par l'ajout supplémentaire d'un méthyle. Bien qu'il s'agisse d'une thématique émergente, plusieurs études mettent déjà en avant que l'identité du premier nucléotide de transcription ainsi que son taux de méthylation ont un impact entre autres sur la stabilité et la traduction du transcrit [62–64].

En plus de l'initiation dépendante de la coiffe m<sup>7</sup>G canonique décrite ci-dessus, il existe chez les eucaryotes une initiation dépendante de la coiffe m<sup>7</sup>G non canonique. On la retrouve notamment dans le cas d'ARNm ayant un court 5'UTR. Il est estimé que cette initiation concerne environ 4% des transcrits chez les mammifères [65], des transcrits notamment dits de « ménage », à qui elle confère une résistance à l'inhibition globale de la traduction en réponse à un stress énergétique et une réponse particulière aux rythmes circadiens et alimentaires [66,67]. Elle dépend de la présence de la séquence consensus TISU (SAASAUGGCGGC) présente au niveau de l'extrémité 5' de l'ARNm. Malgré la dépendance vis-à-vis de la présence d'une coiffe m<sup>7</sup>G, cette initiation non canonique se fait indépendamment de l'étape de scanning [68]. Alors que les mécanismes sont encore méconnus, il est acquis que la formation du complexe 48S serait immédiatement suivie du positionnement de la sous unité 60S sur ce complexe au niveau du codon AUG et du détachement de eIF4F de la coiffe [69].

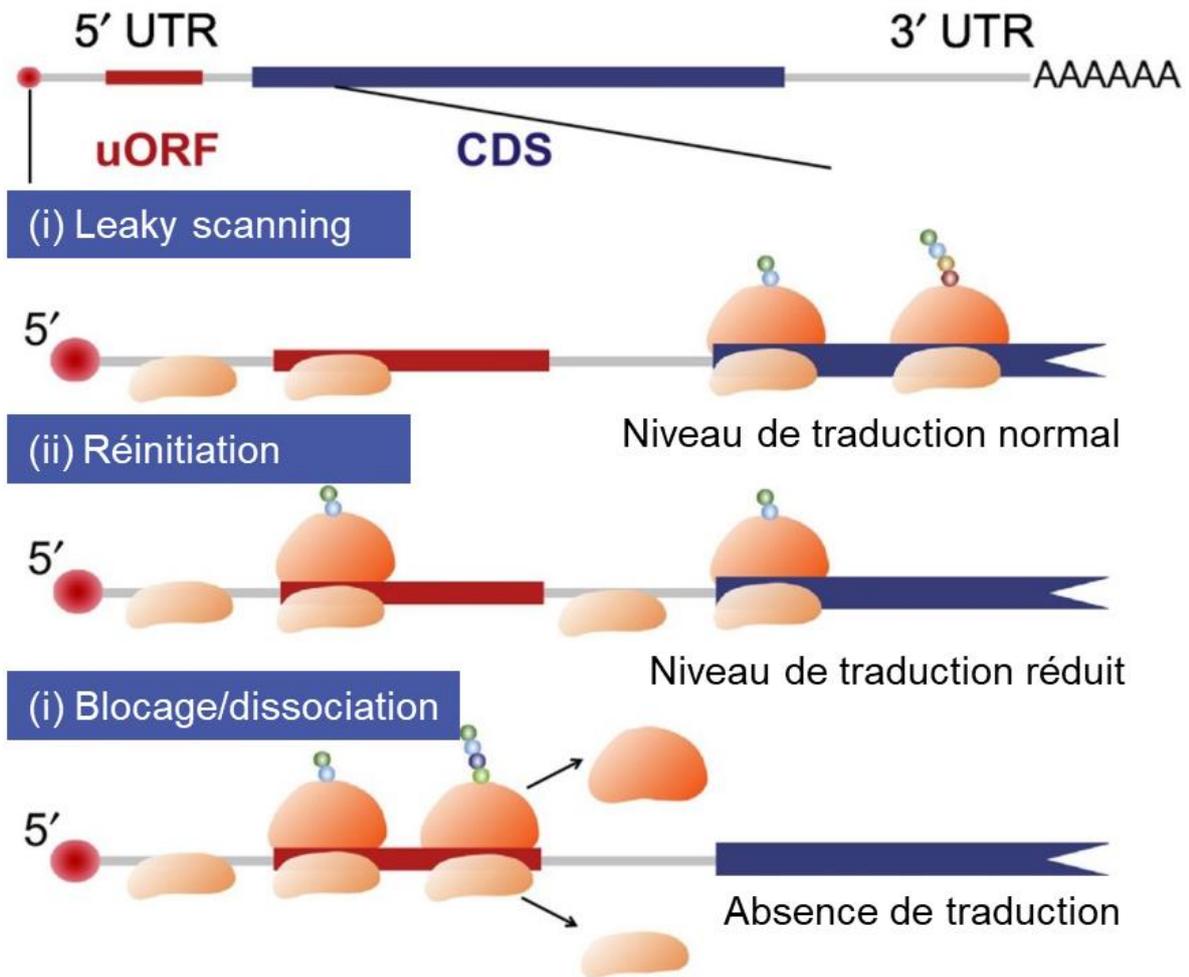


Figure 9 : Régulation traductionnelle médiée par les cadres de lecture ouverts en amont de l'AUG (uORF). Les trois scénarios possibles sont présentés : (i) Le PIC ne parvient pas à reconnaître le codon de départ de l'uORF et continue à chercher le codon de départ suivant ("leaky scanning"). Dans ce cas, la traduction de la séquence codante en aval (CDS) n'est pas affectée. (ii) Le PIC reconnaît avec succès le codon de départ de l'uORF et s'assemble en un ribosome complet pour la traduction. Après la terminaison au codon stop de l'uORF, la sous-unité 40S reste liée à l'ARNm et réinitialise la traduction au codon de départ du CDS ("réinitiation"). La traduction de l'uORF va interférer avec les autres ribosomes de balayage, ce qui entraîne une réduction de la traduction du CDS. (iii) Lors de la fin de la traduction au codon stop de l'uORF, le ribosome peut bloquer et déclencher une digestion de l'ARNm médiée par la voie NMD ("blocage"), ou se dissocier de l'ARNm ("dissociation") sans traduire la CDS. Adapté de [70].

Cela permettrait au codon AUG d'être correctement positionné au site P du ribosome et ainsi de débiter l'élongation.

Alors que l'initiation permet généralement de débiter la traduction au niveau du codon initiateur de la CDS, il est possible que le ribosome débiter l'élongation en amont, avec la traduction d'une séquence dans le 5'UTR appelée « upstream Open Reading Frame (uORF) »[55] (Figure 9). Chez les mammifères, il a été démontré que plus de 45% des ARNm contiennent au moins une uORF, qu'elles sont généralement courtes (<30 nucléotides), fréquemment traduites, et impliquées dans la répression de l'expression de la CDS [71–74]. La quasi-totalité des uORFs présentes dans les ARNm mammifères semble être permissive à la réinitiation, ce qui suggère que ces processus n'impliquent pas de séquences spécifiques. Le mécanisme de réinitiation est encore mal connu, mais des travaux ont mis en évidence l'implication du complexe eIF4F et du complexe stable eIF4G-eIF3-40S [75]. Il a été démontré également que les eIF restent associés au ribosome 80S en début d'élongation, ce qui permettrait d'expliquer la réinitiation facilitée suite à une uORF courte [57]. La présence d'une uORF peut réguler l'expression de la CDS par un mécanisme appelé « leaky scanning », qui se produit lorsque plusieurs codons d'initiation AUG sont présents sur l'ARNm et que le premier AUG se trouve dans un contexte de séquence d'initiation de force intermédiaire. Il dépend de plusieurs facteurs trans tels que les niveaux cellulaires des facteurs d'initiation eIF1 et eIF5 [76–79]. Au cours du leaky scanning, soit les ribosomes initient alors sur le premier AUG de l'uORF, soit parfois le dépassent pour initier sur l'AUG de la CDS en aval [80–82]. Si les deux codons AUG sont dans le même cadre l'un par rapport à l'autre, cela conduit à la traduction de deux isoformes de la même protéine, l'une avec une extension N-terminale. Si les deux codons AUG font partie de deux ORFs qui se chevauchent cela conduit à l'expression de deux protéines différentes à partir du même ARNm. La fréquence du leaky scanning affecte ainsi les quantités relatives des isoformes de protéines courtes et longues ou des deux protéines fabriquées. Chez l'Homme, la protéine CEBPA (CCAAT/enhancer-binding protein alpha) est un très bon exemple de leaky scanning [83]. Dans les cellules saines, une production équilibrée de deux isoformes de 30 et 42 kDa est maintenue par leaky scanning. L'isoforme de 30 kDa, contrairement à l'isoforme 42 kDa, n'inhibe pas la prolifération cellulaire [84]. Dans les cellules cancéreuses un déséquilibre entre les deux isoformes a été retrouvé avec une augmentation du leaky scanning, conférant ainsi un avantage à ces cellules [85].

Même si l'initiation dépendante de la coiffe m<sup>7</sup>G est l'initiation la plus courante *in vivo*, de nombreuses études portent sur la capacité des cellules eucaryotes à utiliser une autre initiation indépendante de la présence de coiffe m<sup>7</sup>G : l'initiation dépendante d'IRES.

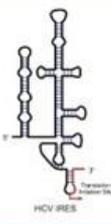
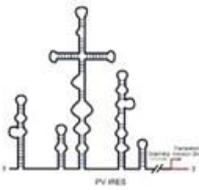
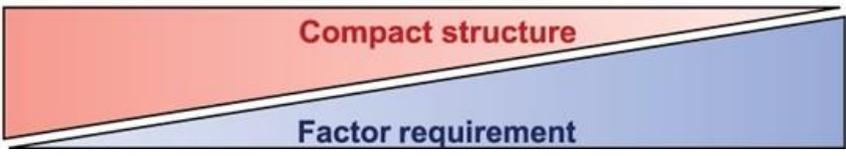
	Group I	Group II	Group III	Group IV
IFs	None	eIF2 eIF3	eIF2 eIF3 eIF4A eIF4B eIF4G	eIF2 eIF3 eIF4A eIF4G
ITAFs	None	None	PTB ITAF <sub>45</sub>	PTB PCBP1 PCBP2
40S recruitment	Direct binding to 40S ribosomal subunit, no IFs requirement	Direct binding to 40S ribosomal subunit, 40S does not scan	Binding to 40S ribosomal subunit through ITAFs, 40S does not scan	Binding to 40S ribosomal subunit through ITAFs, 40S scanning to the downstream AUG
Examples	CrPV, PSIV, TSV	HCV, CSFV, PTV-1	EMCV, FMDV, TMEV	PV, rhinovirus
Secondary Structure				
				

Figure 10 : Les quatre groupes d'IRES virales. Les IRES virales peuvent être classées en quatre groupes en fonction de leur structure et de l'exigence des ITAF et des IF. En général, les IRES à structure plus compacte nécessitent moins de facteurs protéiques auxiliaires. Issu de [86].

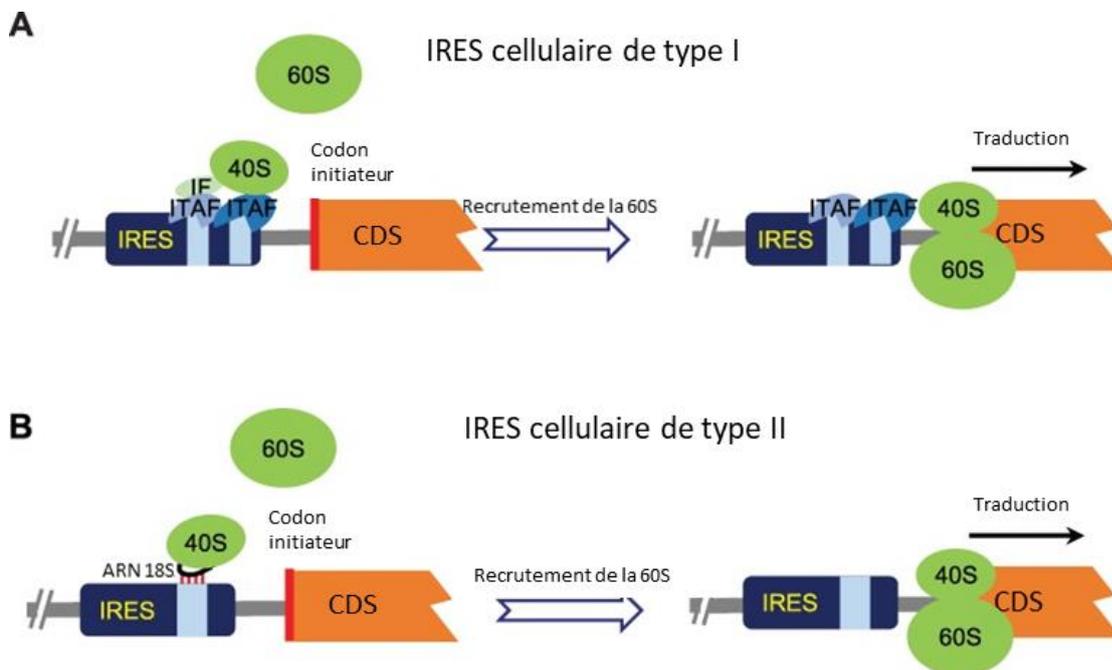


Figure 11 : Mécanismes d'initiation de la traduction piloté par les IRES cellulaires. (A) IRES cellulaire de type I. (B) IRES cellulaire de type II. Les IRES cellulaires de type II contiennent les éléments cis courts qui peuvent s'apparier à l'ARNr 18S. Par conséquent, ces IRES cellulaires peuvent interagir directement avec la sous-unité ribosomique 40S et recruter la sous-unité ribosomique 60S pour commencer la traduction. Adapté de [86].

## **b. L'initiation dépendante d'IRES**

Découverte en 1988 par Pelletier et Sonenberg [87], l'IRES est par définition un élément ARN qui permet le recrutement du ribosome dans la région interne de l'ARNm et cela de manière indépendante de la coiffe m<sup>7</sup>G. Les premières découvertes d'initiations dépendantes d'IRES concernent de nombreux virus, avec des séquences IRES très différentes mais des structures secondaires parfois très similaires [88] (Figure 10). Elles font intervenir différents facteurs, les IRES-Transacting Factor (ITAF) qui sont propres à l'initiation dépendante d'IRES et les Initiation Factors (IF) qui interviennent également dans l'initiation dépendante de la coiffe m<sup>7</sup>G [89,90]. Les initiations virales dépendantes d'IRES ont alors été classées en 4 classes différentes, en fonction des structures secondaires des IRES et des facteurs ITAF et IF qu'elles nécessitent [91]. Les IRES du groupe I peuvent initier la traduction à partir d'un codon initiateur non-AUG, sans ITAF et sans le TC eIF2-GTP/ARNt<sub>Met</sub>. Elles ont une structure secondaire compacte qui comprend trois pseudo-noeuds nommés PKI, PKII et PKIII, et de multiples tiges boucles. Elles peuvent interagir directement avec la sous unité 40S de manière efficace et sont connues pour avoir une forte activité [92–94]. Les IRES du groupe II nécessitent généralement l'aide de plusieurs IF, notamment le TC et eIF3, mais n'ont pas besoin d'intervention d'ITAF pour se lier au 40S [95–97]. Contrairement aux groupes I et II, les deux derniers ont besoin d'ITAF et de nombreux IF pour se lier à la petite sous unité (eIF2, eIF3, eIF4A, eIF4B et eIF4G) [98–100]. Le groupe III nécessite l'intervention des ITAF PTB (Pyrimidine Tract Binding protein) et ITAF<sub>45</sub>, alors que ce sont les ITAF PTB et PCBP 1 et 2 (Poly(rC)-Binding Protein) qui interviennent dans le groupe IV. La principale différence entre le groupe III et IV est l'étape de scanning : comme les groupes I et II, le groupe III initie généralement sans scanning [98], contrairement au groupe IV [101]. Cette différence reste parfois contestée avec des IRES du groupe IV qui sont capables de recruter directement la sous unité 40S [102].

Alors que les IRES virales ont constitué le point de départ des recherches dans le domaine, il est maintenant acquis que dans des conditions de stress cellulaire avec une diminution de l'initiation dépendante de la coiffe m<sup>7</sup>G, des ARNm cellulaires continuent à être exprimés via l'initiation dépendante d'IRES [103] (Figure 11). Les IRES cellulaires sont classées en deux groupes. Les IRES de type I interagissent avec les ribosomes par le biais d'ITAF [104,105] tandis que les IRES de type II contiennent un court élément cis semblable aux séquences Shine-Dalgarno procaryotes, qui s'apparie avec l'ARNr 18S pour recruter des ribosomes [106]. Alors que l'étude des IRES cellulaires est plus compliquée que celle des IRES virales à cause de leur faible niveau d'activité, il a été estimé récemment qu'environ 10% des ARNm de mammifères pourraient initier de manière dépendante d'IRES [107].

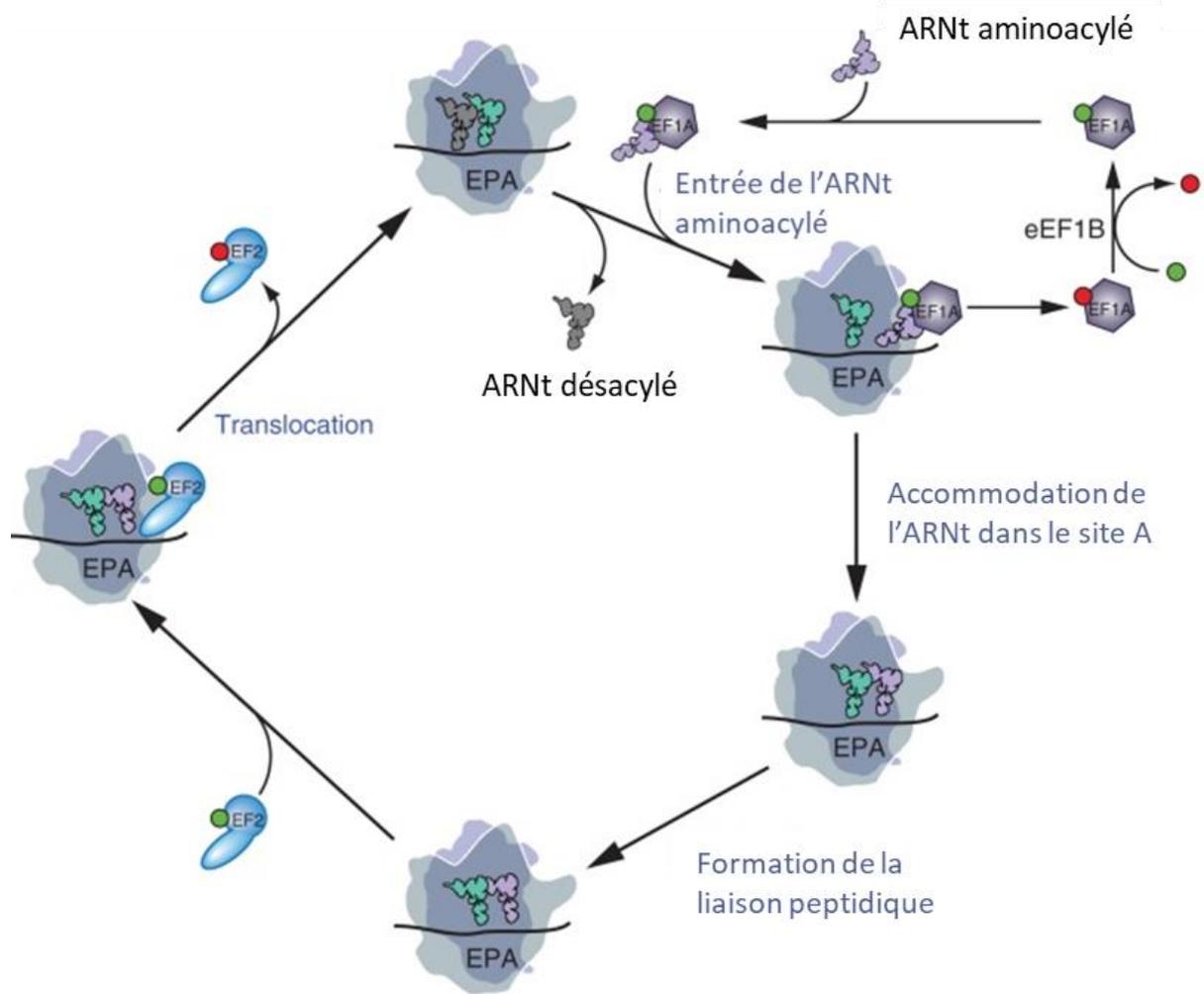


Figure 12 : La phase d'élongation de la traduction eucaryote. Dans ce modèle, la grande sous-unité ribosomique est rendue transparente pour visualiser les ARNt, les facteurs et les ARNm se liant au centre de décodage à l'interface entre la grande et la petite sous-unité et les ARNt interagissant avec le centre de la peptidyl transférase dans la grande sous-unité. Le GTP est représenté par une boule verte et le GDP par une boule rouge. Adapté de [108].

Une fois le ribosome assemblé au niveau du codon initiateur, l'initiation s'achève et laisse place à l'élongation, deuxième étape de la traduction que nous allons à présent aborder.

## **B. L'élongation**

L'élongation est la deuxième étape de la traduction. Le ribosome est présent au niveau du codon initiateur avec l'ARNt<sub>Met</sub> au site P du ribosome, et les facteurs d'initiation ont été relargués. Le ribosome 80S débute alors l'élongation permettant la synthèse de la protéine (Figure 12). L'élongation, contrairement à l'initiation, fait intervenir peu de facteurs appelés eukaryotic Elongation Factors (eEF). Elle repose sur deux étapes principales : la discrimination de l'ARNt correct et le transfert de l'acide aminé de l'ARNt au peptide. Ainsi nous allons commencer par la description de la reconnaissance et l'accommodation de l'ARNt dans le ribosome.

### **a. Reconnaissance et accommodation de l'ARNt**

L'ARNm doit être lu codon par codon, du codon initiateur jusqu'au codon stop, c'est-à-dire l'ensemble de la CDS sans que le ribosome change de cadre de lecture (Figure 12). A la fin de l'initiation l'ARNt<sub>Met</sub> initiateur est présent au niveau du site P du ribosome, apparié avec le codon AUG. Les sites A et E du ribosome sont libres. Le second codon de la CDS se situe alors dans le site A du ribosome. L'élongation débute ainsi par l'entrée de l'ARNt aminoacylé correspondant à ce second codon. Cet ARNt est associée au facteur d'élongation eEF1A-GTP. Lorsque l'ARNt est cognat du codon, les interactions d'appariement des bases entre le codon et l'anticodon présents dans le site A déclenchent l'hydrolyse du GTP par eEF1A : c'est la reconnaissance de l'ARNt. L'hydrolyse du GTP provoque la libération d'eEF1A-GDP et l'accommodation de l'ARNt dans le site A. Par définition, l'accommodation est un changement de conformation du de l'ARNt lui permettant de passer du site A au site P après sa sélection. La reconnaissance et l'accommodation de l'ARNt au sein du site A ont été fortement étudiées afin de déterminer l'origine des règles de décodage par le ribosome. Ces règles reposent sur l'encombrement stérique et la géométrie globale de l'appariement codon-anticodon, et font intervenir des changements de conformation du ribosome qui seront discutés dans le chapitre portant sur le maintien de la fidélité de la traduction. Une fois l'ARNt accommodé au site A, il y a formation de la liaison peptidique entre l'acide aminé entrant et le peptide en cours de synthèse.

### **b. Formation de la liaison peptidique et translocation**

Cette étape se produit au niveau du centre peptidyl-transférase de la sous unité 60S, rapidement après la reconnaissance et l'accommodation de l'ARNt aminoacylé (Figure 12). Le facteur eIF5A, présent au niveau du site E, interagit avec le bras accepteur de l'ARNt peptidyl présent au niveau du site P du ribosome.



eIF5A est supposé promouvoir la formation de la liaison peptidique en induisant un positionnement favorable des ARNt du site P et A [109–112]. Le facteur eIF5A est notamment indispensable afin que le ribosome n'effectue pas une pause en présence de proline, glycine ou d'acides aminés chargés au niveau du centre peptidyl transferase ou à l'entrée du tunnel de sortie [113]. La formation de la liaison peptidique provoque le transfert du néopeptide sur l'acide aminé de l'ARNt présent au niveau du site A. Cette étape conduit à la translocation de l'ensemble des ARNt d'un site ribosomique à un autre, déterminée structuralement par trois phases de rotation des ARNt présents [114]. Après la formation de la liaison peptidique, le ribosome se situe alors dans la première phase de rotation : l'ARNt peptidyl nouvellement formé se trouve dans le site A et l'ARNt désacylé dans un état hybride entre les sites P et E, son anticodon toujours apparié à l'ARNm dans le site P et le bras accepteur dans le site E. Vient ensuite l'état de rotation 2, dans lequel l'ARNt désacylé est toujours dans l'état hybride P/E, alors que l'ARNt peptidyl est repositionné dans un état hybride entre le site A et P, son anticodon apparié dans le site A et le peptide attaché au bras accepteur dans le site P [115,116]. La translocation de l'ARNm et des ARNt vers les sites E (ARNt désacylé) et P (ARNt peptidyl) est favorisée par le facteur d'élongation eEF2, dont le rôle, déduit par similarité avec son orthologue bactérien eFG, serait de défaire l'interaction entre l'hélice h44 de l'ARNr 18S et le duplex codon-anticodon dans le site A [117,118]. Après la translocation, un ARNt désacylé occupe alors le site E, l'ARNt peptidyl est dans le site P et le site A est libre. La libération de l'ARNt désacylé du site E du ribosome reste encore un processus mal compris chez les eucaryotes, avec de premières pistes chez la levure indiquant que ce serait le facteur eEF3 qui faciliterait les étapes tardives de la translocation en stabilisant un état de « non-rotation » du ribosome, et en induisant une conformation ouverte de la tige L1 permettant la libération de l'ARNt du site E [119]. La translocation laisse place à un nouveau codon dans le site A du ribosome, les phases d'élongation vont alors se succéder, le peptide en cours de synthèse sortant via le tunnel de sortie. Ce tunnel a une longueur d'environ 10 nm et permet le repliement de la chaîne peptidique naissante en structures secondaires dans un environnement protégé [120,121]. A la sortie de la grande sous-unité, l'extrémité N-terminale de la protéine partiellement repliée rencontre divers facteurs de maturation tels que des chaperonnes de repliement, des protéines de reconnaissance du signal d'adressage et des enzymes de modifications co-traductionnelle [122]. L'élongation continue ainsi jusqu'à l'arrivée du codon stop dans le site A du ribosome.

### **C. La terminaison**

La terminaison est la dernière étape de la traduction qui implique le décodage d'un codon. Elle est réalisée par un complexe de facteurs de terminaison appelés eukaryotic Release Factors (eRF). Nous allons à présent débiter par les mécanismes impliqués dans la reconnaissance du codon stop.

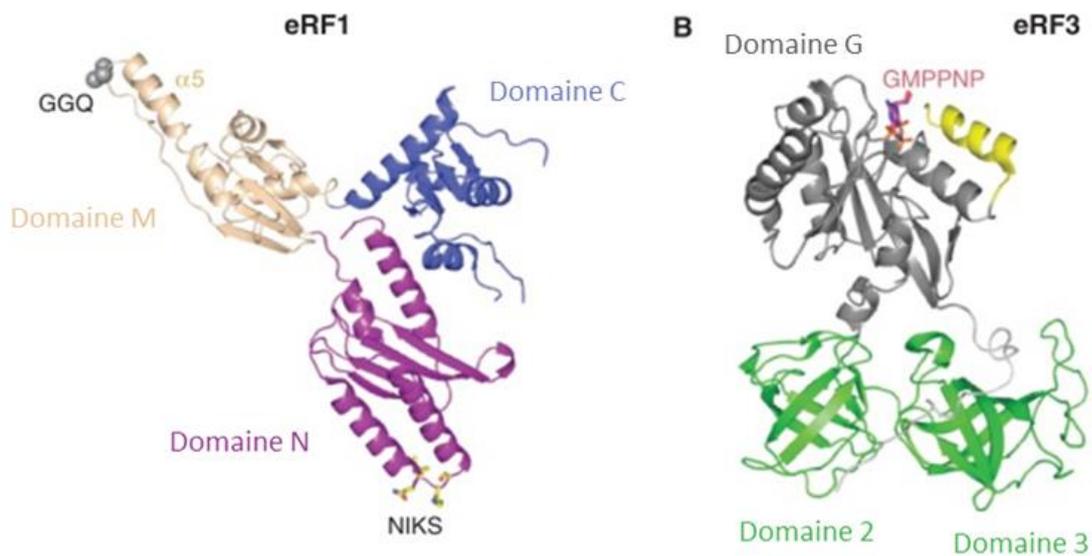


Figure 13 : Terminaison de la traduction chez les eucaryotes. Représentations en ruban de (A) l'eRF1 humain (banque de données sur les protéines [PDB] : 1DT9), avec les atomes C  $\alpha$  du motif GGQ dans le domaine M et le motif NIKS dans le domaine N montrés dans des modèles CPK et bâton, respectivement, et (B) l'eRF3 de *Schizosaccharomyces pombe* (acides aminés 215-662, et donc sans le domaine amino-terminal non conservé) (PDB : 1R5B), avec le GMPPNP lié montré dans un modèle bâton. Adapté de [123].

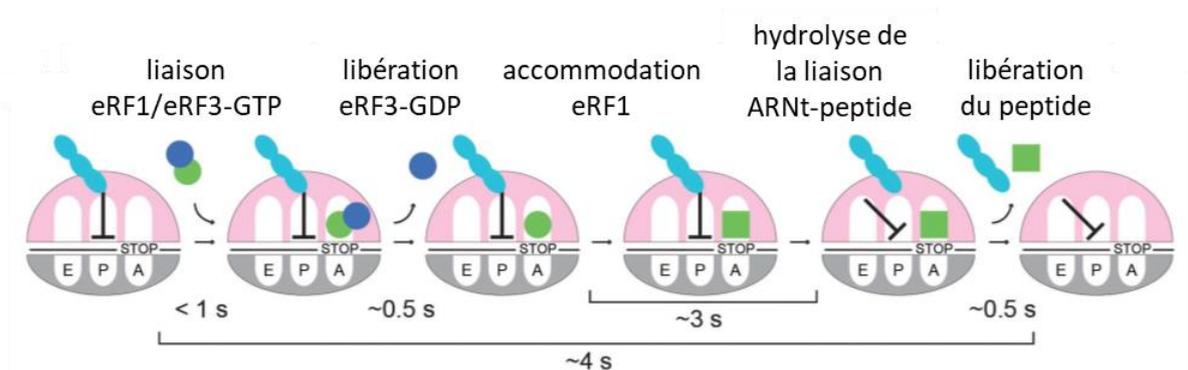


Figure 14 : Les différentes étapes de la terminaison eucaryote et leurs cinétiques respectives. Le peptide est représenté en bleu clair. Le facteur de terminaison eRF3 est représenté en bleu foncé. Le facteur de terminaison eRF1 est représenté en vert avant (rond) et après (carré) accommodation dans le ribosome 80S. Les temps sont indiqués en secondes. Adapté de [124].

### **a. Reconnaissance du codon stop**

La terminaison a lieu lorsqu'un codon stop est présent au niveau du site A du ribosome. Elle est assurée par un complexe ternaire eRF1-eRF3-GTP [125] (Figure 13). eRF1, le facteur de terminaison de classe I, possède notamment un domaine terminal aminé (N) qui mime le bras anticodon de l'ARNt et est capable de reconnaître les trois codons stop UAA, UAG et UGA via son motif NIKS (Asn<sub>61</sub>, Ile<sub>62</sub>, Lys<sub>63</sub>, Ser<sub>64</sub>) formant une boucle entre les deux hélices à l'extrémité du domaine [126]. Le facteur eRF1 possède également un domaine central (M), qui présente des analogies structurales avec le bras accepteur d'un ARNt et possède un motif GGQ permettant la libération du peptide [127]. eRF3 est un facteur de terminaison de classe II, une GTPase permettant de stimuler l'activité d'eRF1. eRF1 inhibe la dissociation du GTP d'eRF3 ce qui permet de stabiliser le complexe eRF1-eRF3-GTP [128].

Une étude récente en molécule unique a permis d'obtenir de nouvelles données sur le déroulement de la terminaison de la traduction [124] (Figure 14). Il a été démontré que, bien que les facteurs de terminaison eRF3 et eRF1 sont capables de se lier individuellement au ribosome, eRF3 agirait comme une chaperonne du facteur eRF1, et c'est sous forme de complexe qu'ils se lieraient de manière optimale au ribosome. Ainsi, la stimulation de l'activité d'eRF1 par eRF3 passerait par son rôle de chaperonne, avec eRF3 qui bloquerait eRF1 dans la conformation la plus appropriée pour se lier au ribosome. Lorsque le complexe de terminaison entre dans le ribosome, le domaine M d'eRF1 est bloqué contre eRF3-GTP, cette conformation inhibant la lyse instantanée du peptide en cours de synthèse [129–131]. En effet, des données structurales indiquent que le motif catalytique GGQ est séquestré à plus de 80Å de la liaison ester de l'ARNt peptidyl. Le domaine N d'eRF1 se situe quant à lui dans le site A du ribosome et s'apparie au codon stop. eRF3 lui se lie au centre de liaison aux GTPases (GAC), entre la boucle sarcine-ricine (SRL) de la sous-unité 60S et les hélices h5 et h14 de l'ARN 18S de la sous-unité 40S. Le modèle de reconnaissance du codon stop par eRF1 a été décrit en deux étapes [132]. La première étape consiste en la liaison d'eRF1 au codon stop situé dans le centre de décodage du ribosome. eRF1 reconnaît alors les deux premiers nucléotides du codon stop (UA ou UG), les troisièmes (A et G) étant décodés sur des sites différents d'eRF1. Durant la deuxième étape, eRF1 effectue une rotation différente en fonction de la nature du second nucléotide ce qui permet la reconnaissance du second et troisième nucléotide (AG, AA ou GA). La reconnaissance du codon stop par eRF1 va alors provoquer la lyse de la liaison peptidique de l'ARNt peptidyl.

### **b. La libération de la protéine néosynthétisée**

La reconnaissance du codon stop par eRF1 a été décrite comme favorisant l'hydrolyse du GTP par eRF3. Bien que non essentielle, cette hydrolyse permet de stimuler l'accommodation d'eRF1 dans le site A du ribosome [125].

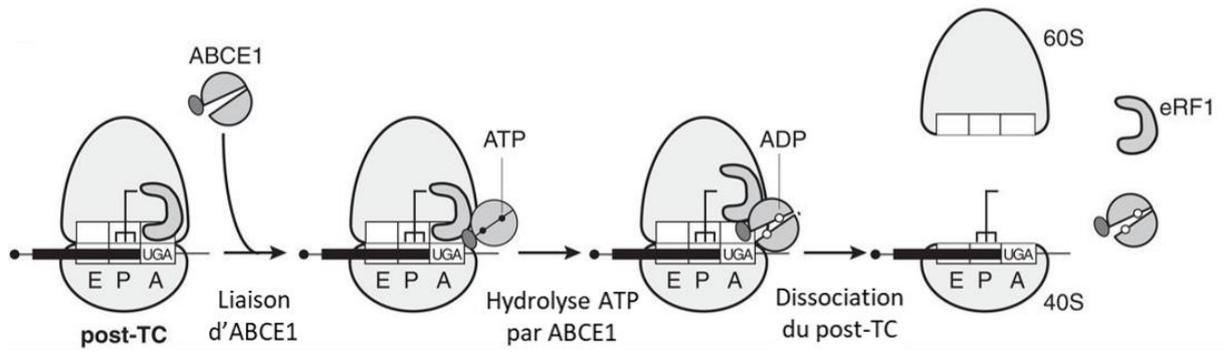


Figure 15 : Le recyclage des ribosomes chez les eucaryotes. Post-TC : complexe post-terminaison. Adapté de [123].

Elle provoque également la libération d'eRF3-GDP du ribosome qui est remplacé par le facteur ATP Binding Cassette subfamily E (ABCE1). eRF1 adopte alors une conformation relâchée qui permet au motif GGQ catalytique de se loger dans le centre peptidyl-transférase. L'accommodation du facteur eRF1 en conformation active a été décrite comme étant l'étape la plus longue du processus de terminaison, avec une durée de 3 secondes sur les 4 nécessaires à l'ensemble du processus [124]. Le motif GGQ se situe alors à proximité de l'extrémité CCA de l'ARNt-peptidyl, avec notamment Q<sub>135</sub> adjacent à la liaison ester reliant l'ARNt au néopeptide. Selon les conformations similaires entre eRF1 et les facteurs de libérations bactériens RF1/RF2, il a été déduit que son mode de lyse était analogue à ces derniers préalablement caractérisés. Chez les procaryotes, le placement du motif GGQ dans le centre peptidyl-transférase induit des changements de conformation qui exposent la liaison ester de l'ARNt peptidyl à une attaque nucléophile par une molécule d'eau. Le groupement amide de la chaîne principale de G<sub>185</sub> forme une liaison hydrogène avec le 3'-OH du dernier A du motif CCA en 3' de l'ARNt[133,134], permettant la libération du peptide synthétisé, ce qui entraîne une rotation entre les deux sous unités ribosomales et le mouvement de l'ARNt désacylé présent au site P en un état hybride P/E [124]. ABCE1 et eRF1 forment alors le complexe de post-terminaison et le recyclage du ribosome débute.

#### **D. Le recyclage**

Le recyclage du ribosome est la quatrième et dernière étape de la traduction. Il s'agit d'un point clé de régulation de la traduction par la cellule, qui est orchestrée par le complexe de post-terminaison eRF1-ABCE1 chez les eucaryotes. Ce mécanisme se décompose en deux étapes : la dissociation des sous unités ribosomiques, puis le recyclage des différents ARN.

##### **a. La dissociation des sous unités ribosomiques**

ABCE1 est une protéine à cœur Fe-S, dont les changements structuraux s'effectuent en fonction de la liaison à l'ATP. En effet, une liaison à l'ATP lui confère un état « fermé », tandis que l'hydrolyse de l'ATP en ADP conduit à un état « ouvert ». Ces changements conformationnels sont capables d'induire des changements structuraux dans les macromolécules associées, notamment eRF1 et le ribosome, ce qui pourrait conduire à la dissociation du ribosome [135] (Figure 15). Alors que de récentes études *in vitro* commencent à préciser ce mécanisme et semblent impliquer principalement le domaine Fe-S d'ABCE1 [136–139], un schéma complet de ces changements conformationnels et structuraux qui ont lieu avant et après l'hydrolyse de l'ATP lié à ABCE1 est toujours manquant. Il existe par ailleurs plusieurs avis divergents sur le modèle à établir pour la dissociation des sous unités ribosomiques.

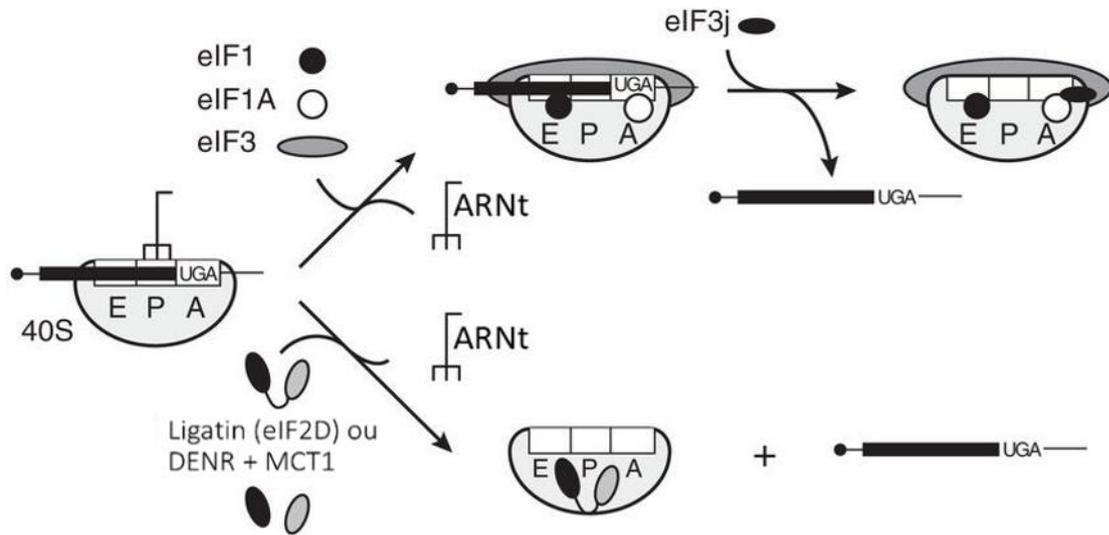


Figure 16 : La libération de l'ARNt désacylé et de l'ARNm de la sous unités 40S recyclée. Trois voies principales sont représentées avec l'intervention des facteurs eIF1, eIF1A, eIF3 et eIF3j (haut), d'eIF2D (bas) ou de DENR et MCT1 les homologues d'eIF2D (bas). Adapté de [123].

Certaines études tendent à montrer que l'hydrolyse de l'ATP par ABCE1 est nécessaire pour la dissociation ribosomique [136,140–142], tandis que d'autres suggèrent que la présence d'ATP est nécessaire pour la dissociation mais que son hydrolyse n'intervient que pour la libération d'ABCE1 de la sous unité 40S [143]. Dernièrement, une nouvelle étude révèle que le maintien d'ABCE1 sur la petite sous unité favoriserait le recrutement des eIFs et ainsi d'amorcer un nouveau cycle de traduction [139]. Bien qu'il reste encore à être élucidé, ce mécanisme provoque la dissociation des sous unités ribosomiques avec la libération de la sous unité 60S, d'eRF1 et potentiellement ABCE1-ADP. La sous unité 40S reste associée à l'ARNm, ainsi que l'ARNt désacylé présent au niveau du site P. Le recyclage entre alors dans la phase de libération des différents ARN.

### **b. Recyclage des différents ARN**

L'éjection de l'ARNt désacylé est indispensable pour le recyclage de l'ARNm et de la sous unité 40S, et peut se faire par différents mécanismes moléculaires (Figure 16). Le premier se fait grâce à l'intervention des facteurs d'initiation eIF1, eIF1a, eIF3 et eIF3j [50,140]. eIF1 et eIF1a se placent respectivement dans les sites P et A, mimant un début d'initiation de la traduction. eIF3 est présent du côté solvant de la sous unité 40S, interagissant avec eIF3j au niveau du site d'entrée de l'ARNm et avec eIF1 au niveau du site de sortie de l'ARNm [21]. eIF1 ayant pour rôle de discriminer les ARNt non initiateurs présents dans le site P, il déstabilise la liaison de l'ARNt désacylé ce qui provoque la libération de ce dernier [144]. Les facteurs eIF1a et plus particulièrement eIF3 permettent de stimuler cette libération [140]. Alors qu'eIF3 provoque la rétention partielle de l'ARNm, eIF3j stimule la libération de l'ARNm, terminant ainsi l'étape de recyclage. Deux autres voies de recyclage des ARN peuvent remplacer cette première. Elles concernent soit la ligatine eIF2D, soit ses homologues MCT-1 (Multiple Copies in T-cell lymphoma-1) et DENR (DENsity-Regulated protein) qui travaillent de concert [145]. La présence du domaine WHD (Winged-Helix Domain) d'eIF2D, ou de MCT1, à l'interface de la sous unité 40S serait incompatible avec la liaison du facteur eIF3. Au niveau du site P, le domaine SUI1/eIF1 de eIF2 ou de DENR aurait pour rôle de déstabiliser la liaison de la boucle anticodon de l'ARNt au site P et ainsi de provoquer l'éjection de l'ARNt désacétylé [146,147]. Alors que dans la première voie décrite eIF3 provoque la rétention partielle de l'ARNm après le départ de l'ARNt, son absence dans cette seconde voie provoque la libération immédiate de l'ARNm suite à celle de l'ARNt.

A la fin de la traduction, l'ensemble des facteurs impliqués (sous unités ribosomiques, ARNm et ARNt) sont ainsi recyclés et vont pouvoir entrer dans un nouveau cycle de traduction. La traduction est donc un processus en 4 étapes. Chacune de ces étapes peut induire des erreurs au cours de la synthèse protéique et des mécanismes ont été mis en place par la cellule pour limiter l'apparition de ces erreurs. Nous allons ainsi aborder dans ce nouveau chapitre la fidélité de la traduction eucaryote.

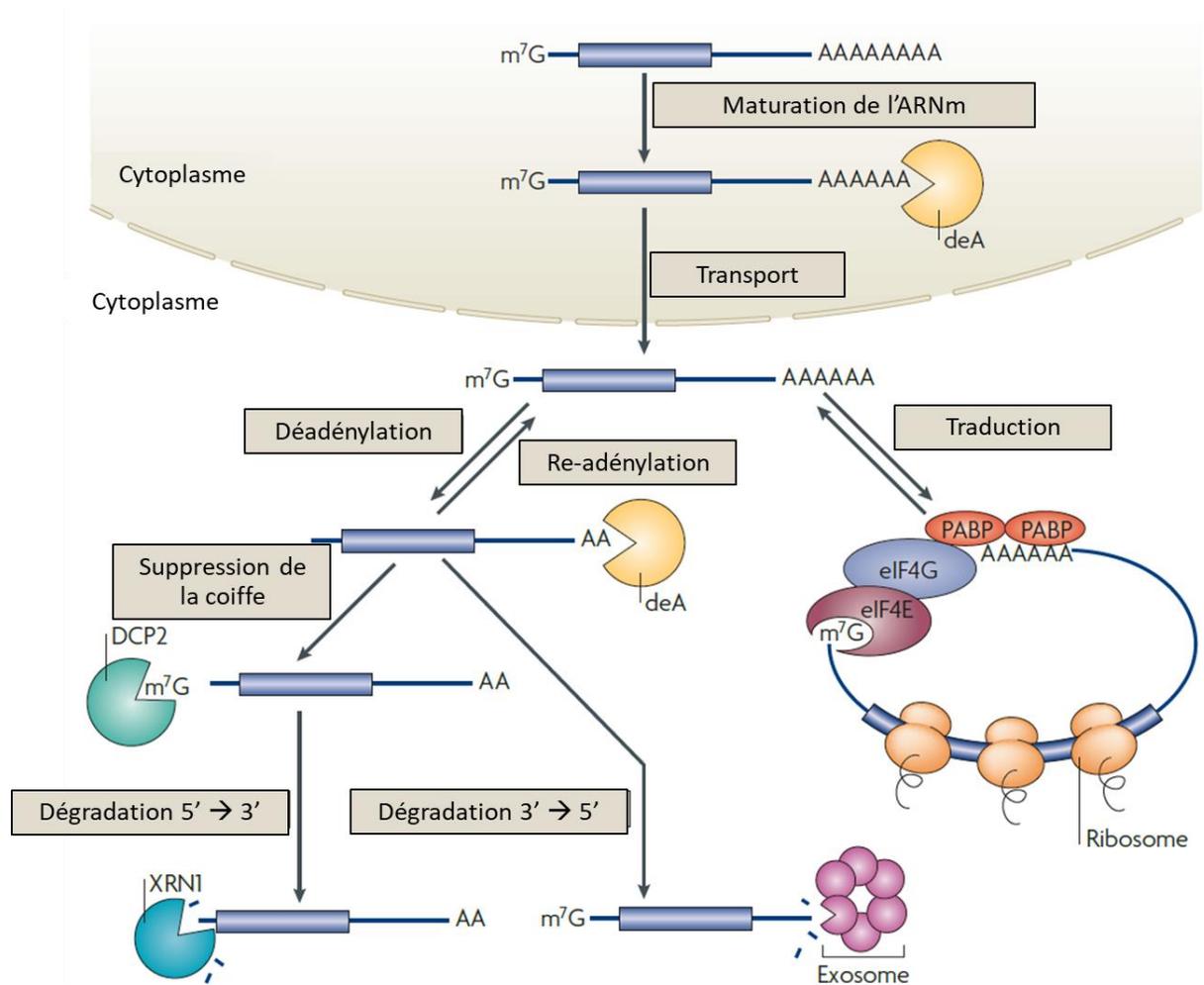


Figure 17 : Régulation de la stabilité des ARNm par la queue poly(A). Chez la levure, la deadénylase (deA) Pan2 taille les queues poly(A) nucléaires à leur longueur adéquate dans le noyau. Une fois dans le cytoplasme la suppression de la queue poly(A) de l'ARNm par les deadénylases (deA) provoquent la digestion de l'ARNm et/ou la répression de la traduction, tandis que la réadénylation peut activer la traduction de certains ARNm. Adapté de [148].

## II. La fidélité de la traduction chez les eucaryotes

Il est nécessaire de maintenir une fidélité de la traduction optimale afin de garantir une expression fiable des gènes. Pour cela différents acteurs entrent en jeu afin de garantir un contrôle qualité de la traduction, soit par l'intervention de voies de surveillance dédiées, soit par le ribosome lui-même. La traduction possède également un taux d'erreurs de  $10^{-3}$  à  $10^{-4}$  conservé chez l'ensemble du vivant [149], et différentes erreurs peuvent survenir que nous décrivons également dans ce chapitre.

### 1. Les mécanismes de contrôle-qualité de la traduction chez les eucaryotes

La dégradation de l'ARNm est importante pour maintenir le contrôle sur l'expression génique. Chez les mammifères, le temps de demi-vie des ARNm varie de 15 minutes à 24 heures ce qui reflète une régulation de ce processus par la cellule [150]. Alors que la dégradation des ARNm permet de réguler le niveau d'expression des gènes, nous allons voir qu'elle permet également d'éliminer les ARNm présentant des défauts de traduction reconnus par différentes voies de surveillance.

#### A. Dégradation cytoplasmique des ARNm

Les ARNm sont des molécules instables majoritairement sous forme simple-brin. Ils sont des cibles privilégiées pour de nombreuses ribonucléases qui clivent ses liaisons phosphodiesters. Comme je l'ai présenté précédemment la stabilité de l'ARNm repose principalement sur la coiffe m7G et la queue poly(A) ajoutées durant sa synthèse. Ils permettent la fixation de protéines cytoplasmiques spécifiques : les Cap-Binding Protein (CBP) et Poly(A)-binding protein (PABP). Ces facteurs ont pour rôle de protéger les extrémités de l'ARNm contre les exonucléases. Par ailleurs, l'ARNm est également associé à d'autres protéines et aux ribosomes au sein de larges particules ribonucléoprotéiques (mRNP) qui le protègent des endonucléases. En modulant la présence de ces facteurs, la cellule peut alors initier la dégradation de l'ARNm.

Chez les eucaryotes, la majorité des ARNm cytoplasmiques sont dégradés suite à une déadénylation, c'est-à-dire un raccourcissement de la queue poly(A) [151] (Figure 17). La dégradation de la queue poly(A) est réalisée soit par le complexe Pan2/Pan3, soit par le complexe Ccr4/Not [152]. La déadénylation a la particularité d'être réversible : si l'ARNm contient des signaux intacts et une séquence Cytoplasmic Poly(A)denylation Element (CPE), une queue poly(A) peut être remise en place par des protéines de la famille des Cytoplasmic Poly(A)denylation Elements Binding proteins (CPEB)Poly(A). Ce processus est connu pour moduler l'expression génique dans des contextes comme la gamétogenèse, le cancer, la plasticité synaptique et l'apprentissage/mémoire cellulaire [153].

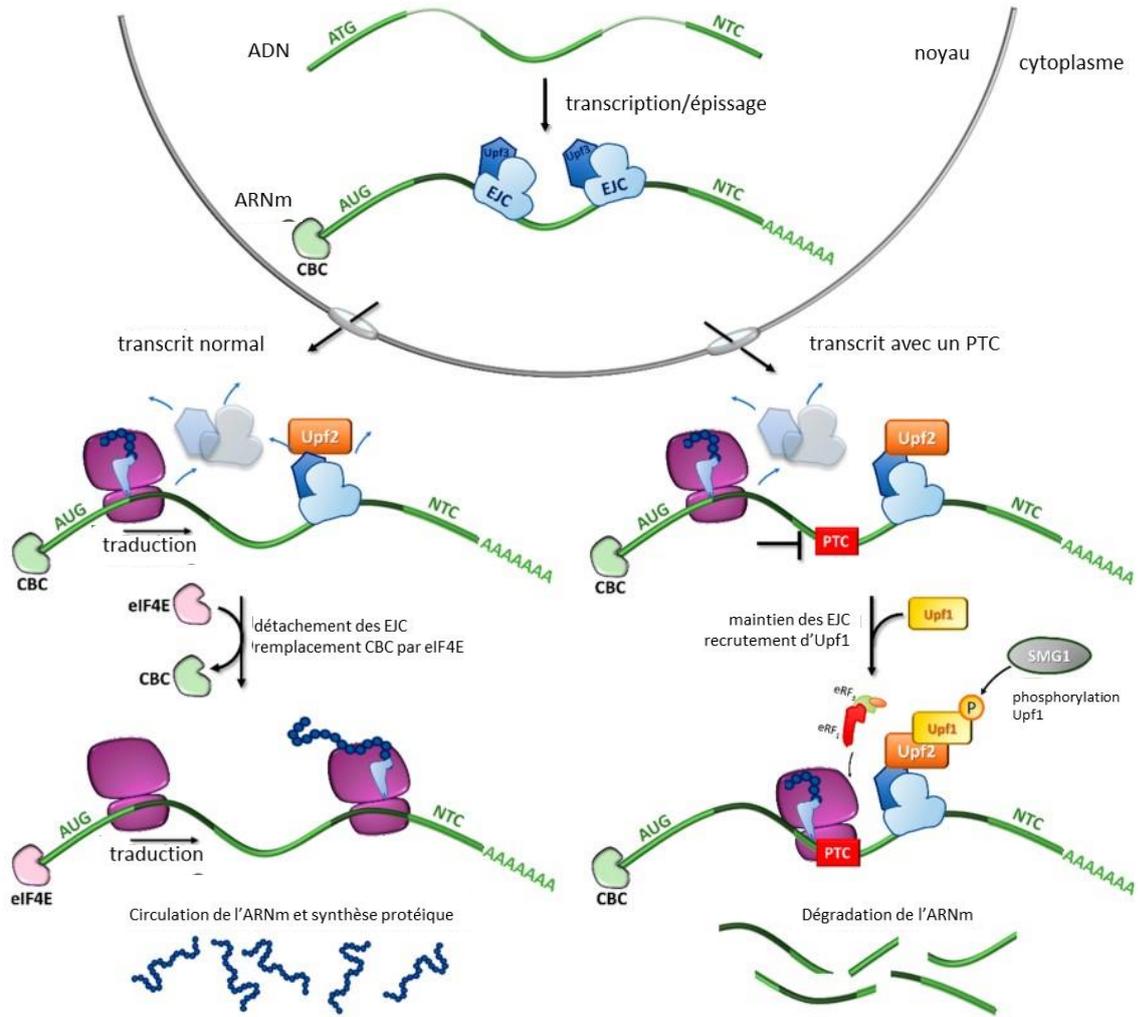


Figure 18 : Déclenchement de la voie Nonsense-mediated Decay (NMD) pour la dégradation d'un ARNm avec un codon stop prématuré (PTC). L'épissage du pré-ARNm dans le noyau donne lieu à un transcrit ARNm mature lié à des complexes exon-jonction (EJC) ainsi qu'à d'autres composants clés tels que les protéines Upf3 (dans le noyau) et Upf2 (dans le cytoplasme). Une fois transporté dans le cytoplasme, l'ARNm subit un premier cycle de traduction : la voie de gauche se produit en absence de PTC et la voie de droite (NMD) en présence d'un PTC. Adapté de [154].

Cependant, chez les mammifères, si moins de 30 à 60 adénines sont présentes après la déadénylation, le processus de dégradation s'engage rapidement de manière irréversible. En effet, après raccourcissement de la queue poly(A) la PABP ne peut plus se fixer sur l'ARNm pour le protéger de la dégradation. Si le complexe Dcp1/Dcp2 clive de manière irréversible la coiffe m7G, le transcrit est alors dégradé par l'exoribonucléase Xrn1 dans le sens 5' vers 3'. Sinon, l'ARNm peut être dégradé dans le sens 3' vers 5' par l'exosome, un complexe d'exonucléases.

Il est possible de réguler l'expression génique par dégradation des transcrits de manières dépendantes de protéines ou d'ARN. Les protéines concernées sont capables de se fixer sur des séquences spécifiques situées dans les régions non codantes 5' UTR et 3' UTR, ainsi que dans la région codante de l'ARNm. Dans les cas les plus documentés on retrouve par exemple les protéines Puf (Pumilio RNA-binding protein family) [155] et les protéines Adenylate-Uridylate-rich elements Binding Proteins (AUBP) [156]. Ces protéines présentent différents modes d'action, elles peuvent induire la dégradation des ARNm notamment en accélérant le processus de déadénylation, en favorisant le blocage des ribosomes ou en facilitant le recrutement des micro-ARN (miRNA). Ces miRNA entrent dans la dégradation dépendante d'ARN appelée dégradation par ARN interférence (RNAi). De nombreux RNAi ont été découverts ces dernières années, et les plus connus sont les petits ARN interférants (siRNA) et les miRNA. Ces RNAi, complémentaires à leur ARNm cible, peuvent réguler son expression soit par recrutement du complexe de dégradation RNA-Induced Silencing Complex (RISC), soit en réprimant la fixation du ribosome en formant un duplex ARN-ARN, soit en accélérant le processus dépoly(A)dénylation.

## **B. Les voies de surveillance de la traduction**

La voie Nonsense-Mediated mRNA Decay (NMD) est connue pour être la voie de dégradation des ARNm contenant un codon stop prématuré (PTC) dans leur séquence (Figure 18). Il a été également démontré qu'elle régule la dégradation d'ARNm possédant une uORF, un intron en 3'UTR, ou une erreur issue de l'épissage alternatif [157,158]. Ces caractéristiques présentes sur les ARNm miment une terminaison prématurée et provoquent la dégradation par la voie NMD d'environ 10% des ARNm eucaryotes. La NMD est un processus nécessitant différents facteurs Up-Frameshift (UPF) et Suppressors with Morphogenetic effects on Genitalia (SMG). UPF1, liée à eRF3 au niveau du codon de terminaison, est activée par UPF2 et/ou UPF3 [159–161]. L'activation de UPF1 implique également sa phosphorylation par SMG1, dont l'activité kinase est contrôlée par SMG8 et SMG9 [162–165]. UPF1 phosphorylée recrute alors l'endonucléase SMG6 qui clive l'ARNm à proximité du PTC [166,167], et l'hétérodimère SMG5/SMG7 qui stimule la dégradation de l'ARNm en recrutant le complexe exonucléase CCR4/NOT [168]. Deux voies d'activation de la NMD ont été décrites.

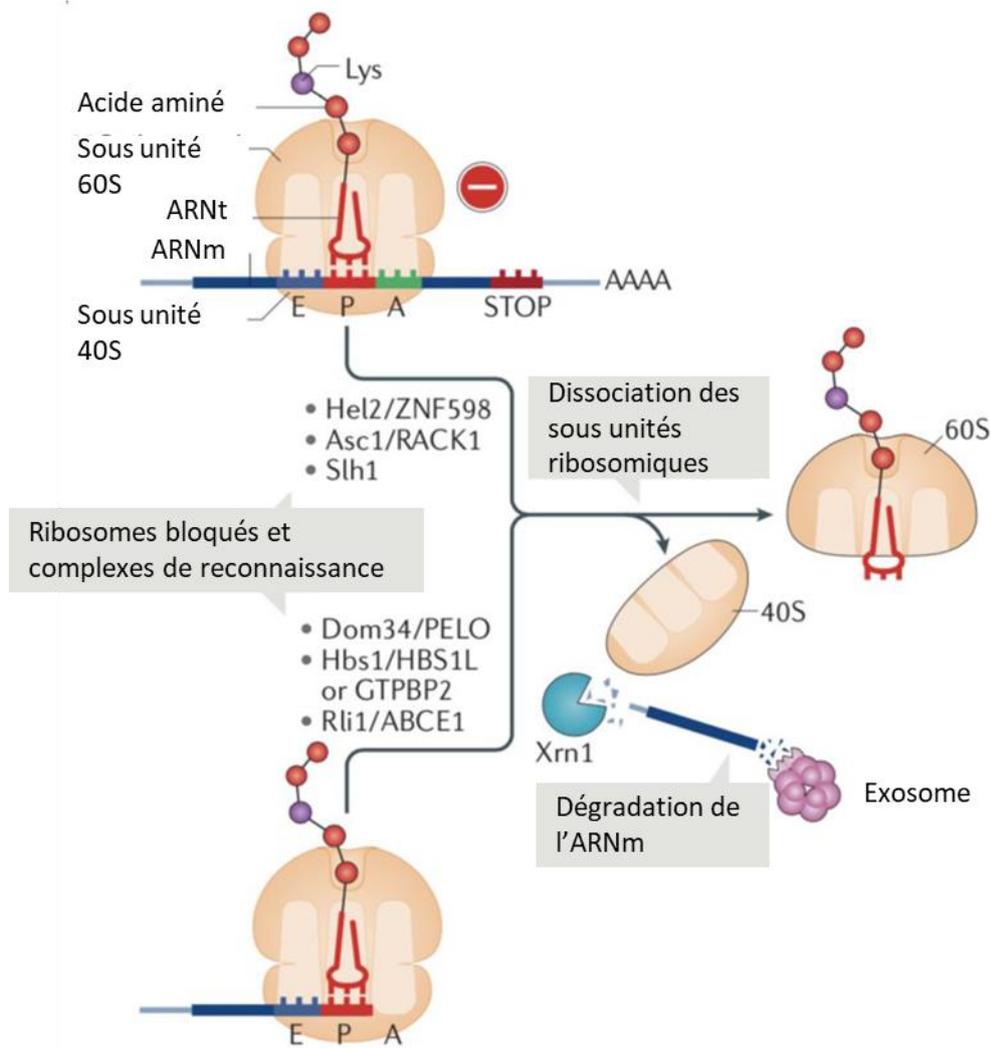


Figure 19 : Le système de détection et de déclenchement des voies de surveillance No-Stop (NSD) ou No-Go (NGD). Haut : le ribosome bloqué en interne de l'ARNm, cas de la NGD. La voie de déclenchement est identique à celle du cas n°2 de la NSD pour laquelle le ribosome serait bloqué dans la queue polyA. Bas : le ribosome est bloqué à l'extrémité de l'ARNm, cas n°1 de la NSD. Adapté de [169].

La première est l'activation par la présence de complexes de jonction d'exon (Exon junction complex, EJC). L'EJC, composé des protéines eIF4A3, Cancer Susceptibility Candidate 3 (CASC3), RNA-Binding Motif protein 8A (RBM8A) et MAGO-nashi Homolog (MAGOH), est déposé à 20-24 nucléotides en amont des jonctions exons-exons durant l'épissage [170,171]. Les EJC sont détachés de l'ARNm au cours de la traduction. La présence d'un PTC provoque la conservation des EJC en aval. Ainsi, un codon stop présent à plus de 50 nucléotides d'un EJC est considéré comme prématuré ce qui déclenche la NMD par interaction de l'EJC avec UPF3 ou UPF2 [172]. La deuxième voie d'activation de la NMD décrite est indépendante de la présence d'EJC. Elle se base sur l'interaction entre eRF3 et une protéine de liaison à la queue poly(A) de l'ARNm : Poly(A) Binding Protein Cytoplasmic 1 (PABPC1). Lorsque le codon stop est suffisamment proche de la queue poly(A), PABPC1 se lie à eRF3, empêchant la liaison d'UPF1 et favorisant alors la terminaison classique à l'activation de la voie NMD [161,173–175]. Dans le cas d'un PTC, la distance accrue avec la queue poly(A) serait le facteur inhibant l'interaction eRF3-PABPC1, favorisant ainsi la liaison eRF3-UPF1 et donc le déclenchement de la voie NMD. La présence de PTC peut avoir des conséquences pathologiques dans le cadre de maladies à mutation non-sens. L'inhibition de la NMD peut alors présenter de grands intérêts dans la démarche thérapeutique de réexpression protéique que nous aborderons dans le paragraphe sur la translecture.

Deux voies de reconnaissances des ribosomes bloqués sur l'ARNm co-existent chez les eucaryotes : la Non-Stop Decay (NSD) et la No-Go Decay (NGD). La NSD est la voie de surveillance déclenchée par l'absence de codon stop dans l'ARNm [176,177]. Deux cas différents de NSD sont rapportés dans la littérature (Figure 19) : (1) le ribosome ne rencontre pas de codon stop alors qu'il est arrivé à l'extrémité 3' du transcrit et son site A est vide (2) aucun codon de terminaison n'est reconnu au cours de la traduction et le ribosome se retrouve dans la queue poly(A), provoquant l'apparition d'une chaîne poly-lysines [169]. Il a été démontré que 10 lysines à la suite seraient suffisantes pour bloquer la traduction, d'une part par l'abondance faible d'ARNt<sub>Lys</sub> dans la cellule, mais également par les liaisons électrostatiques que cela engendre dans le tunnel de sortie chargé négativement du ribosome [178,179]. La voie de surveillance NGD concerne les ARNm avec des ribosomes bloqués en cours d'élongation dans un état hybride (Figure 19). Le blocage du ribosome peut avoir plusieurs origines : une accumulation d'acides aminés chargés positivement dans le tunnel de sortie du ribosome, la présence de structures secondaires telles que des pseudonœuds, la présence de codons rares, la présence de lésions sur l'ARNm ou encore la présence de codons difficilement décodés (proline). Le cas (2) de la NSD et la NGD semblent faire intervenir les mêmes facteurs, la frontière entre les deux s'amincissant au cours des années.

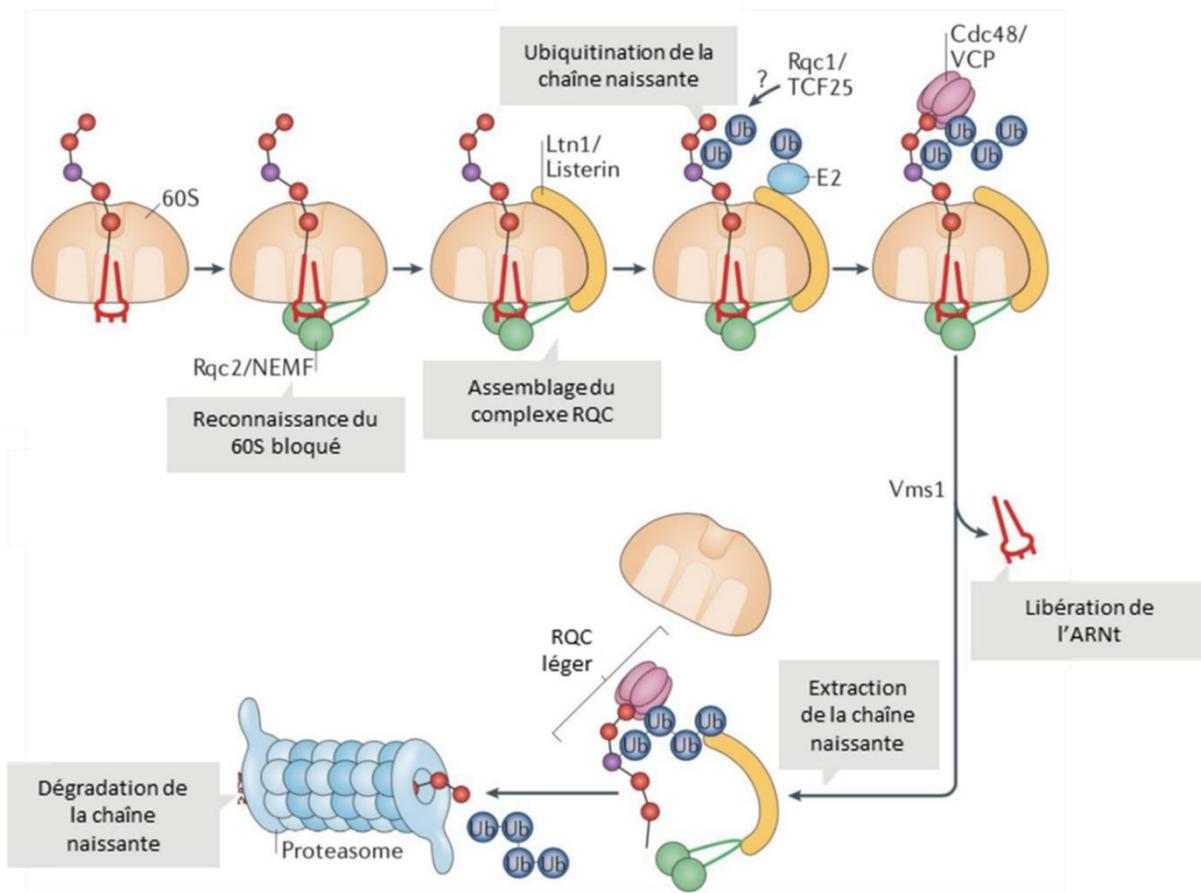


Figure 20 : Le système de surveillance Ribosome-associated protein Quality Control (RQC) dans la levure. Rqc2 et NEMF sont des homologues levure et mammifères respectivement. Le mode de liaison au ribosome de la sous-unité Rqc1 (TCF25 chez les mammifères) du complexe RQC et sa fonction exacte restent inconnus. Vms1 et ANKZF1 sont des homologues levure et mammifère respectivement. Les peptide extrait peut rester associé au complexe RQC sous la forme d'un complexe « RQC léger » avant la dégradation par le protéasome. Adapté de [169].

Pour commencer, le même senseur est retrouvé pour les deux voies chez les mammifères. En effet, le blocage d'un ribosome à l'intérieur de l'ARNm provoque une collision avec celui qui le suit et ainsi l'apparition de disomes, qui sont reconnus par la protéine E3 ligase ZNF598 [180] (Figure 19). La protéine ZNF598 ubiquitine les protéines ribosomales eS10, uS10 et eS3. Il y a alors intervention du complexe Pelota-Hbs1 avec un clivage endonucléolytique de l'ARNm en 5' du ribosome [181,182]. Les ribosomes bloqués en extrémité 3' du transcrit (cas (1) de la NSD) seraient reconnus directement par le complexe Pelota/Hbs1 (Figure 19) [169]. Après le clivage en 5' du ribosome, il y a recrutement du complexe SKI (composé des facteurs Ski2, 3 et 8) au niveau du ribosome sur la partie 3' de l'ARNm [183]. Ces derniers permettent de mobiliser l'exosome et ainsi de dégrader l'ARNm.

Lorsqu'un ribosome est ciblé par un des mécanismes décrits ci-dessus, la sous unité 40S est libérée et prête à effectuer un nouveau cycle de traduction. Ce n'est pas le cas de la sous unité 60S qui contient le peptide tronqué bloqué dans son tunnel de sortie. C'est à ce moment-là que le système de surveillance Ribosome-associated protein Quality Control (RQC) intervient (Figure 20). Le système RQC débute par la liaison spécifique de Rqc2 à la grande sous unité qui expose un ARNt peptidyl, suite au départ de la petite sous unité [184,185]. La liaison de Rqc2 à la sous unité 60S et à l'ARNt prévient la réassociation avec la sous unité 40S et facilite la liaison de l'ubiquitine ligase Ltn1 (E3 ubiquitin-protein ligase listerin). Ltn1 possède une structure hyperhélicoïdale lui permettant de contourner la sous unité 60S, jusqu'au placement de sa partie C-terminale au niveau de la sortie du tunnel de sortie. C'est ainsi qu'elle entre en contact avec les protéines ribosomales eL22 et eL31 qui induisent l'ubiquitination du néopeptide par Ltn1 et E2, l'enzyme associée à l'ubiquitine. Cette ubiquitination est suffisante pour permettre le recrutement de l'endonucléase AZKF1 qui clive la liaison phosphodiester présente au niveau de la tige acceptrice de l'ARNt peptidyl, permettant la libération de l'ARNt du site P. Ensuite, le complexe p97 permet l'extraction du peptide de la sous unité 60S ainsi que de son adressage au protéasome [186].

## **2. Le rôle du ribosome dans le maintien de la fidélité**

Depuis une vingtaine d'années il a été proposé que le ribosome est un « filtre » de la traduction [187] et qu'il existe en réalité des sous populations spécialisées de ribosomes entre chaque tissu de l'organisme, voir au sein même de la cellule, permettant de réguler la traduction. Ainsi, nous allons à présent nous concentrer sur le ribosome et nous intéresser au rôle de sa composition et de sa structure dans le maintien de la fidélité chez les eucaryotes.



## A. Composition variable du ribosome eucaryote

Ces dernières années le ribosome est apparu comme un acteur à part entière dans la régulation de la traduction. Le modèle de ribosomes spécialisés repose sur la présence de populations de ribosomes présentant une composition variable en ARNr, en protéines ribosomales ou en facteurs associés au ribosome, et présentant des capacités de traduction différentes.

Nous avons vu plus haut que le ribosome tient son activité catalytique de l'ARNr. Différentes études ont été faites sur l'impact de la composition en ARNr sur l'activité du ribosome. Il a par exemple été mis en évidence que chez le Zebrafish les embryons expriment des sous-types distincts (maternel et somatique) d'ARNr 5.8S, 18S et 28S pendant l'embryogenèse [188]. Des approches *in silico* ont indiqué que les régions 5'UTR des ARNm maternels pourraient se lier de manière préférentielle à certaines régions présentes dans le sous-type d'ARNr 18S maternel, mais pas dans le sous-type somatique. La composition du ribosome vis-à-vis de l'ARNr peut également être modulée en fonction des modifications chimiques présentes sur l'ARNr, notamment la 2'-O-méthylation et la pseudouridine. Présentes sur une 100<sup>ne</sup> de nucléotides de l'ARNr, nous rentrerons en détails sur les effets de ces modifications dans le chapitre associé au rôle des modifications d'ARN dans la fidélité de la traduction. Il est tout de même intéressant de noter ici qu'une variation dans le nombre de sites 2'-O-méthylés ou pseudouridinés ont été associés à une modulation de la fidélité de la traduction ainsi que de l'initiation dépendante d'IRES [189–192]. Ces variations dans le pattern de modification peuvent être dû à un changement d'expression de la protéine responsable de la modification, ou bien des ARN guides site- spécifiques, et de nombreuses données dans ce domaine tendent à prouver l'existence de ribosomes spécialisés tissus-spécifiques.

La composition du ribosome peut également varier en fonction des protéines ribosomiques. Certaines d'entre elles ne sont pas indispensables au fonctionnement du ribosome, mais leur absence module sa fidélité. C'est notamment le cas de Rpl38 dont la présence est nécessaire à la traduction de certains ARNm *Hox* [193]. Ces transcrits possèdent un élément inhibiteur de la traduction dépendante de la coiffe m7G dans leur 3'UTR, et leur expression dépend d'une IRES nécessitant Rpl38 [194]. L'absence de Rpl38 entraîne le développement d'une vertèbre thoracique à la place d'une lombaire, provoquant un phénotype de queue courte chez la souris. Comme l'ARNr, les protéines ribosomiques sont sujettes à des modifications chimiques, comme l'ajout d'ubiquitine, qui peuvent participer à une hétérogénéité dans les populations de ribosomes. Chez la levure, les niveaux d'ubiquitination de Rpl28 varient pendant le cycle cellulaire et les ribosomes avec Rpl28 polyubiquitiné ont un taux de traduction plus élevé *in vitro* par rapport aux ribosomes avec Rpl28 monoubiquitiné [195].

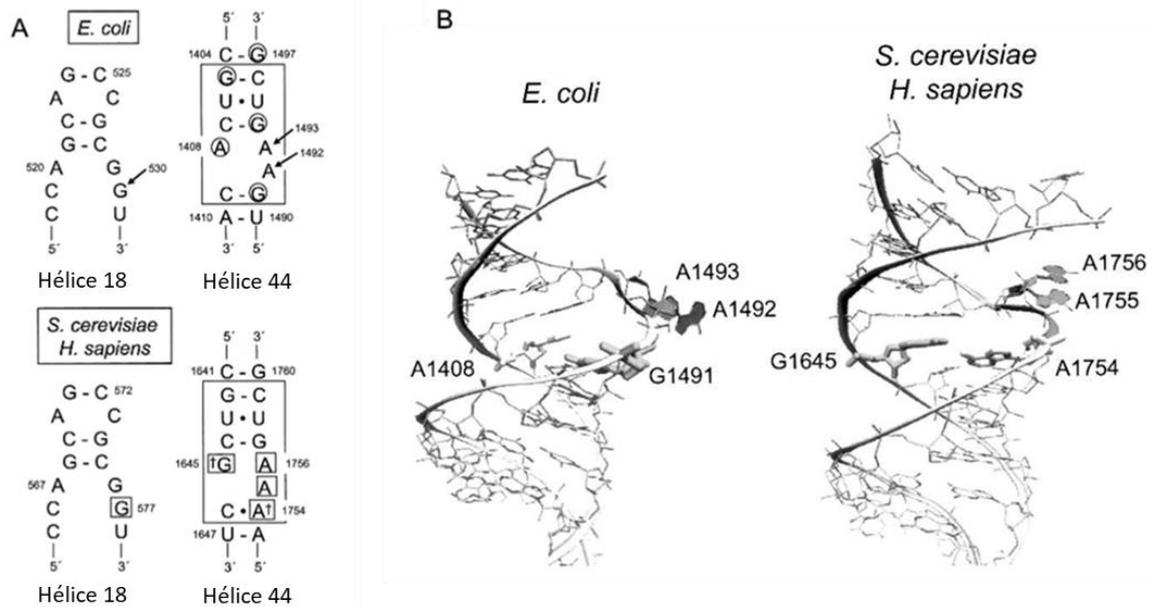


Figure 21 : Les résidus du site de décodage de l'ARNr 16S d'*E. coli* et des ARNr 18S de *S. cerevisiae* et *H. sapiens*. (A) Structures secondaires des ARNr qui contribuent au site de décodage. Portions des hélices 18 et 44 de l'ARNr 16S d'*E. coli* (en haut). La portion de l'hélice 18 contenant le résidu G530 est représentée à gauche. Les résidus de la tige de décodage centrale dans l'hélice 44 sont encadrés à droite. Les bases protégées de la modification DMS par la paromomycine sont encerclées. Les flèches indiquent les résidus du site de décodage G530, A1492, et A1493. Bas : les portions des hélices 18 et 44 dans les ARNr 18S de *S. cerevisiae* et *H. sapiens*. Les résidus correspondant à la tige de décodage centrale de l'hélice 44 d'*E. coli* sont entourés de la grande boîte. Les croix allongées marquent les deux bases du site de décodage non conservées entre procaryotes et eucaryotes. Les petites cases indiquent les résidus mutés dans l'étude de l'article. (B) Comparaison des sites de décodage procaryotes et eucaryotes. Les structures des fragments d'ARN contenant les sites de décodage de *E. coli* (à gauche) ou de *S. cerevisiae* et *H. sapiens* (à droite) sont présentées. Les deux résidus non conservés (A1408 et G1491 dans *E. coli* ; G1645 et A1754 dans *S. cerevisiae*) sont mis en évidence. Les adénines conservées (A1492 et A1493 chez *E. coli* ; A1755 et A1756 chez *S. cerevisiae*) sont repliées et représentées dans la conformation retournée. Adapté de [196].

Dans les cellules de mammifères, il a été également démontré que le taux d'ubiquitination de RPS10, RPS20 et RPS3A est important dans les mécanismes de contrôle de la qualité de la cellule par rapport aux ribosomes présents sur la queue poly(A) du transcrit [197].

Il existe d'autres études suggérant l'existence de ribosomes spécialisés, certaines explorant également d'autres pistes telles que la nature des paralogues présents dans le ribosome, la stœchiométrie de certaines protéines, ainsi que les facteurs associés au ribosome, leurs résultats restent encore très contestés à cause d'effets cellulaires non-spécifiques au ribosome qui ne peuvent être écartés.

### **B. Données structurales sur le site de décodage**

Le centre de décodage du ribosome est présent dans la petite sous unité au niveau de l'ARN 18S eucaryote ou 16S procaryote. Il est situé dans le site A du ribosome où il permet la sélection de l'ARNt aminoacylé. Le centre de décodage eucaryote partage une homologie significative avec celui des procaryotes, ce qui suggère que les fonctions essentielles du processus de décodage seraient fortement conservées (Figure 21).

Des études ont mis en évidence trois nucléotides universellement conservés dans le site de décodage (*E. coli*: G530 de l'hélice 18, et A1492 et A1493 de l'hélice 44 / mammifères : G626, A1824 et A1825 respectivement), qui interagissent avec l'hélice codon-anticodon pour contrôler la précision de l'appariement des bases et stabilisent la liaison ARNt du site A par liaison hydrogène [3,198–201]. Il a été démontré que la liaison d'un ARNt cognat ou proche cognat au ribosome 70S induit des réarrangements identiques de (i) la petite sous unité (déplacement de 2 à 3 Å de l'épaule ribosomique) et (ii) du centre de décodage [202]. Des données structurales ont révélé que l'encombrement stérique provoqué par l'appariement codon-anticodon prévalait sur le nombre de liaisons hydrogène de cet appariement pour l'acceptation de l'ARNt, en raison de la structure 3D de l'ARNt et de l'enveloppe rigide formée par le site de décodage [203].

Une analyse structurale de l'hélice 44 humaine a révélé que les résidus correspondant à A1824 et A1825 ont le potentiel d'effectuer une rotation comme leurs homologues bactériens [204]. Par ailleurs, il a été proposé que l'interaction de G626 (G530 chez les bactéries) puisse remplir une fonction plus cruciale en tant que nucléotide de « verrouillage », qui fixe l'hélice codon-anticodon dans le centre de décodage du ribosome[201].



Trois modèles principaux existent alors sur le rôle de l'ARN 16S dans la sélection de l'ARNt. Le premier propose que les deux premiers appariements de l'hélice codon-anticodon sont étroitement vérifiés par les bases A1492 et A1493 du ribosome pour discriminer entre des appariements de type Watson-Crick et des mésappariements, alors que l'environnement de la troisième position (position « wobble ») semble adapté pour accepter d'autres géométries d'appariements [199]. Le deuxième modèle a été proposé suite à l'analyse des données structurales complètes du ribosome, et non uniquement sur celles de la petite sous unité 30S [205]. Les appariements U-G en première ou deuxième position du codon seraient possibles en formant un duplex de type Watson-Crick avec une interaction avec les bases A1492, A1493 et G530 identique aux appariements canoniques Watson-Crick. De plus, le mouvement des bases A1492 et A1493 formerait une partie statique du centre de décodage, définissant ses propriétés spatiales et chimiques et créant des contraintes : les distorsions induites par les mésappariements se propageraient dans toute la structure de l'ARNt, ce qui le déstabiliserait et faciliterait son rejet hors du ribosome. Dans le troisième modèle, l'hélice 69 de la grande sous unité joue un rôle crucial [203]. Cette hélice forme un pont entre les deux sous unités. La sélection de l'ARNt débute par la liaison de l'ARNt au site A libre où le résidu A1493 est en position pour interagir avec le duplex codon/anticodon. Le résidu A1492 est positionné sur le résidu A1913 de l'hélice 69 de l'ARNr 23S. L'entrée d'un ARNt cognat ou non cognat provoque alors deux réarrangements majeurs. Tout d'abord, lors de la fixation de l'ARNt, la boucle anticodon entre en contact avec l'hélice 69 et casse l'interaction A1492- A1913. Ceci provoquerait le repositionnement de la base A1492 qui interagirait alors avec le codon en deuxième position. Le second réarrangement correspond au rapprochement du résidu G530 au niveau de la seconde position du codon et finalise la formation du centre de décodage. L'ARNt est alors bloqué, de telle sorte que la complémentarité du duplex codon-anticodon soit testée en fonction de la force d'appariement entre leurs bases. Un ARNt serait alors efficacement incorporé en raison de sa capacité à former un appariement Watson Crick stable sur les deux premières positions du codon. A contrario, la plupart des ARNt proche-cognat seront éjectés en raison de l'encombrement stérique au niveau de l'appariement codon-anticodon ou du coût énergétique nécessaire pour que l'anticodon de l'ARNt soit correctement positionné dans le centre de décodage. Toutefois, certains ARNt proche-cognat échapperaient à cette discrimination du fait de leur capacité à former une interaction Watson-Crick via une forme tautomère rare de leurs bases.

Aux abords du centre de décodage, d'autres hélices de l'ARNr de la petite sous unité semblent être importantes dans le maintien de la fidélité : les hélices 8, 14, 27 et 34 [206–209].



Des protéines ribosomiques ont aussi été impliquées dans le maintien de la fidélité. La protéine eS12 (sous-unité 30S) se situe à proximité du centre de décodage et interagit avec l'ARNr 16S. Une approche par mutagenèse dirigée réalisée chez *E. coli* a montré que deux hélices très conservées de cette protéine ont un rôle dans la sélection de l'ARNt au site A. Les mutants obtenus présentent une augmentation de la fidélité de la traduction [210]. Des mutants dans les protéines RPS2 et RPS9, appelés RAM (Ribosome Ambiguity Mutants), sont responsables d'une diminution de la fidélité incluant le passage de codon stop, l'incorporation de mauvais acides aminés et des décalages de phase de lecture [211].

### **3. Le recodage**

Malgré l'ensemble des mécanismes mis en place pour garantir la fidélité de la traduction, cette étape reste la moins fidèle de l'expression génétique avec un taux d'erreurs d'environ  $10^{-4}$  [212]. Il existe par exemple l'erreur de chargement des ARNt par les aminoacyl-ARNt-synthases. *In vitro*, les taux d'erreurs de ces enzymes sont de  $10^{-4}$  à  $10^{-5}$  dans la sélection des acides aminés, et de  $10^{-6}$  dans celle des ARNt [213]. Alors que l'identité de l'acide aminé chargé sur l'ARNt n'est pas vérifiée par le ribosome, seulement la moitié des aminoacyl-ARNt synthases seraient strictement fidèles. Ces erreurs de chargements des ARNt participent à l'existence du code génétique alternatif (Figure 7), et auraient des rôles cellulaires tels que la réponse au stress oxydatif [214]. Ce taux d'erreurs, conservé dans l'ensemble du vivant a été sélectionné par la présence de signaux dans certains ARNm, qui altèrent la lecture du code génétique en modifiant les règles de décodage et qui ont pour conséquence la synthèse d'une protéine différente. Ces erreurs programmées de la traduction sont regroupées sous le terme « recodage » et sont en compétition avec les règles classiques de la traduction. Ce processus découvert initialement chez les virus a par la suite été mis en évidence chez tous les organismes dans lesquels il a été recherché. Deux types d'événements de recodage peuvent se produire au cours de la traduction : le décalage du cadre de lecture et le passage du codon stop.

#### **A. Le décalage du cadre de lecture**

Le décalage du cadre de lecture, appelé « Frameshift », correspond au décalage du ribosome en cours d'élongation d'un nucléotide en 5' (Frameshift -1) ou en 3' (Frameshift +1). L'élongation continue alors dans une phase de lecture différente de celle de départ, définie par le codon AUG. Tous les événements de frameshift programmés nécessitent à la fois une séquence glissante qui favorise le glissement des ARNt et un élément stimulateur qui augmente l'efficacité du glissement en induisant une pause du ribosome. En absence d'éléments stimulateurs, la fréquence de cet événement est relativement faible, environ  $3 \cdot 10^{-5}$  par codon.

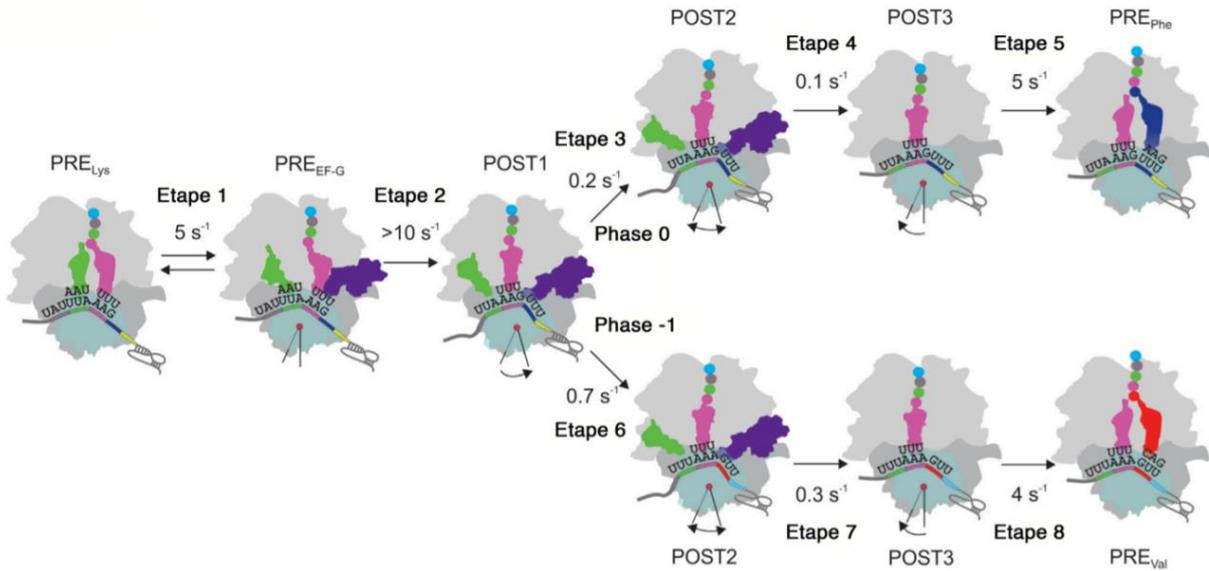


Figure 22 : Modèle cinétique de frameshift -1 en traduction procaryote. Le changement de phase de lecture se produit au cours de la translocation des deux ARNt liés à la séquence glissante (ARNt-Leu et du ARNt-Lys peptidyl). Etape 1 : recrutement de EF-G sur le complexe de pré-translocation (PRE). Etape 2 : mouvement rapide de l'ARNt dans un état chimérique de post-translocation (POST1). Puis, l'ARNt-Leu poursuit son mouvement résultant en un décalage du cadre de lecture -1 (bas) ou non (haut). Tout d'abord, l'ARNt-Leu se déplace sur la sous-unité 50S dans un état POST2 sans que sa distance à la sous-unité 30S ne change (étapes 3 et 6 respectivement en phase 0 et -1). Deuxièmement, l'ARNt-Leu et la sous-unité 30S s'écartent (étapes 4 et 7) dans un état de POST3. Les étapes 3 et 4 sont particulièrement lentes pour l'ARNt qui reste dans en phase 0, ce qui limite le taux de fixation de l'ARNt-Phe (étape 5). En revanche, le mouvement de l'ARNt-Leu sur les ribosomes qui se sont décalés en phase -1, est plus rapide (étape 6). Ce mouvement est suivi de la dissociation de l'ARNt-Leu de la sous-unité 30S, puis de la rotation de la du domaine tête de la sous-unité 30S, de la dissociation de EF-G (étape 7) et enfin de la liaison de l'ARNt-Val (étape 8). Adapté de [215].

Le frameshift -1, dont le mécanisme est le mieux compris, est un événement durant lequel le ribosome recule d'une base en 5' sur l'ARNm (Figure 22). Le ribosome a besoin d'au minimum deux signaux pour avoir une efficacité de décalage significative: une séquence glissante et une structure secondaire située immédiatement en 3'. La séquence glissante correspond à la région sur laquelle le ribosome va effectuer une pause et se décaler d'un nucléotide en 5'. Cette séquence est généralement heptamérique et de consensus X XXY YYZ, où X=N, Y=A ou U et Z≠G. L'élément stimulateur peut être une structure secondaire de type pseudonœud ou tige-boucle située en 3' de la séquence de glissement. La présence de cet élément permet d'augmenter de 10 à 30 fois l'efficacité de décalage de la séquence glissante. Chez les eucaryotes, la distance optimale entre ces 2 signaux varie entre 6 et 8 nucléotides et permettrait au ribosome d'exercer son activité hélicase au niveau de la structure secondaire. Les données issues de la structure cryo-EM d'un ribosome de mammifère au cours d'un événement de frameshift -1 ont montré que le pseudo-nœud bloquait le ribosome à une étape intermédiaire de la translocation, provoquant un repliement de l'ARNt au site A et contraignant le ribosome à se décaler d'un nucléotide en 5'. De plus, ces résultats ont mis en évidence l'interaction directe du facteur d'élongation eEF2 avec l'ARNt au site P. Bien qu'il s'agisse d'un mécanisme essentiellement viral, il existe plusieurs gènes mammifères connus pour être régulés par frameshift -1. On retrouve notamment PEG10 (Human Paternal Expression Gene 10) [216], PNMA3 and PNMA5 (ParaNeoplastic Mammalian Antigen-like genes) [217] et Mar3 (Mammalian Retrotransposon-derived 3) [218].

Le frameshift +1 est un événement durant lequel le ribosome avance d'une base en 3' sur l'ARNm. Le décalage de cadre de lecture se produit sur une séquence minimale tétramérique XXX X. Dans la plupart des cas, le codon adjacent stimule le décalage et la séquence forme alors un heptamère XXX XXX X. Contrairement au frameshift -1, il n'y a pas de séquence consensus [219], mais les codons favorisant le glissement ont été identifiés [220]. Deux mécanismes peuvent expliquer un événement de décalage du cadre de lecture en +1 : soit l'ARNt au site P se décale d'un nucléotide en 3', soit l'ARNt rentrant au site A s'apparie directement dans la phase +1 [221]. La différence réside dans l'affinité de l'ARNt au site P pour son codon et dans sa capacité à glisser. Quel que soit le mécanisme impliqué, le frameshift +1 nécessite une reconnaissance lente du codon présent au site A ainsi qu'un décodage efficace du codon situé en phase +1. Celle-ci peut être déclenchée par la présence d'un codon rare, par un codon stop peu efficace ou par une structure secondaire de type tige boucle ou pseudonœud. Plusieurs gènes cellulaires sont régulés par frameshift +1 chez les eucaryotes : c'est le cas de l'antizyme OAZ1 qui régule la concentration en polyamines, des molécules indispensables au bon fonctionnement de nombreux processus cellulaires et à la croissance [222].

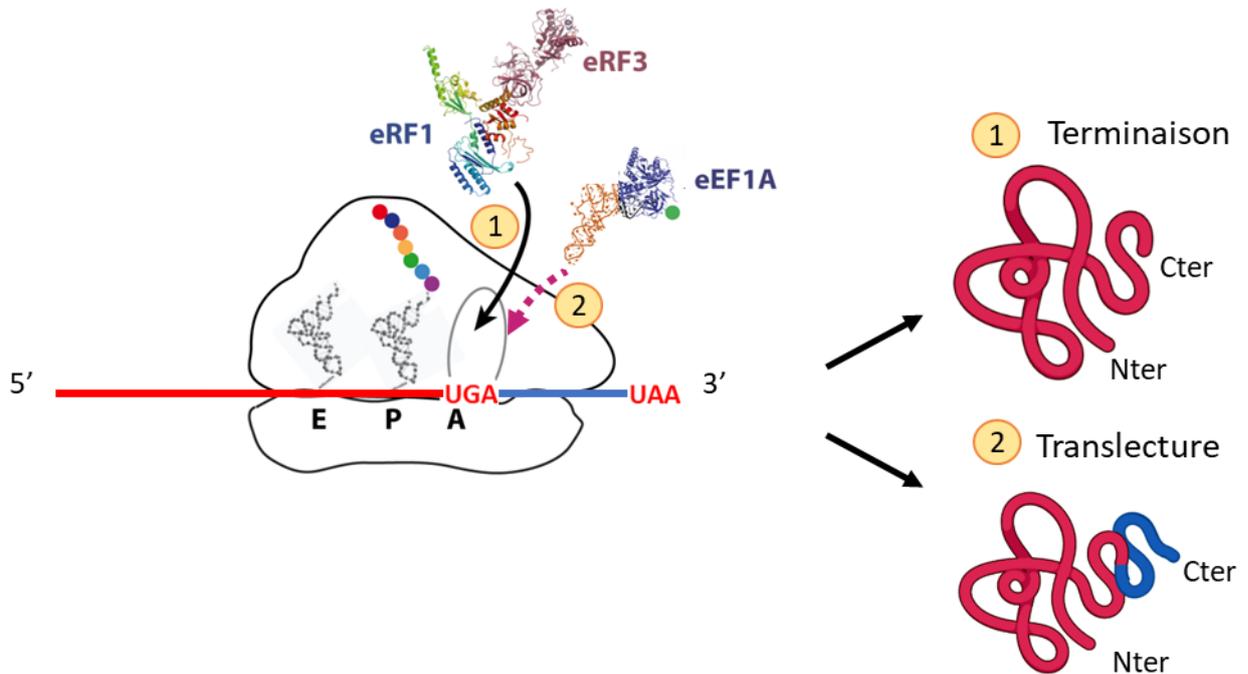


Figure 23 : La translecture d'un codon stop chez les eucaryotes. (1) Terminaison de la traduction : le complexe de terminaison eRF1-eRF3 est incorporé au niveau du codon stop ce qui déclenche la terminaison de la traduction. (2) Translecture du codon stop : un ARNt cognat (ARNt Sélénocystéine) ou proche cognat (ARNt supprimeur) est incorporé au niveau du codon stop. La traduction continue jusqu'au prochain codon stop provoquant la production d'une extension peptidique au niveau de la partie C-terminale.

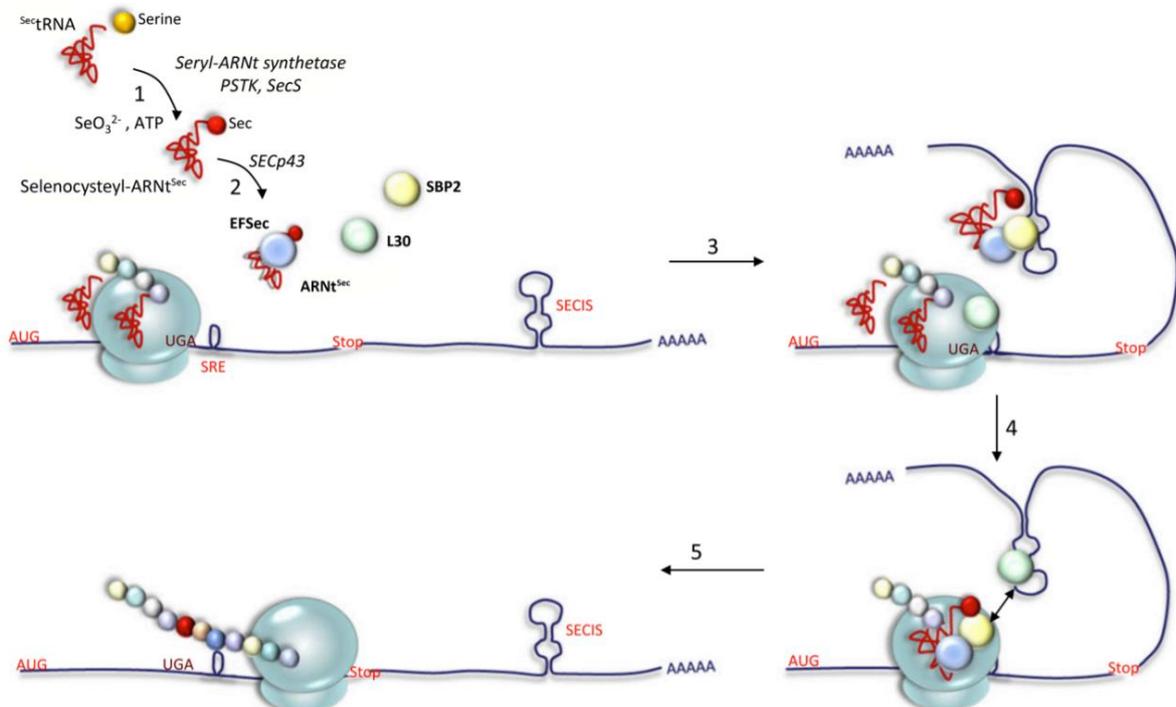


Figure 24 : Synthèse et incorporation de la sélénocystéine chez les eucaryotes La synthèse de la sélénocystéine a lieu sur son ARNt-Sec spécifique 22. L'interaction de l'ARN-tSec avec les facteurs de traduction Sec-spécifiques EFSec et SBP2, en présence de SECp43 (2, 3), permet son recrutement au site A du ribosome. Lorsque le résidu sélénocystéine est incorporé dans la séquence du peptide en réponse au codon reprogrammé UGA, la traduction reprend jusqu'au prochain codon stop (5). Adapté de [223].

Chez la levure, les gènes YFS1, EST3 et ABP140 et le rétrotransposon Ty1 sont aussi régulés par ce mécanisme [219–221]. Pour ces deux gènes, le frameshift +1 est nécessaire à l'expression du gène et à la synthèse de la protéine active.

## **B. La translecture**

Les facteurs de terminaison sont capables de reconnaître les codons stop très efficacement, ce qui permet la synthèse de protéines complètes et fonctionnelles. Toutefois, plusieurs événements modifient le décodage du codon stop en permettant au ribosome de lire ces codons et d'y incorporer un ARNt cognat ou proche cognat, on parle de translecture. Le ribosome continue alors sa traduction en phase jusqu'au prochain codon stop (Figure 23).

### **a. Incorporation d'un ARNt cognat : cas de l'ARNt Sélénocystéine**

En plus des 20 acides aminés standards, il existe deux autres acides aminés incorporés uniquement par translecture : la pyrrolysine et la sélénocystéine. Alors que le premier n'est présent que chez certaines archées et bactéries, le second est également retrouvé chez les eucaryotes. Afin de traiter un exemple d'incorporation d'un ARNt cognat par translecture chez les eucaryotes, nous allons à présent nous intéresser à la sélénocystéine.

Le sélénium, découvert par le chimiste Berzelius en 1817, a été considéré comme un élément toxique pour les mammifères pendant près de 150 ans [224]. Il s'agit pourtant du seul sel minéral ayant pour particularité d'être incorporé directement dans les protéines en tant que 21<sup>ème</sup> acide aminé naturel du code génétique, la sélénocystéine (Sec). Une déficience en sélénocystéine est ainsi associée à des pathologies telles que des maladies cardiovasculaires et myodégénératives, de l'infertilité et de l'ostéochondropathie [225]. Le séléno-protéome comprend 25 membres chez l'humain, impliqués notamment dans la protection contre le stress oxydatif.

La production d'ARNt<sub>[Ser]Sec</sub> nécessite quatre enzymes, contre une seule pour les autres ARNt eucaryotes (Figure 24) [226]. Ainsi, la protéine Seryl-tRNA synthetase (SerRs) permet de charger une Sérine (Ser) sur l'ARNt, puis la protéine PhosphoSeryl-tRNA<sub>[Ser]Sec</sub> Kinase (PSTK) phosphoryle la Sérine. L'acide aminé est ensuite converti en Sélénocysteine par la protéine O-Phosphoserine tRNA<sub>Sec</sub> tRNA Synthase (SepSecS). Pour finir, l'activation du sélénium est réalisée par SPS1/SPS2.

L'ARNt<sub>[Ser]Sec</sub> possède l'anticodon UCA, cognat du codon stop UGA (Figure 24). Il est donc impératif qu'il ne soit pas lié par le facteur d'élongation eEF1A et incorporé au niveau de tous les codons stop UGA, ce qui fait de son incorporation la plus régulée des ARNt. Ainsi, à la place d'eEF1A, c'est le facteur d'élongation eEFSec qui se fixe spécifiquement sur l'ARNt<sub>[Ser]Sec</sub> [227].

Grâce au facteur SECIS Binding Protein 2 (SBP2), il permet l'incorporation de l'ARNt<sub>[Ser]<sup>Sec</sup> au niveau des codons stop UGA de transcrits possédant une séquence structurale Sec Insertion Sequence (SECIS) dans le 3'UTR [228].</sub>

### **b. Incorporation d'un ARNt proche cognat**

Chez les eucaryotes, les acides aminés incorporés par translecture ont tout d'abord été identifiés chez la levure *S. cerevisiae* : le tryptophane (Trp), la Cystéine (Cys) et l'Arginine (Arg) sont retrouvés au niveau du codon UGA ; la Tyrosine (Tyr), la Glutamine (Gln) et la Lysine (Lys) sont retrouvés au niveau des codons UAA et UAG [229]. Ces résultats mettent en avant la possibilité d'un appariement U-G en première position avec l'incorporation d'Arginine au niveau du codon UGA et de Glutamine au niveau des codons UAG et UAA ; et celle d'appariements G-A, G-G et C-A lorsqu'ils sont en troisième position sur la base wobble, avec l'incorporation de Cystéine et de Tryptophane au niveau du codon UGA. Une étude récente réalisée en cellules humaines a permis de montrer l'incorporation d'un nouvel acide aminé au niveau du codon UGA, la Leucine (Leu), première preuve d'incorporation d'un ARNt avec un mis-appariement central [230].

La translecture a initialement été identifiée chez les virus pour qui elle est indispensable et permet la production des polymérase et protéines d'enveloppe [231,232]. Chez les eucaryotes il existe des gènes qui sont également naturellement régulés par ce mécanisme, pour qui l'extension C-terminale induirait des défauts d'activité, une relocalisation cellulaire, une agrégation, voir une dégradation protéique [233–235]. Plusieurs exemples existent chez l'Homme, avec par exemple le gène *OPRL1*, codant le récepteur aux opiacés, et ayant un taux naturel de translecture estimé entre 17 et 31% [236–238]. Pour les autres gènes, on considère que le taux de translecture de base se situe généralement entre 0,01% et 0,1% chez les mammifères et 0,3% chez la levure [239]. Cette fréquence variable est connue pour être influencée par des éléments en cis :

- La nature du codon stop

il est clairement établi que les trois codons stop présentent une efficacité de terminaison différente avec UAA > UAG > UGA [236,240,241]. En effet, les trois codons stop ne sont pas reconnus par les facteurs de terminaison avec la même efficacité [242,243]. Ces efficacités de terminaison différentes s'expliqueraient également par le nombre d'ARNt naturels supprimeurs proches cognats présents dans la cellule et par leur potentiel à être incorporés au niveau du codon stop malgré leur mis-appariement.

- Le contexte nucléotidique entourant le codon stop

De nombreuses études ont mis en évidence l'importance du contexte nucléotidique entourant le codon stop dans l'efficacité de terminaison [244]. Par exemple, lorsque le codon terminateur est dans un contexte « naturel », il est estimé que son efficacité de terminaison est de 0.001–0.1%, contre 0.01–1% pour un codon stop prématuré apparue suite à une mutation non-sens [240,245,246].

Un des éléments les plus importants connu est le contexte 3', notamment la base qui suit immédiatement le codon stop (position +4). Il a d'ailleurs été proposé que le signal de terminaison est en fait un tétranucléotide [247]. Chez les eucaryotes, des études structurales du complexe ribosomal contenant le facteur de terminaison eRF1 ont permis d'expliquer ces observations : lorsqu'un codon stop est placé dans le site A du ribosome, la base en position +4 est tirée par eRF1 dans le site A, où elle est stabilisée contre G626 de l'ARNr 18S [4,248,249]. L'empilement de G626 est moins stable pour les pyrimidines, ce qui contribue à expliquer que les bases cytosine et uracile en position +4 permettent un meilleur taux de translecture. En parallèle, il a été démontré que la base +5 interagit avec la base C1698 de l'ARNr 18S et qu'une purine à cette position permettait de faciliter l'incorporation d'un ARNt proche cognat [200]. La base +8 a aussi été clairement identifiée comme pouvant influencer l'efficacité de translecture dans le cas d'un codon UGA [250]. Dans certaines études, le contexte 3' influençant l'efficacité de terminaison du codon stop s'étend jusqu'à la position +9, avec l'identification de la séquence la plus favorable chez la levure : CARNBA (où R=A/C, N=A/T/G/C, B=C/G/T) [239,251]. En ce qui concerne la séquence en 5' du codon stop, les positions -2 et -1 influencent l'efficacité de terminaison. Il a été notamment démontré que deux adénines à ces positions permettaient d'augmenter le taux de translecture, et cela serait la conséquence du contact direct de ces bases avec l'hélice 44 de l'ARNr 18S [252,253].

- Les structures secondaires

Dans certains cas, la présence de structures secondaires en 3' du codon stop peut moduler l'efficacité de translecture. Ces structures de type pseudo-nœud sont souvent utilisées comme élément stimulateur dans les événements de recodage, en particulier dans les cas de frameshift. Il a été proposé que cette structure, en contraignant la flexibilité de l'ARNm, favoriserait l'incorporation d'un ARNt au détriment du facteur de terminaison [254].

- Les modifications d'ARN

Lors de la traduction, trois ARN jouent un rôle majeur : l'ARNr, l'ARNm et l'ARNt. Depuis 1951, il est acquis que ces ARN sont sujets à des modifications chimiques regroupées globalement sous l'appellation « épitranscriptome » [255]. Depuis leur découverte, ce n'est pas moins de 150 modifications qui ont été mises en évidence chez l'Homme [256,257]. Leur étude a permis de mettre en évidence différents rôles de l'épitranscriptome dans la cellule, dont l'un est de moduler la fidélité d'incorporation des acides aminés durant la traduction. De nombreuses modifications sont ainsi concernées, modifications auxquelles je consacre mon prochain chapitre.

### **III. Les modifications d'ARN dans la fidélité de la traduction**

Alors que l'épigénétique et les modifications post-traductionnelles ont longtemps été mises en avant dans la modulation de l'expression génique, l'épitranscriptomique est de plus en plus étudiée pour ses différents rôles cellulaires. Les modifications d'ARN possèdent deux casquettes dans la fidélité de la traduction : certaines participent à son maintien tandis que d'autres favorisent l'apparition d'erreurs au cours de la traduction. Afin de débiter ce chapitre, je me permets de vous renvoyer à la revue « The Importance of the Epi-Transcriptome in Translation Fidelity » que j'ai écrite, et qui reprend l'ensemble des modifications d'ARN de transfert, ribosomiques et messagers connues pour avoir un impact sur la fidélité de l'incorporation des acides aminés chez les eucaryotes [258]. Je vais ensuite revenir en détails sur les modifications Galactosyl-queuosine-34 (GalQ<sub>34</sub>) et Pseudouridine-35 ( $\Psi_{35}$ ) de l'ARN<sub>t<sub>Tyr</sub></sub>, et la 2'-O-méthylation de l'ARN<sub>r</sub>, en réalisant l'état de l'art des connaissances à leur sujet ainsi que de l'intérêt thérapeutique de leur étude.

#### **1. Rôle des modifications d'ARN dans la fidélité d'incorporation des acides aminés**

Review

## The Importance of the Epi-Transcriptome in Translation Fidelity

Charlène Valadon and Olivier Namy \* 



Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Saclay,  
91198 Gif-sur-Yvette, France; charlene.valadon@i2bc.paris-saclay.fr

\* Correspondence: olivier.namy@i2bc.paris-saclay.fr

**Citation:** Valadon, C.; Namy, O. The Importance of the Epi-Transcriptome in Translation Fidelity. *Non-coding RNA* **2021**, *7*, 51. <https://doi.org/10.3390/ncrna7030051>

Academic Editors: Clément Carré and Damien Brégeon

Received: 27 July 2021

Accepted: 22 August 2021

Published: 27 August 2021

**Abstract:** RNA modifications play an essential role in determining RNA fate. Recent studies have revealed the effects of such modifications on all steps of RNA metabolism. These modifications range from the addition of simple groups, such as methyl groups, to the addition of highly complex structures, such as sugars. Their consequences for translation fidelity are not always well documented. Unlike the well-known m<sup>6</sup>A modification, they are thought to have direct effects on either the folding of the molecule or the ability of tRNAs to bind their codons. Here we describe how modifications found in tRNAs anticodon-loop, rRNA, and mRNA can affect translation fidelity, and how approaches based on direct manipulations of the level of RNA modification could potentially be used to modulate translation for the treatment of human genetic diseases.

**Keywords:** RNA modifications; ribosomes; tRNA; translation fidelity; m<sup>6</sup>A; PSI; Inosine; 2'-O-methylation

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



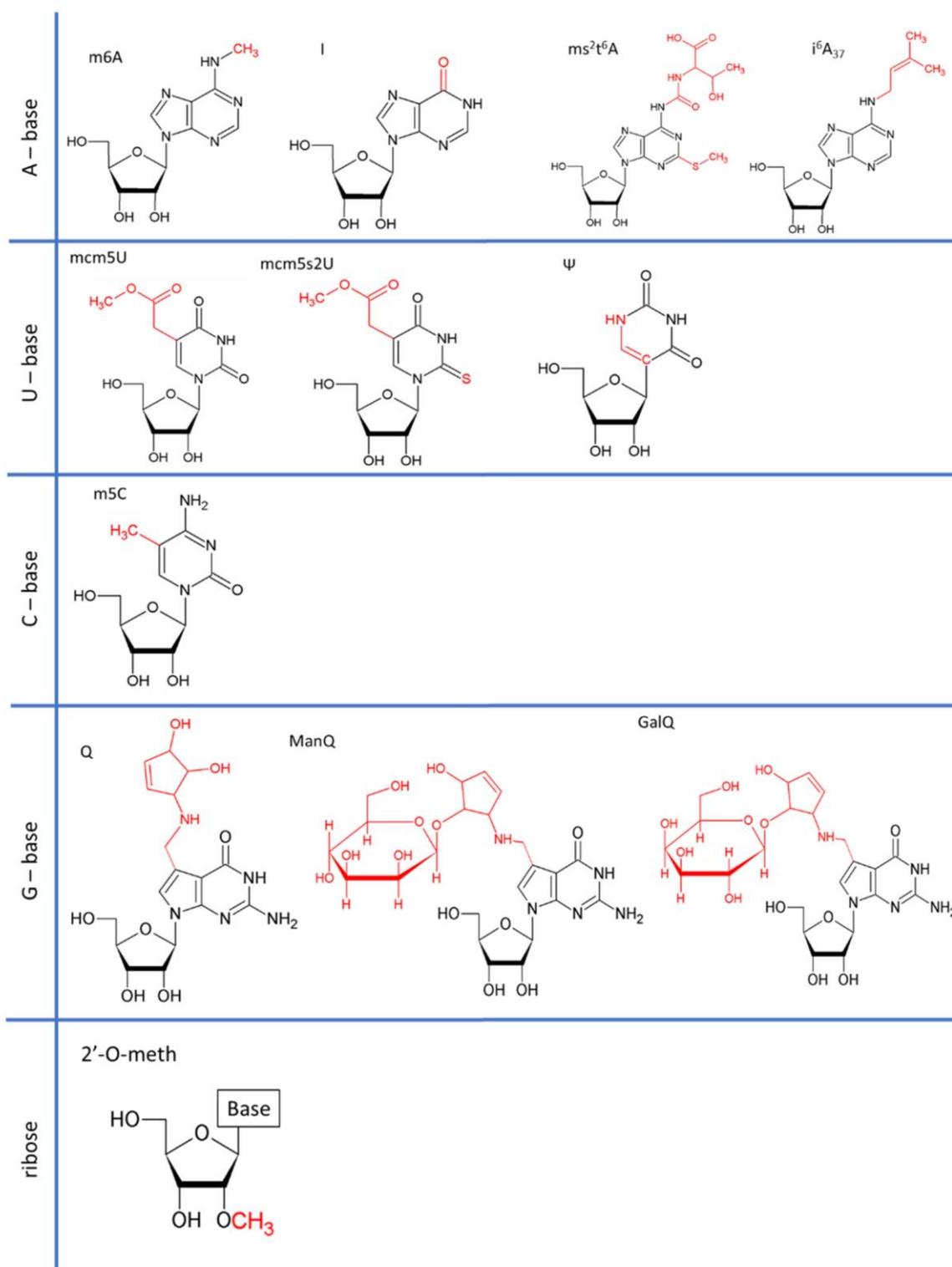
**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

### Introduction

All types of RNA are subject to post-transcriptional modification. Since the discovery of RNA modifications in 1951, more than 150 RNA modifications have been found in coding and non-coding RNAs ranging from the addition of simple groups to the addition of highly complex structures (Figure 1) [1–3].

Their biological consequences are largely unknown, but the discovery that RNAs undergo dynamic, reversible chemical modifications marked the birth of the era of epitranscriptomics. All aspects of RNA metabolism can be affected by RNA modifications [4], either directly, through changes to RNA folding or stability, or indirectly, through the action of “reader” proteins [5–7]. Several excellent reviews have already described in detail the role of RNA modifications in cancers and in cell-fate determination [8–12]. In this review, we mainly focus on the importance of RNA modifications in the two most abundant noncoding RNA families (transfer RNAs and ribosomal RNAs) and their consequences for translation fidelity. Translation proceeds via four steps: initiation, elongation, termination and recycling [13]. Generally, it begins with the fixation of the 43S ribosomal complex to the cap-binding protein eIF4F with the help of numerous translation initiation factors (eIFs) [14,15], followed by scanning of the pre-initiation complex (PIC) to the start codon [16]. The two ribosomal subunits are assembled when the PIC is present at the initiation codon by the universally conserved GTPase eIF5B [17,18]. Elongation begins with the delivery of a tRNA to the ribosomal A-site by elongation factors [19] and continues codon-by-codon until the ribosome reaches a stop codon. When a stop codon enters the ribosome, it is recognized by the termination complex eRF1/eRF3, causing the release of the peptide [20,21]. RLI1/ABCE1 splits the ribosome into its two subunits [22,23], which are then available for a new translation cycle [24]. Translation is not perfectly accurate, as it has a median error rate of 0.01% in humans [25,26]. There are 30 codons in the human code that depend on the incorporation of a near-cognate tRNA (pairing of two of the three bases). Codon-anticodon pairing is known to be flexible at the third position of the codon, but it is clear that RNA modifications alter

translation accuracy [27]. The incorporation of a near-cognate tRNA can occur during elongation or termination, in which case it is known as readthrough [28].



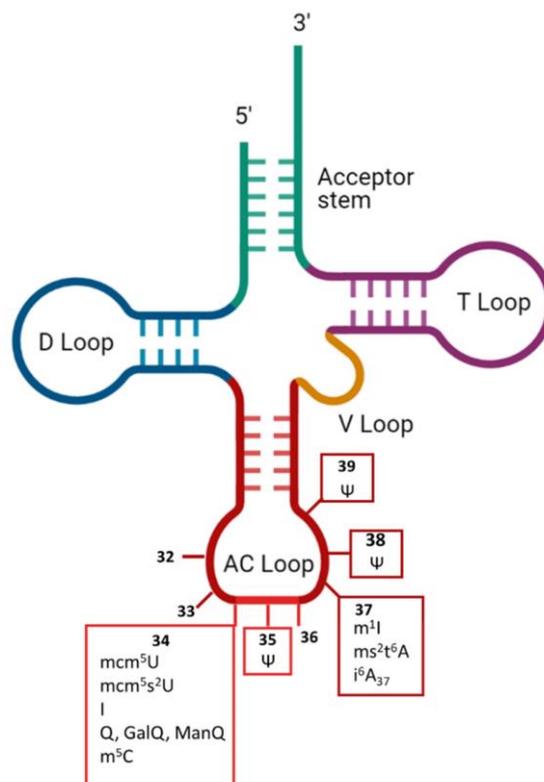
**Figure 1.** Structure of different RNA modifications discussed in this review. In black: basic structure of the base or ribose concerned. Red: chemical modification. m6A: N6-methyladenosine. I: inosine. ms<sup>2</sup>t<sup>6</sup>A: 2-methylthio-N6threonylcarbamoyladenosine. I<sup>6</sup>A: N6-isopentenyladenosine. mcm<sup>5</sup>U: 5-methoxycarbonylmethyluridine. mcm<sup>5</sup>s<sup>2</sup>U: 5-

methoxycarbonylmethyl-2-thiouridine.  $\Psi$ : pseudo-uridine. m<sup>5</sup>C: 5-methylcytosine. Q: queuosine. ManQ: mannosylqueuosine. GalQ: galactosyl-queuosine. Nm: 2'-O-methylation.

In this review we will describe modifications affecting two abundant types of noncoding RNAs—tRNAs and rRNAs—and will consider how such modifications fine-tune translational accuracy. We will also discuss the importance of certain mRNA modifications affecting ribosome fidelity. There is a striking difference in the mode of action of modifications between coding and non-coding RNAs, in that modifications to rRNAs and tRNAs act directly on the folding and activity of the molecule, whereas most of the modifications to mRNAs act via reader proteins. This simplified presentation needs to be modulated because some tRNA modifications are required for the proper action of aminoacyl-tRNA synthetases (aaRS) [29].

### A. Control of Translation Fidelity by Modifications to Cytosolic tRNAs

The type of RNA most frequently modified in cells is tRNA, in which about 17% of nucleotides are modified [30]. Modifications have been found in all five domains of tRNA (i.e., acceptor stem, D-loop, T-loop, V-loop and anticodon loop) (Figure 2).



**Figure 2.** Overview of human cytosolic tRNA anticodon loop modifications known to shape translation fidelity. AC loop: anticodon loop.  $\Psi$ : pseudo-uridine. I: inosine. m<sup>1</sup>I: N<sup>1</sup>-methylinosine. mcm<sup>5</sup>U: 5-methoxycarbonylmethyluridine. mcm<sup>5</sup>s<sup>2</sup>U: 5-methoxycarbonylmethyl-2-thiouridine. Q: queuosine. GalQ: galactosyl-queuosine. ManQ: mannosyl-queuosine. ms<sup>2</sup>t<sup>6</sup>A: 2-methylthioN<sup>6</sup>-threonylcarbamoyladenine. m<sup>5</sup>C: 5-methylcytosine. i<sup>6</sup>A: N<sup>6</sup>-isopentenyladenine. 32–39: anticodon loop positions.

Exceptions exist, but most of the modifications to the D- and T-loops affect the stability or folding of the molecule, whereas those in the anticodon region can affect either the recognition by aaRS or the fidelity of genetic code translation. The anticodon region occupies positions 34–35–36 (Figure 2) that directly base pair to the mRNA codon but are also recognized (together with the acceptor stem) by some aaRS [29]. Some modifications found in the anticodon loop specifically alter the reading frame maintenance. This is the case at position 37, which has been reported to affect the maintenance of the reading frame with the wybutosine yW modification [31]. Interestingly, in humans, no modification has yet been found at position 36, which pairs with the first position in the codon, whereas position 34 of tRNAs, which pairs with the third nucleotide of each codon, is one of the positions at which the various chemical modifications are most numerous [3,32,33]. The reasons for these

differences remain unclear, but may reflect the importance of strict base-pairing at certain positions to prevent incorrect amino-acid incorporation, whereas flexibility may be more acceptable at other positions, at which it may not necessarily cause a change of amino acid, thanks to the redundancy of the genetic code. For a long time tRNA modifications have been considered irreversible. However, in 2016, the work of Fange et al. demonstrated that *ALKBH1* can remove the methyl group from m<sup>1</sup>A<sub>58</sub> in tRNAs [34], opening the possibility that tRNA modifications would be more dynamic than anticipated. Below, we review modifications of the anticodon loop known to affect the translation of the genetic code by affecting the efficacy of cognate or near-cognate tRNA incorporation.

## mcm<sup>5</sup>U<sub>34</sub> modifications

The modifications observed at the U<sub>34</sub> position of tRNA are 5-methoxycarbonylmethyluridine (mcm<sup>5</sup>U), 5-methoxycarbonylmethyl-2-thiouridine (mcm<sup>5</sup>s<sup>2</sup>U) and their derivatives. Catalysis begins with the addition of 5-methoxycarbonylmethyl (cm<sup>5</sup>) to uridine by the elongator complex [35,36]. The cm<sup>5</sup>U is then further modified by the addition of a methyl group by a heterodimeric complex, Trm9-Trm112 (*ALKBH8* in mammals) [37]. Finally, the oxygen atom attached to the C2 atom of the uracil ring may undergo thiolation in a subset of tRNAs, catalyzed by the Ncs2-Ncs6 complex and resulting in a final mcm<sup>5</sup>s<sup>2</sup>U modification [38]. Ultimately, the U<sub>34</sub> position is modified in most eukaryotic tRNAs.

The roles of mcm<sup>5</sup>U and mcm<sup>5</sup>s<sup>2</sup>U in codon–anticodon recognition have been studied in depth in yeast [39–41]. The rate of amino-acid misincorporation has been assessed by dual-luciferase reporter assays in a *S. cerevisiae* *TRM9* mutant (absence of mcm<sup>5</sup>U and mcm<sup>5</sup>s<sup>2</sup>U) [40]. The Trm9 deletion decreases the fidelity of translation specifically for the Arg, Gln, Glu and Lys tRNAs. U<sub>34</sub> modifications, thus promoting discrimination between some cognate and near-cognate codons. A second study in *S. cerevisiae* used mass spectrometry to specify amino-acid misincorporation during readthrough, in the presence or absence of U<sub>34</sub> modifications. A similar phenotype was observed for tRNA<sub>Arg</sub> under *TRM9* gene deletion. However, the results obtained clearly differed from those for tRNA<sub>Gln</sub>, for which U<sub>34</sub> modifications of Gln enhance misincorporation. The impact of *ALKBH8* protein deficiency has been investigated in mammals [37], through the generation of *Alkbh8*<sup>-/-</sup> mice. This mouse line has no mcm<sup>5</sup>U, mcm<sup>5</sup>s<sup>2</sup>U, or mcm<sup>5</sup>Um modifications to tRNAs, these modifications being replaced by the corresponding acid/amide forms: cm<sup>5</sup>U and/or ncm<sup>5</sup>U/ncm<sup>5</sup>s<sup>2</sup>U. Hypomodification of the selenocysteine tRNA (tRNA<sub>Sec</sub>) impairs its ability to decode the UGA stop codon in vitro. U<sub>34</sub> modifications therefore play a role in codon–anticodon recognition in mammals, as in the yeast model. Despite the availability of this mammalian model, studies of the impact of U<sub>34</sub> modifications on other tRNAs have yet to be published.

The physiological importance of U<sub>34</sub> modifications has been demonstrated for mcm<sup>5</sup>s<sup>2</sup>U. This is especially well illustrated by the finding that loss of U<sub>34</sub> modifications can lead to ribosome pausing, promotes proteotoxic stress and protein aggregation [42]. Indeed, the presence of the sulfur atom provides an extended chemical group, which stabilizes A-U or G-U pairing [43]. In conclusion, the weak interaction between A-U is strengthened by U<sub>34</sub> modification, favoring translation fidelity, whereas stabilization of the unconventional G-U base-pairing favors the incorporation of near-cognate tRNAs.

## I<sub>34/37</sub>

Inosine (I) results from deamination of the C6 of adenine [44]. Its editing is catalyzed by adenosine deaminases, which act directly on tRNA (ADATs) [45]. Inosine is present at tRNA positions 34 (8 tRNAs) and 37 (tRNA<sub>Ala</sub>) in eukaryotes [45–47]. I<sub>34</sub> is catalyzed by the heterodimeric enzyme ADAT (hetADAT), consisting of ADAT2 and ADAT3. I<sub>37</sub> is deaminated by *ADAT1* and further modified by methylation (m<sup>1</sup>I<sub>37</sub>) catalyzed by the tRNA methyltransferase *TRM3* [48]. A deficiency of I<sub>34</sub> has been reported to affect human health, with patients presenting bi-allelic *ADAT3* mutations displaying intellectual disability [49].

I<sub>34</sub> has been shown to enable tRNAs to pair with U, A and C nucleotides at the wobble position [27]. The chemical origin of the I<sub>34</sub> base-pairing effect is the replacement of the hydrogen-donating amino group in the C6 position by a hydrogen-accepting oxygen [44].

Inosine contributes to the extension of the genetic code [50].

Presence of inosine in humans has been linked to a deviation of codon usage between prokaryotes and eukaryotes [51]. Indeed, bioinformatic analysis has shown that TAPSLIVR amino-acid stretches are more efficiently

expressed with I<sub>34</sub>-tRNAs. This has resulted in a greater abundance (eight-fold) and length of such proteins in eukaryotic proteomes than in prokaryotic proteomes, which have only one I<sub>34</sub>-tRNA.

## Q<sub>34</sub> and Its Derivatives

Queuosine (Q), or 7-deazaguanosine, is a modified analog of guanosine incorporated at G<sub>34</sub> of GUN anticodon tRNAs [52]. In mammals, Q<sub>34</sub> is hypermodified at the Tyr and Asn anticodons, by the addition of a sugar (a galactose (GalQ) and a mannose (ManQ), respectively) to the C4 hydroxy group of the cyclo-pentenediol (Figure 1) [53–55]. Eukaryotes cannot synthesize Q de novo, and must therefore obtain it (or its derivative, queuine) as a micronutrient from the gut microbiota or through dietary intake [56–58]. Human cells must, therefore, take up queuine, to replace guanosine by queuosine. Interestingly, queuosine levels change during development [59,60]. It remains unclear whether these changes reflect differences in the need for translation fidelity during development or whether they simply reflect the availability of queuine in the diet and/or the possibility of its synthesis by the microbiota. Queuosine has also been associated with cell differentiation, the symptoms of poly-phenyl-ketonurea disease, cancer progression and microbiote diversity. Mannosyltransferase responsible for the generation of ManQ has been isolated from rat liver, but no galactosyltransferase has been identified yet for GalQ [61]. It remains unclear whether these sugar-modified tRNAs are involved in translation (i.e., are they still able to enter the ribosome?) or whether they act as regulatory RNAs, as reported for tRNA-derived fragments [62].

Despite its key position in the anticodon, the role of Q in translation fidelity has been little studied. The impact of Q on codon–anticodon pairing has been physically measured; the C-G pairing is slightly more stable than Q-G, and the Q-U pairing is more than twice as stable as C-U [63]. However, it remains unclear how Q is distinguished from G in front of C or U [64]. Q has been shown to modulate translation accuracy in *E. coli* [65]. However, a study in eukaryotes investigated the role of Q in tRNA<sub>His</sub> decoding in *Xenopus laevis* oocytes [66]. GUG or QUG–tRNA<sub>His</sub> from *D. melanogaster* was injected into the oocytes, and their ability to decode CAC or CAU codons was assessed. The results underline that Q<sub>34</sub>-tRNA<sub>His</sub> decoded CAU more efficiently than the CAC codon, contrary to what was found for G<sub>34</sub>-tRNA<sub>His</sub>. Additional studies are required to clarify the role of Q and to characterize the role of its hypermodified derivatives, GalQ and ManQ, in translation fidelity.

## m<sup>5</sup>C<sub>34</sub>

5-methylcytosine (m<sup>5</sup>C) is an additional methyl group on C5 present on the cytosine at the position 34 of tRNA<sub>Leu</sub>(CAA). In humans, the methyltransferase hTrm4 is responsible for this methylation position [67], whereas its yeast homolog Trm4 also modifies positions 48 and 49 [68]. In *S.cereviae*, tRNA<sub>Leu</sub>(CAA) is surnamed “tRNASUP53” for its abilities to suppress the UAG stop codon. Interestingly, the presence of m<sup>5</sup>C<sub>34</sub> on tRNASUP53 depends on the integrity of a 32 bases intron in the pre-tRNA.

The importance of m<sup>5</sup>C<sub>34</sub> in tRNA<sub>Leu</sub>(CAA) mis-incorporation has been assessed in the yeast model [68]. The suppressor activity of tRNASUP53 has been tested on medium minus tryptophan, using a trp1-J(Am) marker. The results showed that in the absence of m<sup>5</sup>C<sub>34</sub> the cells do not grow on the medium, highlighting a decrease in the tRNASUP53 suppressor activity. Thus, m<sup>5</sup>C<sub>34</sub> in tRNA<sub>Leu</sub>(CAA) is considered to act as an enhancer of tRNA misincorporation modification. Unfortunately, for now, no replica of these results exists in a human or another eucaryotic model.

The lack of interest in this modification could be due to the low level of leucine incorporation at the UAG stop codon which has not been reported in any recent study; moreover, hTrm4 has not been linked to any disease in humans.

## ms<sup>2</sup>t<sup>6</sup>A<sub>37</sub>

2-methylthio-N6-threonylcarbamoyladenine (ms<sup>2</sup>t<sup>6</sup>A) biosynthesis is well known: the N6-threonyl carbamoyl adenosine (t<sup>6</sup>A) is methyl-thiolated to generate the ms<sup>2</sup>t<sup>6</sup>A at position 37 of the tRNA<sub>Lys</sub>(UUU) which is the only tRNA concerned by this modification in humans. The methyl-thiolation is done by *Cdkal1* [69]. There is some structural evidence that ms<sup>2</sup>t<sup>6</sup>A at position 37 inhibits a noncanonical U33-A37 interaction and is likely to be needed to compensate for the relatively weak U-turn remodeling properties of mcm5s2U<sub>34</sub> [70].

For now, studies of the role of  $m^2t^6A$  in translation fidelity have been performed only in a bacteria model [71]. In *B. subtilis*, absence of  $mcm5s2U_{34}$  ( $\Delta yqeV$ ) revealed that  $m^2t^6A_{37}$  in  $tRNA_{Lys}(UUU)$  prevents the misreading of its cognate codons AAA and AAG, especially when the rate of translation is high. Unfortunately, ten years after this study, these encouraging data have not yet been reproduced in a eukaryotic model. However, in the same study, a knock-out of *Cdkal1* in mouse has been performed and resulted in reduction of glucose-stimulated proinsulin synthesis. Thus, it has been hypothesized that, in the absence of *Cdkal1*,  $tRNA_{Lys}(UUU)$  misreading increases, limiting the generation of mature insulin and C-peptide which relies on Lys 1 and 2 of proinsulin. This phenotype could explain the molecular pathogenesis of type 2 diabetes in patients carrying *Cdkal1* risk alleles, new evidence for the importance of tRNA modifications in human health.

### $i^6A_{37}$

N<sup>6</sup>-isopentenyladenosine ( $i^6A_{37}$ ) results in the addition of an isopentenyl group onto N<sup>6</sup> of  $A_{37}$  by an isopentenyl-transferase (IPTase). The nature of concerned tRNAs varies from one organism to another. In humans,  $i^6A_{37}$  has been found to be synthesized by *TRIT1* in cytosolic  $tRNA_{Ser}(UGA, AGA, \text{ and } CGA)$  and  $tRNA_{[Ser]Sec}$  [72]. The last is relatively hypomodified (~40%). As a member of near-wobble modification,  $i^6A_{37}$  is proposed to enhance A:U stacking by stabilizing the Watson-Crick base pair but the chemical reasons are not yet clear [73].

The role of  $i^6A_{37}$  on translation fidelity was assessed in a *S. cerevisiae* model

(Blanchet 2017). Mod5p enzyme synthesizes the  $i^6A_{37}$  in the Tyr and Cys tRNAs. The ability of Tyr and Cys tRNAs to readthrough respectively UAA/UAG and UGA stop codons has been assessed in a  $\Delta Mod5p$  strain by mass spectrometry. In this strain, both tRNAs are less efficient at being mis-incorporated, revealing that  $i^6A_{37}$  modification acts as an enhancer of translation plasticity in *S. cerevisiae*. In another study in *S. pombe*, the role of  $i^6A_{37}$  has been assessed using a  $\beta$ -galactosidase codon-swap reporter [74]. This confirms that  $i^6A_{37}$  increases the incorporation of  $tRNA_{Tyr}$  at a near-cognate codon. Moreover, it shows that the modification enhances the incorporation of  $tRNA_{Cys}$  at a cognate codon. Altogether, these data suggest that  $i^6A_{37}$  promotes decoding activity generally.

To date, no study has been made in humans about the role of  $i^6A_{37}$  in translation fidelity. However, it is clear that *TRIT1* mutation is associated with severe diseases. Indeed, it has been defined as a tumor suppressor [75]. The mutation of this protein is also associated with encephalopathy and myoclonic epilepsy pathology [76].

### $\Psi_{35/38/39}$

Pseudouridine ( $\Psi$ ) is an isomer of uridine in which uracil binds the ribose via a C1'-C5 rather than a C1<sup>0</sup>-N1 bond [77]. The pseudouridine synthases responsible for the catalysis of tRNA  $\Psi$ s belong to the PUS RNA-independent family.  $\Psi$ s are mainly found in the anticodon loop of tRNAs, at positions 35, 38 and 39. Catalysis is performed by two enzymes: PUS7 and PUS3. The effects of  $\Psi_{35}$ ,  $\Psi_{38}$  and  $\Psi_{39}$  on translation fidelity probably stems from their ability to improve the stacking of double-stranded RNA over that achieved with the uridine isomer [78]. More precisely, the presence of  $\Psi$  stabilizes the C3'-endo conformation, creating an additional N1 H-bond donor [79]. Thanks to these chemical characteristics,  $\Psi$ s involved in ( $\Psi_{35}$ ) or next to ( $\Psi_{38}$ - $\Psi_{39}$ ) the anticodon enhances the stability and structure of pairing.

Studies in yeast (*S. cerevisiae*) have shown that the deletion of *PUS3* triggers an increase in misincorporation events relative to the WT strain [39,80]. Even  $\Psi_{38}$  and  $\Psi_{39}$ , at the end of the anticodon loop increase miscoding frequency. We have also studied the contribution of  $\Psi_{35}$  to the ability of  $tRNA_{Tyr}$  to act as a near-cognate tRNA for stop-codon readthrough in a  $\Delta PUS7$  strain [39]. Mass spectrometry has shown that tyrosine incorporation is less efficient in the absence of the *PUS7* gene. This finding highlights the importance of  $\Psi_{35}$  for the ability of  $tRNA_{Tyr}$  to read through UAA and UAG codons. Together, these studies demonstrate that  $\Psi$  modifications to the anticodon loop of tRNAs have a major effect on translation fidelity. The tRNA pseudo-uridine synthases *PUS3* and *PUS7* have both been implicated in human health problems; they cause different diseases, and notably intellectual disability [81,82].

It is interesting to mention that some modifications away from the anticodon loop seems to play a role in translation. For example, absence of  $ac4C_{12}$  and  $dU_{20}$  of Leucine et Serine tRNAs causes a reduction in A-site occupancy at the corresponding codons. Widespread changes in the A-site occupancy have been also observed in

the absence of  $m^2_2G_{26}$  [83]. Since these modifications are not in the anticodon loop, they may influence either the charging of these tRNAs or their ability to bind stably the ribosomal A-site. Several studies suggest that interactions inside the anticodon loop are crucial for a correct modification of the tRNAs. These interconnections between modifications create a complex network. For example, in eukaryotes, the  $\gamma W$  formation at  $m^1G_{37}$  of tRNA<sup>Phe</sup> is greatly stimulated by the presence of  $Cm_{32}$  and  $Gm_{34}$  [84–86]; the presence of  $i^6A_{37}$  or  $t^6A_{37}$  in tRNA<sup>Ser</sup> stimulates the formation of  $m^3C_{32}$  [87];  $m^5C_{38}$  is stimulated by prior  $Q_{34}$  formation [88]; and  $I_{34}$  editing in tRNA<sup>Thr</sup>(AGU), is stimulated by prior C to U editing at position 32 [89]. Future structural studies of tRNA-modification enzymes complexes will probably explain such dependencies.

It is also possible that tRNA modifications act through other processes such as RNA sequestration, or the generation of tRNA-derived fragments (tRFs/tiRFs) [90] obtained by the endonuclease cleavage of tRNAs (mainly in the D-Loop, T $\Psi$ C-Loop and anticodon-Loop), either in normal or stressed conditions. Such tRFs/tiRFs are involved in various physiological and pathological processes by modulating RNA stability or translation [62]. There is a clear link between tRNA modifications and tRFs/tiRFs generation. For example, Dnmt2, which methylates the  $C_{38}$  of tRNA<sup>Asp</sup> [91], protects tRNA from the degradation [92]. It has been also shown that *ALKBH3* activity results in removal of  $m^1A$  and  $m^3C$  modifications, leading to a sensitivity to angiogenin cleavage [93].

## B. Role of rRNA Modifications in Translation Fidelity

Ribosomal RNA is the most abundant non-coding RNA in the cytoplasm. It is the main constituent of the ribosome. In total, 200 modification sites have been mapped on the human ribosome, in which about 2% of the nucleotides are modified [30,94–96]. These modifications can modulate all stages in the life of the rRNA, from ribosome biogenesis to translation accuracy [97]. The most frequent modifications observed are pseudo-uridines and 2'-O-methylations, although base methylation and acetylation have been reported [30]. We focus here on the description of the two main modifications of rRNAs known to affect translation fidelity.

### 2'-O-methylation (Nm)

2'-O-methylation (Nm) is a modification in which a sugar is added to the 2'C hydroxyl group of the nucleotide. The chemical impact of Nm on RNA has been investigated by several studies. It has been reported that Nm biases the sugar pucker equilibrium in favor of the C3'-endo conformation of pyrimidines [98]. Intra-residue steric repulsion occurs between the Nm, the 3'-phosphate, and the 2-carbonyl groups in the C2'-endo conformation, favoring the C3' form. The Nm modification may, therefore, either stabilize or modulate RNA structures.

In human cells, Nm is mediated by the ribonucleoprotein complex consisting of the methylase fibrillarin (FBL) and the guide RNA (C/D box snoRNA) specific to the methylation site [99]. FBL is an essential protein, but it can be partially inactivated, leading to a decrease of up to 50% in the number of methylation sites in human cells [100,101]. More than 100 2'-O-methylation sites have been mapped on rRNAs, independently of nucleoside identity [102,103].

The role of Nm in miscoding has been explored in human cancer cells [104]. FBL overexpression, leading to hypermethylation of the ribosome, has been shown to trigger an increase in amino-acid incorporation at cognate or near-cognate codons. It is difficult to identify the 2'-O-methylation sites responsible for this phenotype, because site-specific inactivation experiments have not been performed yet on human cells. As FBL methylates all the sites, the only solution would be to inactivate each snoRNA specifically, one-by-one. A study of this type has been performed in yeast, in which knockouts of the various guide C/D box snoRNAs have been performed [105]. The impact of the loss of each snoRNA was evaluated by measuring stop codon readthrough efficiency. Nm- $C_{1639}$  was identified as the most important of the Nm sites tested. The abolition of Nm at this P-site triggers a slight increase in UAG readthrough. This work revealed a role for Nm- $C_{1639}$  in the maintenance of ribosome fidelity during termination. There is now a need to reproduce such systematic analyses of Nm sites in humans.

The role of rRNA's Nm extends beyond miscoding events. The downregulation of FBL has been shown to alter IRES-dependent initiation and frameshifting. A single deletion of  $Am_{398}$  or  $Gm_{3745}$  in the 28S rRNA or of  $Am_{163}$  in the 18S rRNA is embryo-lethal in zebrafish [106]. Moreover, FBL overexpression has been reported during the differentiation of human stem cells, and in several cancer studies, suggesting a central role in these processes [100,104,107,108].

### Pseudouridine

With the exception of position 50 in the 5S rRNA that is catalyzed by the enzyme *PUS7*, the formation of  $\Psi$ s in rRNA is catalyzed by a ribonucleoprotein complex composed of the pseudo-uridine synthase DKC1 associated with H/ACA box snoRNAs [109,110]. In human rRNAs,  $\Psi$ s are mapped with a  $\Psi$ /U ratio of 5–7%, with a total of about 100 sites [109,111–114]. DKC1 is an essential protein, and mutations of its gene have been linked to X-linked dyskeratosis congenita disease. Patients may display alterations to skin color, nail dystrophy, bone marrow failure, and an increase in the risk of developing cancer and pulmonary fibrosis, although it is not clear whether these effects are related to the absence of  $\Psi$  from rRNA [110].

The role of  $\Psi$ s in miscoding has been investigated in human cells [115]. SNORA24 (ACA24), a H/ACA box snoRNA guiding the  $\Psi_{609}$  and  $\Psi_{863}$  on the 18S rRNA, has been downregulated in HCC cells [116]. An analysis of ribosomal pre-translocation complex dynamics by sm-FRET indicated changes in tRNA conformation in the A-site in ribosomes lacking  $\Psi_{609}$  and  $\Psi_{863}$  relative to wild-type ribosomes, depending on the tRNA entering the ribosome. It has also been shown that lower levels of SNORA24 expression increase amino-acid misincorporation by 10%–20% and readthrough by 15% at UGA, but not at UAG codons.

The way in which  $\Psi$ s in rRNAs decrease the accuracy of translation seems to depend on their abundance in the peptidyl transferase and decoding centers of the ribosome [77].  $\Psi$ s are known to generate an additional N1 H-bond donor and to stabilize the C3<sup>0</sup>-endo conformation [79]. This enables  $\Psi$ s to increase RNA–RNA stability in the fidelity centers of the ribosome [117]. A decrease in the number of  $\Psi$  sites is, thus, accompanied by ribosome destabilization, resulting in a decrease in ribosome fidelity.

### C. mRNA Modifications Influence the Reading of the Genetic Code

Many studies over the last decade have revealed the importance of mRNA modifications. These modifications are highly dynamic, with eraser proteins able to eradicate the modifications from the mRNA. The dynamic aspect of the modifications allows integration in a very efficient manner of the RNA metabolism and translation to the physiological state of the cell, considering the appearance of possible stresses.

#### Inosine

The formation of inosine on mRNAs is catalyzed by the adenosine deaminases ADAR1 and 2 [118]. The inosines of mRNAs, like those of tRNAs, play a major role in expansion of the genetic code, with 5072 identified editing sites in human coding sequences [119].

One of the best known examples of the importance of A-to-I editing in mRNA is the modification of the glutamate receptor subunit B (GluRB) precursor messenger RNA: CAG (Gln) → CIG (Arg) in exon 11. This site is modified by ADAR2 and is essential to ensure the impermeability of the glutamate receptor to Ca<sup>2+</sup> ions [120]. A defect of this inosine site has, notably, been shown to contribute to neuronal death in amyotrophic lateral sclerosis [121]. The dysregulation of ADAR1 and 2 has also recently been observed in human hepatocellular carcinoma [122]. Patients with an upregulation of ADAR1 and a downregulation of ADAR2 have higher incidences of tumor recurrence and liver cirrhosis, and shorter disease-free survival times. These dysregulations are linked to changes in the number of inosine sites, with, in particular, hyper-editing of the *FLNB* mRNA and hypo-editing of the *COPA* mRNA [122]. Finally, ADAR1 seems to act as an oncogene, whereas ADAR2 acts as a tumor suppressor, in hepatocellular carcinoma.

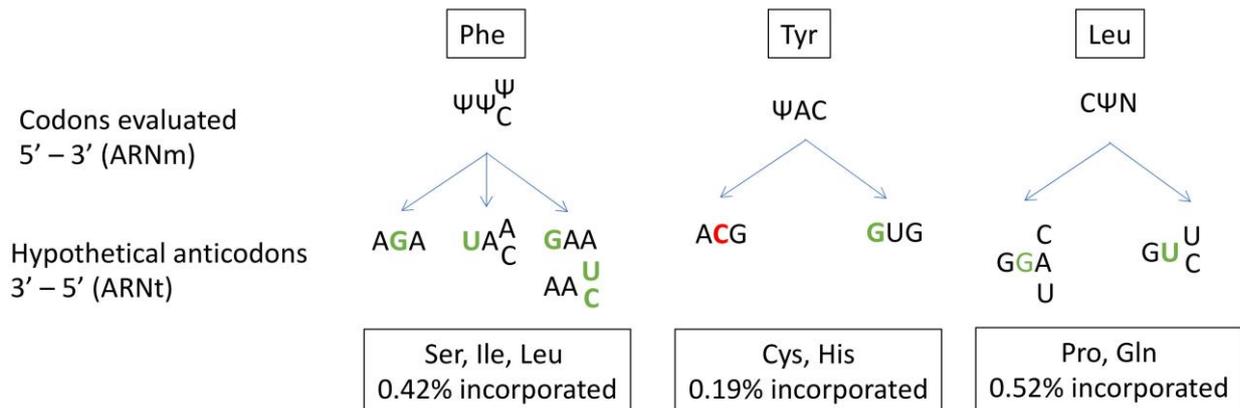
Inosine in mRNAs is known to modulate alternative splicing and stability, but it clearly also plays an essential role as an enhancer of near-cognate tRNA incorporation, ensuring the activity of some proteins [118]. On the other hand, we did not find any significant difference in ribosome profiling between edited and non-edited mRNA in term of translation efficiency in *A. Thaliana* mitochondria [123]. The conservation of these essential CDS sites, rather than the cognate codon with a G, remains to be evaluated.

#### Pseudouridine

Unlike the  $\Psi$ s found in rRNA, the reaction generating those found in mRNA is catalyzed by pseudo-uridine synthases, which are RNA-independent proteins [113,124–126], although the existence of some box H/ACA snoRNAs complementary to mRNAs raises the possibility that RNA-dependent pseudo-uridylation of mRNAs also occurs [127,128]. mRNA  $\Psi$ s are known to be modulated under cellular stress and during development, but no  $\Psi$  reader or eraser has yet been described [129]. Within the translated and untranslated regions of mRNAs, pseudo-

uridine is present with a  $\Psi$ /U ratio of 0.2–0.6%, and 1889 sites have been identified by N3-CMC–enriched pseudo-uridine sequencing [113]. More than 60% of pseudo-uridine residues are located within the coding sequence, suggesting a link with translation [130,131].

In prokaryotes, several studies have described the ability of  $\Psi$  to alter base-pairing and induce misincorporation [132–134]. However, far fewer studies have been performed on human cells [131]. Amino-acid misincorporation in front of a “U-codon” has been shown to occur at a rate of 1%. The presence of  $\Psi$  in mRNA induces the substitution of Ser, Ile or Leu for Phe at UUU/C codons; Cys or His substitution for Tyr at UAU/C codons; and Pro or Gln substitution for Leu at CUA/U/C/G codons (Figure 3).



**Figure 3.** Prediction of anticodon substitution in front of codons with pseudo-uridine ( $\Psi$ ), based on the amino acid misincorporated in the study of Eyer et.al. (2019). Nucleotides in green indicate a mismatch in front of  $\Psi$ . Nucleotides in red indicate a mismatch next to  $\Psi$ . N: A, U, C or G base. Possibilities of codon-anticodon pairings with more than one mismatch are not represented.

Given the high frequency of  $\Psi$  in mRNA and its role in near-cognate tRNA recognition,  $\Psi$  modifications probably make a major contribution to translation fidelity. A closer look at codon/anticodon base-pairing in the case of the misincorporation of Cys at a Tyr codon reveals a central mismatch between an A and a C. This unfavorable interaction is probably compensated for by the strong ability of  $\Psi$  to stabilize the codon/anticodon structure by stacking interactions. Indeed,  $\Psi$  is known to enhance RNA structure stability. Despite its ability to form a supplementary N1-hydrogen bond,  $\Psi$  has the same Watson-Crick base-pairing properties as U [79].

## $m^6A$

The N6-methyladenosine ( $m^6A$ ) modification involves the addition of a methyl group to the N atom linked to the C6 of adenosine. Chemical predictions of the impact of  $m^6A$  on RNA–RNA base-pairing suggest a disruption of this interaction due to the methyl group [135]. Indeed, this group must adopt an anti-conformation in the context of A–U pairing. This conformation is less energetically favorable than the syn conformation, leading to destabilization of the RNA–RNA accommodation. In humans, more than 12000  $m^6A$  sites are estimated to be present on 7000 mRNAs [136,137]. About 35% of  $m^6A$  sites are located within the coding region [138].  $m^6A$  is a dynamic modification that has been reported to interact with several enzymes called readers [139,140]. A heterodimeric methylase complex (*METTL3-METTL14*) is responsible for adding the methyl group. Once modified, the site can be recognized by reader proteins, most of which belong to the YTH-domain protein family (YTHDC and YTHDF), or eraser proteins, which are demethylases (such as *FTO* and *ALKBH5*).

The impact of  $m^6A$  at the first or second position of the codon has been measured by quench flow techniques [141]. This modification delays tRNA incorporation, by slowing tRNA accommodation at site A of the ribosome. However, it has been reported that  $m^6A$  at the middle position of the codon has a lesser effect on pairing for near-cognate than for cognate tRNAs [142]. This difference in kinetics suggests that tRNA misincorporation rates are likely to be higher in the presence of  $m^6A$  at the middle position. However, contrary to these findings for prokaryotic systems, mass spectrometry assays in eukaryotes (wheat germ and HEK293T) identified no miscoding effect of the  $m^6A$  modification [143,144]. The method used for eukaryote systems may be insufficiently sensitive to detect misincorporation in the context of cognate/near-cognate competition. Indeed, the same study found no

miscoding effect of  $\Psi$  modification, contradicting the findings of another team published in the same year [131]. In the face of these conflicting data, further studies are required to clarify the impact of  $m^6A$  on miscoding events.

$m^6A$  is one of the most commonly studied RNA modifications because of its broad influence on RNA maturation and degradation, RNA-protein interactions and translation efficiency, implicating this modification in a number of different biological processes [138,145–149]. Focusing on human health, altered  $m^6A$  levels have been implicated in the regulation of the expression of genes relating to cancer pathogenesis and development [150].

## $m^5C$

5-methylcytosine ( $m^5C$ ) is a cytosine with an additional methyl group on C5. Like  $m^6A$ ,  $m^5C$  is a dynamic modification, with writer, reader and eraser proteins. *NSUN2* and the Aly/REF export factor are the principal  $m^5C$  mRNA writer and reader proteins, respectively [151].  $m^5C$  has been mapped on several transcriptomes in humans [152–155]. Although *NSUN2* and *NSUN6* are well-known tRNA-modification enzymes, they also appear to modify mRNA. The number of  $m^5C$  sites in mRNA has been estimated at about a thousand by bisulfite RNA sequencing [156]. Interestingly, viral RNAs are particularly rich in  $m^5C$  modifications, suggesting that it could play a role in the discrimination of endogenous and exogenous RNAs.

The question of the impact of  $m^5C$  on translation has been addressed by ribosome profiling in Hues9 human embryonic stem cells with a knockout of *NSUN6* gene [157]. No global translational defect was observed, but the absence of *NSUN6* was found to trigger stop codon enrichment at the P-site of the ribosome, possibly after readthrough, and an increase in ribosomes bound to the 3'UTR of mRNAs modified by *NSUN6*. These data suggest that  $m^5C$  sites in the 3'UTR of mRNA enhance translation termination efficiency by decreasing the readthrough rate. It remains unclear how  $m^5C$  in the 3'UTR affects termination. Another study in HEK293T cells assessed the impact of  $m^5C$  at the three codon positions by mass spectrometry [144]. None of the three positions was found to modulate the misincorporation of amino acids.

$m^5C$  is linked to human health. Indeed, *NSUN2* mutations are associated with growth retardation, neurodevelopmental defects, and have been identified as a possible treatment target for tumors [153,158–161]. Moreover, the  $m^5C$  reader and eraser proteins cited above are known to display altered expression levels in various types of cancer [151].

## D. Manipulation of RNA Modifications to Treat Human Diseases

The field of RNA modifications is undoubtedly a very promising area in human therapy. Synthetic modified mRNAs can be used in diverse therapeutic contexts, including cardiac regeneration, asthma, cystic fibrosis or lung diseases [162–165]. The best-known application is probably the current COVID-19 vaccines of Pfizer/BioNtech and Moderna. In these mRNA-based vaccines, all the uridine residues are replaced by N1-methyl-pseudouridines to prevent the recognition of the vaccine mRNA by host RNA sensors and to stimulate translation initiation by attenuating eIF2 $\alpha$  phosphorylation [166,167]. For those interested in eIF2 $\alpha$  stress response and translational regulations, please see the following review [168]. It is also possible to target mRNAs directly, through the use of artificial snoRNAs to replace a U residue with a  $\Psi$  at a specific position [132]. In this example, changing the U to a  $\Psi$  at the first position of a premature termination codon leads to the incorporation of several amino acids rather than a stopping of translation. This could restore production of the full-length protein, thereby correcting the genetic defect.

From another standpoint, RNA modifications affect diverse biological processes, and the correct incorporation of many of these modifications, at the correct sites, is required for normal development. Alterations to these modifications have been implicated in several diseases, including cancers and resistance to therapy of melanoma cells [169]. The role of  $m^6A$  in cancer is very well documented, and  $m^5C$  has also emerged as a major player in cancer development [170–172]. Given the crucial roles of writer, reader and eraser proteins in cell homeostasis, these proteins have naturally emerged as potential treatment targets [173]. Ribosome modifications are also of potential interest in this context, and *DKC1* and *FBL* may serve as potential anticancer targets, as shown by the changes in their levels of expression in many cancers [104,174].

As discussed above, it is possible to target mRNA with an H/ACA snoRNA for the incorporation of a  $\Psi$  at a specific position. This approach could be used in genetic diseases caused by the presence of a premature

termination codon (PTC). Proof-of-concept has been obtained through the demonstration that replacing the U of the stop codon with  $\Psi$  converts the stop codon into a sense codon [132]. Indeed, serine and threonine were found at  $\Psi$ AA and  $\Psi$ AG codons, whereas tyrosine and phenylalanine were found at  $\Psi$ GA codons. In principle, it should be possible to change the modification status of tRNAs to modulate translation fidelity. This would be particularly useful in diseases linked to the appearance of a premature stop codon, which are treated with readthrough-inducing molecules. These molecules, such as aminoglycosides, target the ribosome, enabling it to read through the stop codon, but it should be possible to improve the incorporation of specific tRNAs by altering their modifications [175]. However, in this case, a delicate balance must be found between promoting high levels of readthrough without compromising normal tRNA usage. The recent publication describing the stimulation of UGA readthrough by inhibiting the Cm<sub>34</sub> modification on tRNA<sub>Trp</sub> with 2,6-diaminopurine (DAP) paves the way for the development of such therapeutic approaches [176]. We are still at the dawning of the epi-transcriptomic era, particularly as concerns human treatments, but this field promises to yield extraordinary advances.

## Conclusions

With so many unanswered questions both in terms of molecular mechanisms and physiological consequences, the field of RNA modifications gains more and more interest. One of the current limitations is the difficulty to identify and quantify RNA modifications, especially in highly structured molecules such as tRNAs or rRNAs. Mass spectrometry approaches are extremely accurate, but require highly purified molecules and are hardly quantitative on a large scale. Deep-sequencing (NGS) approaches require either chemical modification of the RNA or immunoprecipitation with a specific antibody, with the associated problems of specificity [177]. Direct RNA sequencing (nanopore) holds a lot of promise with the possibility of directly detecting modified positions. However, this still requires the development of bioinformatics tools to allow a reliable and quantitative analysis. There is no guarantee that all tRNAs will be accessible, although preliminary report exists suggesting that some will be [178].

**Author Contributions:** Writing the manuscript C.V. and O.N.; Funding acquisition, O.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** ANR ANR Rescue\_ribosome program (17-CE12-0024-01) and ActiMeth (ANR-19-CE120004). This work was also supported by the epiRNA program from I2BC.

**Institutional Review Board Statement:** Not relevant.

**Informed Consent Statement:** Not relevant.

**Data Availability Statement:** Not relevant.

**Acknowledgments:** The English of this manuscript was corrected by Alex Edelman & Associates. We acknowledge support from the ANR ANR Rescue\_ribosome program (17-CE12-0024-01) and ActiMeth (ANR-19-CE12-0004). CV is supported by a grant from FRM. This work has also been supported by the I2BC epiRNA program.

**Conflicts of Interest:** The authors declare no conflict of interest.

Références de la revue : voir annexes

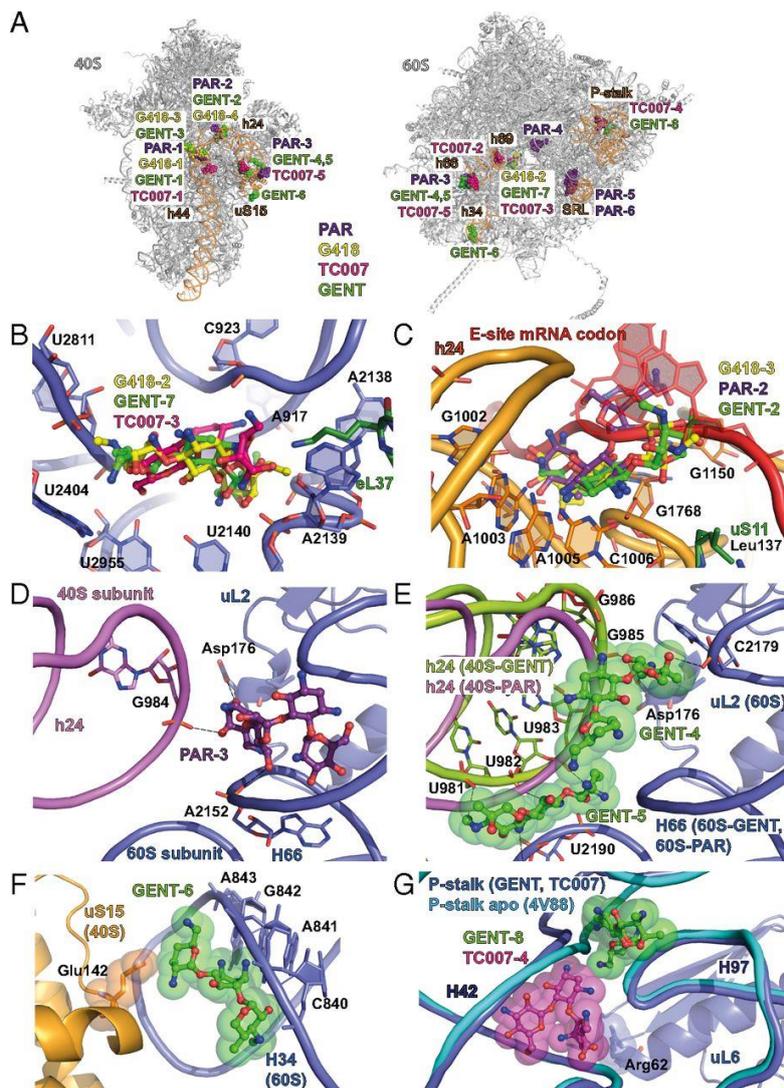


Figure 25 : Les sites de liaison secondaires des aminoglycosides dans le ribosome 80S. (A) Sites de liaison de la gentamicine (GENT), du G418, du TC007 et de la paromomycine (PAR) dans le ribosome 80S de *S. cerevisiae*. Le ribosome est coloré en gris clair ; les éléments des poches de liaison des aminoglycosides sont colorés en orange clair ; la gentamicine est colorée en vert ; le G418 est coloré en jaune ; le TC007 est coloré en magenta ; et la paromomycine est colorée en violet. (B) Liaison de GENT-7, G418-2, et TC007-3 au tunnel de sortie des peptides du ribosome 80S. Le G418 est coloré en jaune, la gentamicine en vert clair, le TC007 en magenta, la grande sous-unité ribosomale en bleu et la protéine eL37 spécifique aux eucaryotes en vert. Des fixations similaires sont observées pour G418-2 et GENT-7, mais l'orientation de TC007-3 est différente. Les atomes d'oxygène sont colorés en rouge, et les atomes d'azote sont colorés en bleu. (C) Sites de liaison de PAR-2, GENT-2, et G418-3 dans le site E de la petite sous-unité ribosomale chevauchant la position de l'ARNm. L'ARNm est coloré en rouge ; les éléments du ribosome 70S sont omis pour plus de clarté. La sous-unité 40S est colorée en orange ; la paromomycine est colorée en violet ; la protéine universelle uS11 est colorée en vert ; et les autres codes de couleur sont ceux de A. (D) Interactions de PAR-3 avec les éléments de l'interface des sous unités ribosomiques. Les contacts établis entre la paromomycine et G984 de la sous-unité 40S, A2152 de la sous-unité 60S, et Asp176 de uL2 sont marqués par des lignes pointillées. La sous-unité 40S est colorée en rose clair ; la paromomycine est colorée en violet ; la sous-unité 60S est colorée en bleu ; et la protéine universelle L2 (uL2) est colorée en bleu. Les atomes d'oxygène sont colorés en rouge, et les atomes d'azote en bleu. (E) GENT-4 et GENT-5 stabilisent des conformations particulières de l'interface des deux sous unités. Rose : structure 80S-paromomycine. Les contacts 80S-gentamicine sont marqués par des lignes pointillées. La sous-unité 40S de la structure 80S-gentamicine est colorée en citron vert ; les autres codes de couleur sont les mêmes que dans A-C. (F) GENT-6 cible le pont B4 formé par la protéine universelle uS15 et H34 de la grande sous-unité. uS15 est représentée en jaune ; les autres codes de couleur sont ceux de D. La glutamine 142, qui interagit avec GENT-6, est représentée par des sphères. (G) Interactions de GENT-8 et TC007-4 avec les éléments de la tige P ribosomale. Les différentes conformations des hélices 42 et 97 de la tige P sont représentées en bleu pour les structures 80S-GENT et 80S-TC007 et en cyan pour la structure 80S-apo (code PDB ID 4V88). H42 dans la conformation apo entretrait en conflit avec GENT-8. Arg62 dans la protéine uL6 s'approchant de TC007 est représentée par des sphères. Le code de couleur est le même que dans A-E. Issu de [259].

## **2. Les modifications de l'anticodon de l'ARNt<sub>Tyr</sub> humain**

Les modifications de la boucle anticodon de l'ARNt sont donc connues pour leur impact sur les règles de décodage [258]. Dans cette nouvelle partie nous allons commencer par l'intérêt thérapeutique de leur étude, avant de revenir sur les deux modifications auxquelles je me suis intéressée durant ma thèse : le GalQ<sub>34</sub> et la  $\Psi_{35}$  de l'anticodon de l'ARNt<sub>Tyr</sub> humain.

### **A. Intérêt thérapeutique des modifications de la boucle anticodon des ARNt dans le cadre des maladies à codon stop prématuré**

Alors que la translecture des codons stop est un événement rare au niveau des codons stop naturels avec un taux estimé à environ  $10^{-4}$  à  $10^{-3}$  chez les mammifères, elle représente une véritable opportunité thérapeutique [251]. En effet, 11% des maladies génétiques sont dues à l'apparition d'un PTC dans la séquence codante [260]. Lorsque la translecture se fait au niveau d'un PTC, la traduction se poursuit jusqu'au codon stop naturel, permettant la formation d'une protéine complète.

Le premier objectif pour restaurer l'expression des protéines est d'induire plus de translecture. En 1985, la première preuve qu'une molécule, la gentamicine, est capable d'induire la translecture était publiée [261]. Depuis, il a été démontré que par exemple dans le cas de la mucoviscidose, la production de 1 à 5 % de protéine CFTR fonctionnelle pourrait suffire à réduire la gravité de la maladie, faisant des inducteurs de translecture un sujet de recherche majeur pour le traitement des maladies dues à des mutations non-sens [262,263]. Après la Gentamicine, d'autres aminoglycosides comme la Généticine (G418), la Paromomycine, la Néomycine, la Lividomycine et le NB124 ont été décrits comme des inducteurs de translecture [240,264,265]. Les aminoglycosides sont des antibiotiques dont la principale caractéristique est d'être des molécules chargées positivement capables de se lier à différents types d'ARN, dont les ARNr. Une étude récente a analysé la fixation de différents aminoglycosides dans le ribosome 80S à une résolution de 3Å afin d'élucider leur mode d'action sur la fidélité de la traduction [259] (Figure 25). La structure obtenue indique que les aminoglycosides interagissent largement avec l'interface des sous-unités du ribosome. Ils ont constaté que les aminoglycosides, composés d'un maximum de trois cycles aromatiques ont la capacité de se lier au tunnel de sortie du peptide. Les aminoglycosides ont également été retrouvés à proximité du site A et du site E, à la base de la tige P du ribosome, et au niveau de la boucle sarcine-ricine. Des molécules non aminoglycosidiques ont également été décrites comme inductrices de translecture comme l'Ataluren (et ses analogues), la Négamycine, les RTCs, l'Amlexanox, la Clitocine et l'Escin [266–274]. Malgré le nombre de candidats potentiels, seuls deux d'entre eux (Ataluren et gentamicine) ont été testés dans des essais cliniques.



Par la suite, leur utilisation dans le cadre des maladies à mutation non-sens a été arrêtée, l'Ataluren ne permettant pas d'obtenir un effet chez les patients testés [275], et la gentamicine étant toxique lorsqu'elle est administrée de manière prolongée [276,277].

Le deuxième objectif est de restaurer l'activité de la protéine. Pour cela l'identité de l'acide aminé incorporé par translecture est primordiale. Chez l'homme, les identités des acides aminés incorporés par translecture sont connues mais contrairement à la levure, les règles d'incorporation des différents ARNt correspondants ne sont pas claires [229,230,278]. Les ARNt supprimeurs naturels des codons UAA et UAG sont les mêmes : Tyrosine (Tyr), Glutamine (Gln), Lysine (Lys). En ce qui concerne, le codon stop UGA ce sont des ARNt supprimeurs différents : Tryptophane (Trp), Cystéine (Cys), Arginine (Arg) et la Leucine (Leu). Contrairement aux autres ARNt supprimeurs cités, l'ARNt<sub>Leu</sub> n'avait pas été retrouvé comme étant un ARNt supprimeur naturel chez la levure. Une meilleure compréhension des mécanismes de reconnaissance des ARNt supprimeurs naturels chez l'Homme permettrait de potentialiser les traitements des patients atteints de maladies à codon stop prématurés, en favorisant l'incorporation d'un acide aminé compatible avec l'activité de la protéine.

L'épitranscriptome apparaît ainsi comme une voie potentielle d'action dans le traitement des maladies à codon stop prématuré. En effet, comme je l'ai présenté en début de chapitre, les modifications d'ARN sont capables de moduler la translecture d'un point de vue quantitatif mais également qualitatif. On retrouve notamment l'exemple récent de la molécule 2,6-diaminopurine (DAP) [279]. En effet, en 2020 une équipe a démontré que la DAP, présente dans l'extrait du champignon *Lepista inversa*, était un inducteur de la translecture spécifique du codon UGA. L'augmentation de la translecture est induite par une inhibition de la protéine ftsJ homologue 1, qui catalyse la méthylation du C<sub>34</sub> de l'ARNt<sub>Trp</sub>. Ainsi, en présence de DAP, l'ARNt<sub>Trp</sub> hypométhylé est incorporé d'avantage par translecture au niveau du codon UGA. L'étude de la DAP a également permis de démontrer son efficacité en tant qu'inducteur de translecture en cellules humaines et en modèle murin, avec une faible toxicité. Ces résultats seront ainsi suivis d'essais cliniques plus avancés de la DAP qui pourrait devenir la première molécule affectant l'épitranscriptome utilisée dans la thérapie des maladies à PTC.

Les modifications GalQ<sub>34</sub> et  $\Psi_{35}$  sont les deux modifications présentes au niveau de l'anticodon de l'ARNt<sub>Tyr</sub>, qui est l'ARNt majoritaire supprimeur des codons stop UAA et UAG [229,278]. Tout comme le C<sub>34</sub> de l'ARNt<sub>Trp</sub>, ces deux modifications pourraient apporter de nouvelles pistes thérapeutiques.

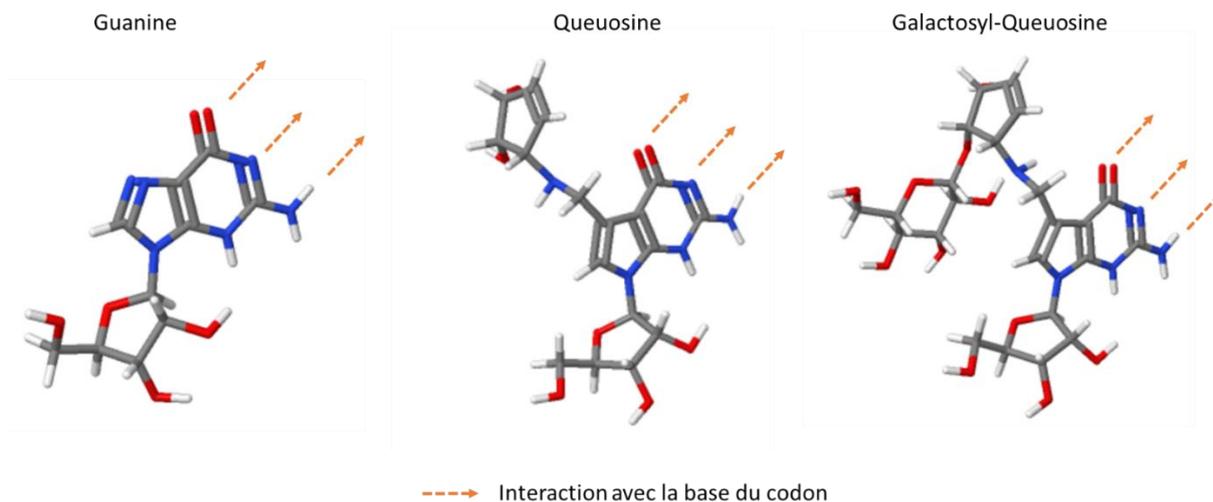


Figure 26 : Représentation atomique des bases retrouvées en position 34 de l'anticodon de l'ARNt Tyrosine. De gauche à droite : la guanine, la queuosine et le galactosyl-queuosine. Flèches : groupes atomiques impliqués dans l'interaction avec la troisième base du codon de l'ARNm.

## B. Galactosyl-queuosine-34

La queuosine (Q), est un analogue modifié de la guanosine (Figure 26). Elle est incorporée par l'ARN guanine transglycosylase (TGT) à partir d'une base queuine [280–282], à la place du G<sub>34</sub> des ARNt à anticodon GUN, soit l'Asparagine (Asn), Aspartate (Asp), Histidine (His), et Tyrosine (Tyr) [283]. Il s'agit d'un nucléotide conservé dans l'ensemble du vivant, mise à part la levure *S. cerevisiae* qui n'en incorpore pas dans ses ARN [284]. Chez les mammifères, la Q<sub>34</sub> cytosolique est hypermodifiée pour les ARNt Tyr et Asn par l'ajout d'un sucre, un galactose (GalQ) et un mannose (ManQ) respectivement, au niveau du groupe hydroxy en C4 du cyclo-pentenediol [285–287] (Fig. 27). La mannosyltransférase responsable de la catalyse de ManQ a été isolée chez le rat, mais aucune galactosyltransférase n'a encore été identifiée pour GalQ [288]. Une étude réalisée en cellules humaines a retrouvé trois formes d'ARNt<sub>Tyr</sub> cytosolique, G<sub>34</sub>, Q<sub>34</sub> et GalQ<sub>34</sub>, alors que l'ARNt<sub>Asp</sub> a toujours été trouvé sous formes G<sub>34</sub> ou ManQ<sub>34</sub> [289]. Cela suggère que l'ARNt<sub>Asp</sub> est modifié par ajout d'un ManQ, alors que l'ARNt<sub>Tyr</sub> serait d'abord modifié par ajout d'un Q puis d'un galactose, une autre hypothèse pouvant être la compétition entre l'ajout d'un Q ou d'un GalQ au niveau de l'ARNt<sub>Tyr</sub>. Alors que les procaryotes sont capables de synthétiser Q *de novo*, les eucaryotes doivent l'obtenir, et notamment la base queuine, en tant que micronutriment à partir du microbiote intestinal ou par l'alimentation [287,290,291].

GalQ<sub>34</sub> (Figure 26) est donc spécifique de l'ARNt<sub>Tyr</sub>, l'ARNt suppresseur majoritaire incorporé au niveau des codons UAA et UAG. Il s'agit d'une modification chimique peut étudiée malgré le fait qu'elle soit de taille imposante et présente au niveau de la position wobble (Fig. 28). En revanche, le rôle de la modification Q a été plus étudié, ce qui a permis d'obtenir des données préliminaires sur son impact sur l'incorporation des ARNt.

En 1978, les premières règles de décodage d'un ARNt portant un Q étaient publiées, avec la mesure physique que l'appariement non canonique Q-U est deux fois plus stable que celui du G-U [292]. Ensuite, la première étude *in vivo*, s'est concentrée sur les règles de décodage de l'ARNt<sub>His</sub> en fonction de la présence de Q [293]. Des ARNt<sub>His</sub> avec l'anticodon G/QUG issus de *D. melanogaster* ont été injectés dans des ovocytes de Xénope et évalués pour leur capacité de décodage des codons cognat CAC et proche cognat CAU. Il a été constaté que l'ARNt<sub>His</sub> G<sub>34</sub> décode plus favorablement le codon CAC, alors que le Q<sub>34</sub> présente une légère préférence pour le codon CAU. La présence de Q<sub>34</sub> permettrait donc de faciliter l'incorporation de l'ARNt proche cognat. Une autre étude plus récente a confirmé ces résultats [294]. Partant du principe que le taux de modification Q varie au cours du développement de *D. melanogaster* [295], ils l'ont comparé à celui de l'expression de gènes présentant des biais de codons [294].

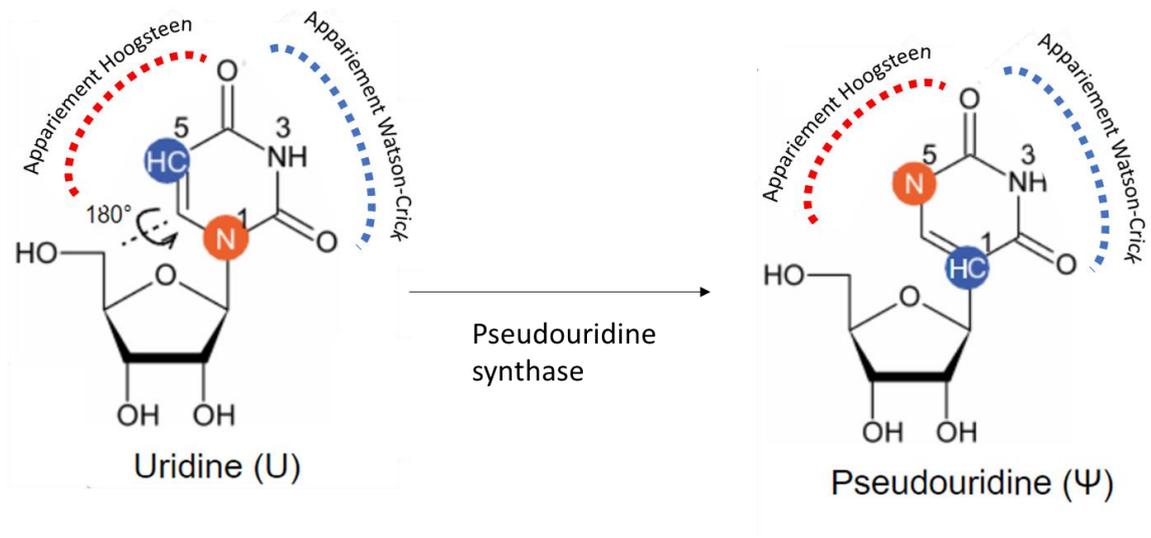


Figure 27 : Structure du nucléoside uridine (U) et de son isomère, la pseudouridine (Ψ). La rotation de la base provoque le remplacement de la liaison glycosidique C1-N1 par une liaison C1-C5. La réaction est catalysée par une pseudouridine synthase.

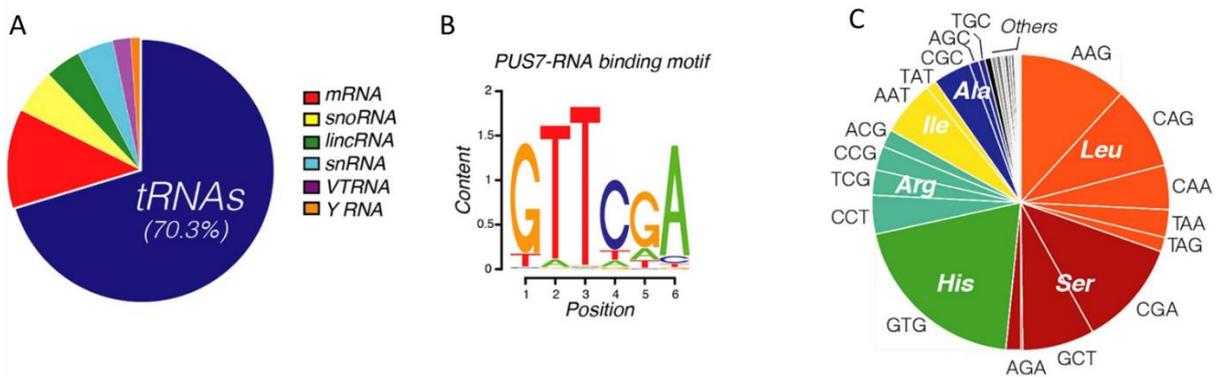


Figure 28 : Les ARN cibles de la protéine Pus7 chez l'Homme. (A) Le graphique montre les catégories distinctes d'espèces d'ARN liées à Pus7 dans les cellules humaines identifiées par iCLIP-seq. (B) L'analyse de la séquence des produits iCLIP spécifiques à PUS7 identifie un nouveau motif putatif de liaison ARN-Pus7. (C) Le graphique montre l'enrichissement des ARNt individuels liés à Pus7 identifiés par iCLIP-seq, normalisé à leur abondance dans les cellules déterminée par smRNA-seq. Issu de [296].

Leur analyse a permis de mettre en évidence que le taux de modification Q influe directement le niveau d'expression de ces gènes à cause de l'augmentation de l'incorporation des ARNt proches cognats. A ce jour, l'impact sur les mécanismes de décodage des hyper-modifications ManQ et GalQ n'a toujours pas été étudié. Il est pourtant connu que la présence de queuine a un effet sur la production de l'acide aminé Tyr. En effet, des souris privées de queuine survivent en conditions classiques, jusqu'à ce qu'elles soient également privées de l'apport alimentaire de Tyrosine [297]. Ces souris peuvent être sauvées par l'apport de queuine dans leur alimentation. Il a été démontré par la suite que cela découle de l'oxydation accrue de la tétrahydrobioptérine en absence de queuine, qui compromet alors la production de Tyrosine par la cellule [298]. Ce phénotype peut être annulé par l'apport de Tyrosine dans le milieu cellulaire.

### C. Pseudouridine-35 de l'ARNt<sub>Tyr</sub>

La pseudouridine ( $\Psi$ ), isoforme de U ou « cinquième nucléoside », est la première des modifications chimiques de l'ARN à avoir été découverte [255], la plus abondante du répertoire épitranscriptomique, et présente dans les trois domaines du vivant [257,299]. Elle est le résultat de l'isomérisation de l'uridine : la liaison glycosidique C1-N1 qui relie la base au ribose est remplacée par une liaison C1-C5 [6,7] (Figure 27). Cette isomérisation a pour conséquence de libérer N1, ce qui augmente les possibilités d'appariements Hoogsteen et facilite alors le repliement de l'ARN, la formation de structures secondaires, la formation de liaisons ARN-ARN ou ARN-protéine [300–302]. Les pseudouridines peuvent être catalysées par des pseudouridines synthases dépendantes d'ARN guides appelés snoRNAs à boîte H/ACA dans le cas des ARNr, ou bien par des pseudouridines synthases indépendantes, dans le cas de la majorité des sites  $\Psi$ s. Ces dernières ciblent leurs sites de manière séquence et/ou structure spécifique [303].

Tout comme GalQ<sub>34</sub>, la  $\Psi_{35}$  est présente au niveau de l'anticodon de l'ARNt<sub>Tyr</sub> cytoplasmique. Chez les eucaryotes, elle est connue pour être catalysée par l'enzyme Pus7 [304]. Il a été démontré chez l'Homme que Pus7 était responsable de la présence de  $\Psi$  dans différentes populations d'ARNt, mais également dans d'autres populations d'ARN comme les ARNm, snoRNA, lincRNA et snRNA [296] (Figure 28). Cette étude a permis d'attribuer notamment un nouveau rôle à Pus7, qui active par pseudourylation un nouveau réseau de petits fragments dérivés de l'ARNt (tRF), ciblant le complexe d'initiation de la traduction. Ainsi, l'inactivation de Pus7 dans les cellules souches embryonnaires affecte la régulation de la traduction médiée par les tRF, ce qui entraîne une augmentation de la synthèse des protéines.

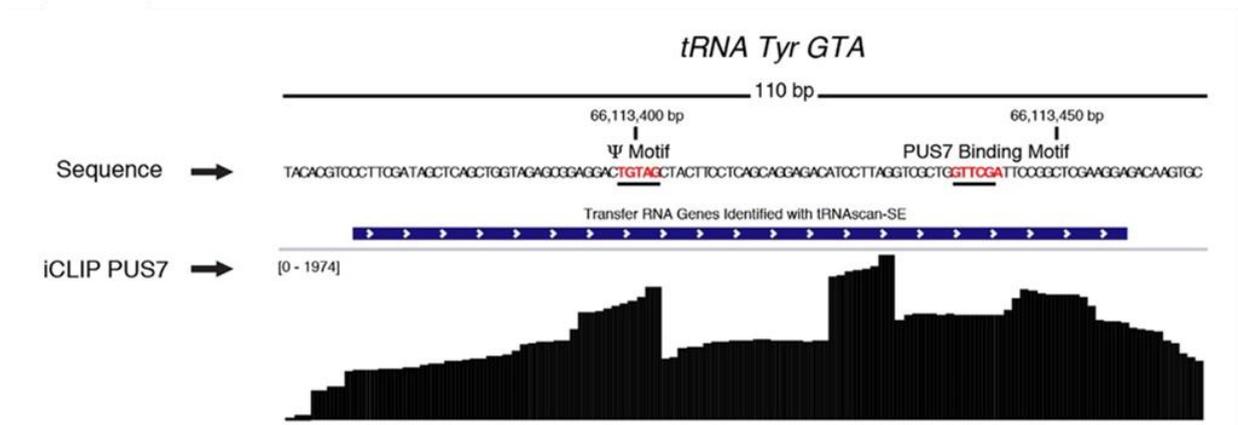


Figure 29 : iCLIP-seq de l'ARN-Tyr (GTA) lié à Pus7 dans les cellules humaines. Une séquence génique indique la séquence génomique linéaire de l'ARN-Tyr-GTA-5-1 (axe des x, milieu) ainsi que le nombre de lectures iCLIP spécifiques de Pus7 par million (axe des y, bas). En haut, la séquence génomique met en évidence la position des motifs Pus7 Ψ-consensus (UGUAG) et de liaison (GTTCTCGA). Issu de [296].

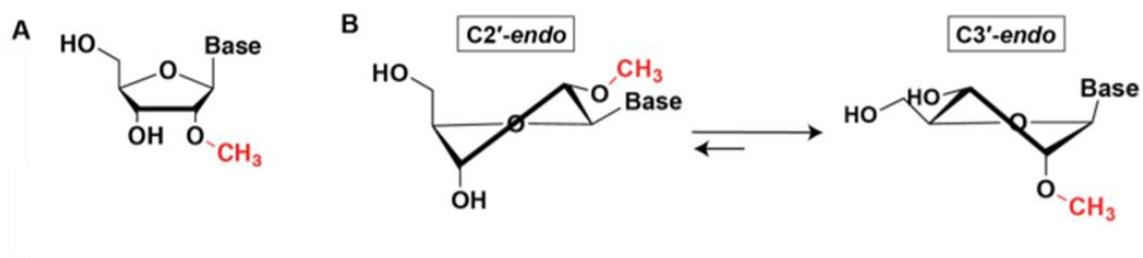


Figure 30 : Structure chimique d'une base portant une modification 2'-O-méthylation. A) Représentation plane d'un nucléoside portant une 2'-O-me. B) Induction d'une conformation C3'-endo du sucre du nucléoside par la présence de la 2'-O-me. Issu de [305].

La modification  $\Psi_{35}$ , étant présente également chez la levure, elle a déjà été étudiée au laboratoire pour son rôle dans l'incorporation de l'ARNt<sub>Tyr</sub> par translecture [306]. Des analyses par spectrométrie de masse ont permis de mettre en évidence que dans des conditions sauvages chez *S. cerevisiae*, l'ARNt<sub>Tyr</sub> est l'ARNt suppresseur majoritaire des codons stop UAA et UAG, alors que la délétion du gène PUS7 provoque la chute de l'incorporation de l'ARNt<sub>Tyr</sub> au niveau des deux codons stop. En revanche, un ribosome profiling n'a pas permis de retrouver des modulations de l'incorporation de l'ARNt<sub>Tyr</sub> au niveau des codons sens tyrosine. Chez la levure  $\Psi_{35}$  ne semble donc pas jouer de rôle dans l'incorporation proche cognat de l'ARNt<sub>Tyr</sub> au codon UAU alors qu'elle est importante pour l'incorporation au niveau des codons UAA et UAG.  $\Psi_{35}$  a également été étudié dans le modèle procaryote, pour l'ARNt<sub>Tyr</sub> d'*E. coli* qui est très proche de l'humain avec un anticodon Q<sub>34</sub> $\Psi_{35}$ A<sub>36</sub> et GalQ<sub>34</sub> $\Psi_{35}$ A<sub>36</sub> respectivement [307]. Une délétion de la pseudouridine synthase responsable, rluF, a été réalisée et la capacité d'incorporation de l'ARNt<sub>Tyr</sub> a été évalué au niveau des codons sens Tyrosine. Les résultats ont permis de mettre en évidence qu'en absence de rluF, une séquence de 10 codons Tyrosine était suffisante pour diminuer la traduction par rapport à la souche WT. Il est connu que les modifications d'ARNt agissent comme un véritable réseau, avec des connections possibles entre elles. Une étude de la protéine Pus7 réalisée en cellules humaines a déjà mis en avant par iCLIP-seq que la protéine Pus7 se lie également à l'ARNt<sub>Tyr</sub> humain, et que la délétion du gène PUS7 ne provoque pas de différences significatives dans la fréquence des différents isoaccepteurs d'ARNt<sub>Tyr</sub> matures [296] (Figure 29). Une autre étude, toujours en cellules humaines, a démontré que l'absence d'expression de la protéine Pus7 engendrait l'absence de  $\Psi_{35}$  de l'ARNt<sub>Tyr</sub> [308]. Il reste en revanche encore à déterminer si, chez l'Homme avec la présence de GalQ<sub>34</sub>,  $\Psi_{35}$  conserve son rôle dans l'incorporation de Tyrosine par translecture ou au niveau des codons sens.

### **3. 2'-O-méthylation de l'ARNr humain**

La 2'-O-méthylation (2'-O-me) a été décrite par de nombreuses études chez les deux ARN majoritaires de la cellule, l'ARNt et l'ARNr, avant d'être aussi retrouvée dans l'ARNm [309]. Alors que nous avons vu que son rôle dans la fidélité de l'incorporation des acides aminés a été récemment établi [258], nous allons revenir en détails sur cette modification dans le cadre de l'ARNr, et de l'intérêt thérapeutique de son étude dans le cadre de la cancérologie.

#### **A. 2'-O-méthylation de l'ARNr : structure, fonctions et formation**

La 2'-O-méthylation (2'-O-me) consiste en l'ajout d'un groupe méthyle sur le groupement hydroxyle du second carbone du ribose du nucléotide, à partir du métabolite SAM (S-Adénosyl-Méthionine) (Figure 30). Elle est retrouvée sur les 3 ARNr dérivant du même précurseur, l'ARNr 47S, (i.e., 5.8S, 18S et 28S) [310].

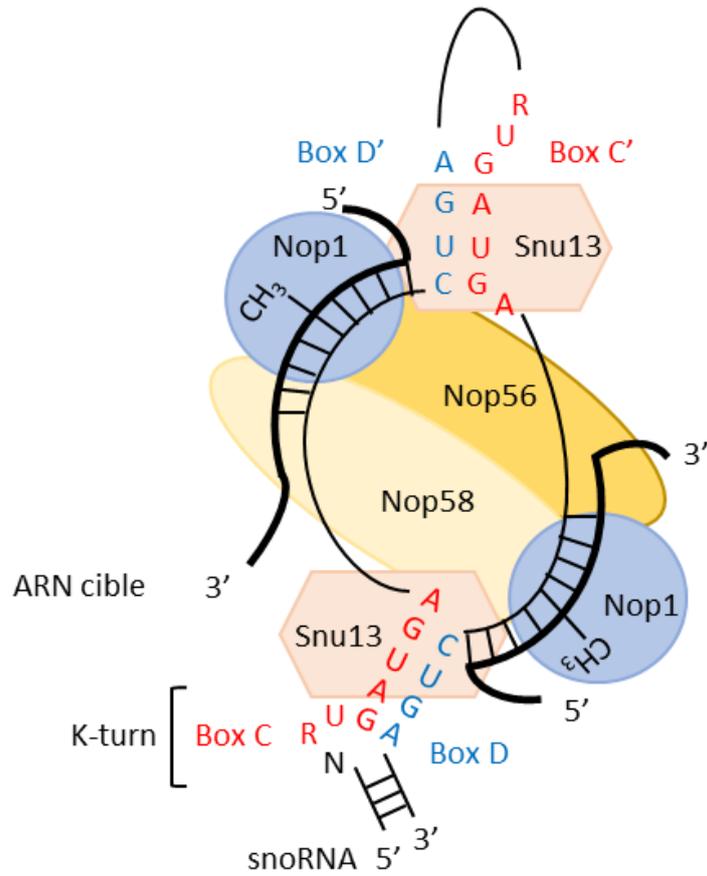


Figure 31 : Structure du complexe snoRNP de la boîte C/D. Deux régions d'ARNr cibles sont représentées en gras et sont liées à des régions complémentaires du snoRNA guide à boîte C/D. Les nucléosides méthylés se trouvent à 5 nucléotides en amont d'une boîte D ou D' et sont indiqués (-CH<sub>3</sub>). Les boîtes C et C' sont surlignées en rouge et les boîtes D et D' sont surlignées en bleu. Les protéines associées sont indiquées en couleur. A partir de [311].

La présence de 2'-O-me crée une répulsion stérique intra-résidu entre le 3'-phosphate et les groupes 2-carbonyle quand ils sont dans la conformation C2'-endo, favorisant ainsi la conformation C3' (Figure 30) [305]. Cette modification chimique peut affecter les ARN de multiples façons [312], en augmentant leur hydrophobie, en les protégeant des attaques des nucléases [313,314], en stabilisant les structures hélicoïdales, et en affectant leurs interactions avec les protéines ou d'autres ARN [315–317]. D'un point de vue du ribosome, on retiendra notamment que la 2'-O-me augmente la stabilité thermodynamique des paires de bases ARN:ARN [318–321], peut perturber les structures tertiaires de l'ARN [322], et inhiber les interactions ARN-protéine par des effets stériques [323,324] ou par un impact sur la liaison hydrogène [325,326].

La 2'-O-me est catalysée par un complexe de méthylation appelé snoRNP, contenant quatre protéines différentes et un petit ARN nucléolaire non codant (snoRNA) (Figure 31) [310,327,328]. Bien qu'elle soit indépendante de la nature du nucléotide, l'ajout de 2'-O-me se produit à des sites spécifiques le long de l'ARNr. Un site de méthylation est spécifié par le snoRNA correspondant, qui sert de guide pour l'ensemble du complexe snoRNP. Ces snoRNA, appelés SNORD, présentent des motifs appelés boîtes C/D [327,329], qui interagissent directement avec l'ARNr par homologie de séquence. Les interactions entre les SNORD et l'ARNr dépendent de la structure particulière du complexe de méthylation fourni par NOP58, NOP56 et NHP2L1 chez l'Homme. La méthyltransférase du complexe, la Fibrillarin (FBL), catalyse la 2'-O-me du cinquième nucléotide en aval de la boîte D. Il existe des méthylation d'ARNr indépendantes de ce système, réalisées par des méthyltransférases individuelles spécifiques du site à méthyler [309]. Ces méthyltransférases site-spécifiques, ainsi que leurs sites 2'-O-me associés, sont conservés chez les bactéries, où quatre sites uniques de 2'-O-me ont été décrits [330]. On retrouve notamment 16S-Cm1402, 23S-Gm2251, 23S-Cm2498, et 23S-Um2552, catalysées respectivement par RsmI (YraL), RlmB (YfjH), RlmM (YgdE), et RlmE (Ftsj) [309].

Les niveaux de modification de 2'-O-me à des sites distincts peuvent changer de façon dynamique en réponse à des voies de signalisation et représentent une source considérable d'hétérogénéité du ribosome dans les cellules humaines [192]. Il a été démontré que la suppression d'un seul site de méthylation important pouvait entraîner une modulation du niveau d'expression traductionnel de certains ARNm, avec ainsi l'émergence de ribosomes ayant une fonction spécialisée.

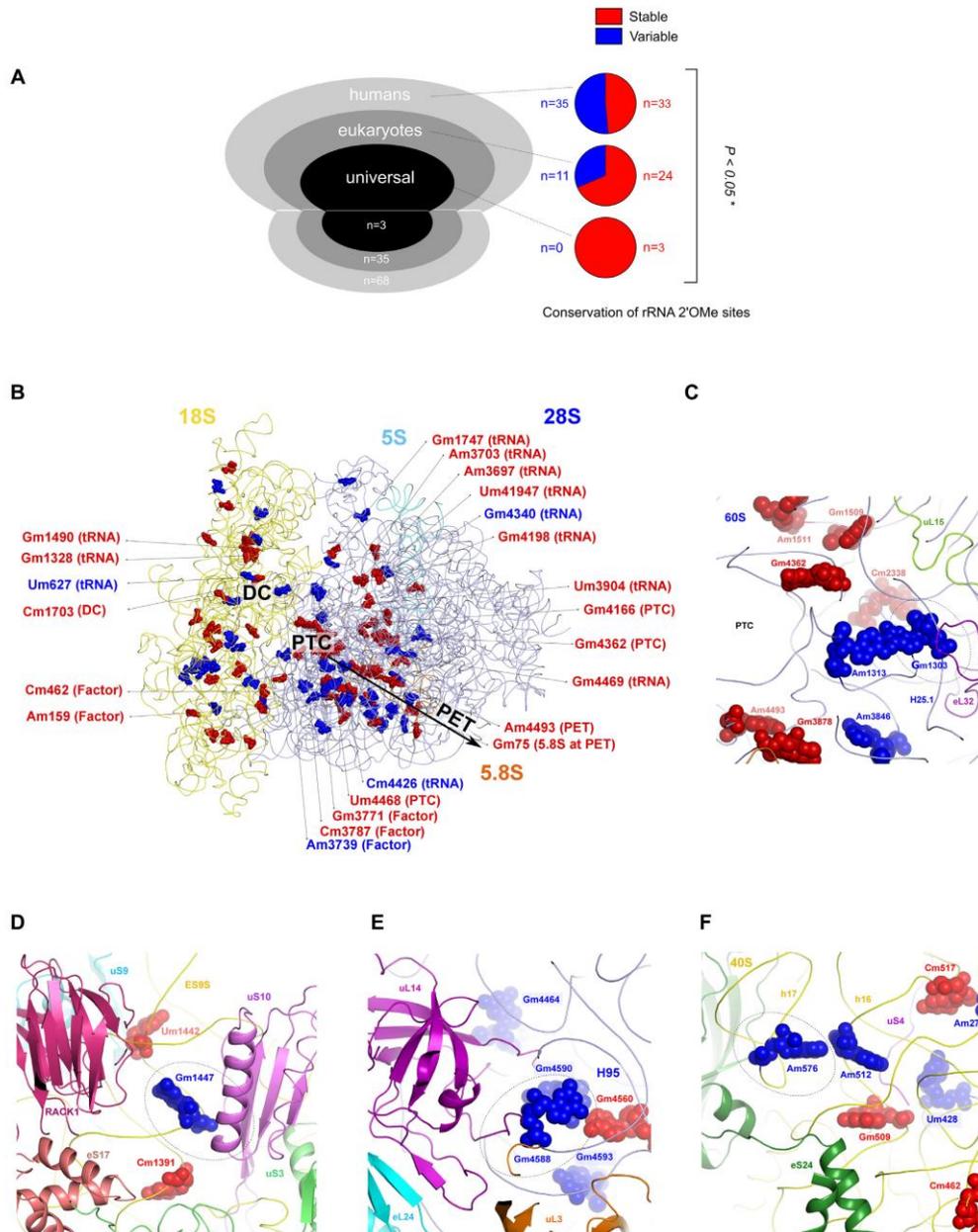


Figure 32 : Propriétés structurales des sites de 2'-O-méthylation stables et variables de l'ARNr. (A) Parmi les 106 sites 2'-O-méthylés de l'ARNr chez les humains, 68 sont spécifiques aux humains (humains), 35 sont présents chez les eucaryotes (eucaryotes) et 3 sont conservés des bactéries aux humains (universels) (panneau de gauche). La proportion de sites stables et variables parmi ces trois groupes est représentée par des diagrammes circulaires (panneau de droite). (B et C) Cartographie des deux classes de sites 2'-O-me d'ARNr sur le ribosome humain (B) et au niveau du centre peptidyl transférase (C). Seules les positions d'ARNr 2'-O-me directement impliquées dans les domaines de fonction ou dans l'interaction avec les composants de la machinerie traductionnelle ont été annotées. (D-F) Environnement structural des sites variables de 2'-O-me de l'ARNr significativement associés aux caractéristiques du cancer du sein. Rouge : sites stables ; bleu : sites variables ; cercles pointillés : sites d'intérêt ; DC : centre de décodage ; PTC : centre peptidyl-transférase ; PET : tunnel de sortie du peptide. Issu de [331].

La protéine FBL est une protéine essentielle, son knock-out ayant été associé à la létalité dans des modèles de levure, poisson zèbre et souris [332,333]. En effet, il a été démontré qu'elle est requise dans le développement précoce [334], et qu'elle est d'ailleurs fortement exprimée dans les cellules souches pluripotentes embryonnaires dans lesquelles elle module la différenciation [335]. En plus de son rôle de 2'-O-me transférase, FBL est connue pour participer au clivage des pré-ARNr [336], et à la régulation de l'activité de l'ARN Pol I sur les promoteurs des gènes ribosomiques en méthylant un résidu glutamine de l'histone H2A [337]. De récentes études montrent l'importance de la 2'-O-me transférase dans le bon fonctionnement cellulaire avec la mise en évidence d'une augmentation du niveau de FBL dans différents types de cancers [338,339], la surexpression étant accompagnée d'une diminution du taux de survie chez les patients concernés.

## **B. Variabilité dans le cadre de cancers**

Depuis quelques années, le taux de méthylation des ARNr est devenu un sujet d'études dans le domaine de la cancérologie. Les premières évidences de variation du taux de 2'-O-me ribosomique ont été trouvées par RTL-PCR (Reverse Transcription Low dNTPs-PCR) dans des cas de patients atteints de leucémie myéloïde aiguë (LMA), en fonction de l'oncogène associé [340]. Une augmentation de 2'-O-me a été retrouvée au niveau de deux sites de l'ARNr 18S chez 3 des 5 patients analysés atteints de LMA associée à ETO1, tandis qu'une baisse de 2'-O-me a été retrouvée au niveau de 4 sites de l'ARNr 28S chez 14 dans 16 patients analysés atteints de LMA associée à NPM1c<sup>+</sup>. Avec l'avancée de la technologie et notamment l'apparition du RiboMeth-Seq, de nouvelles études de plus grandes envergures ont permis de confirmer ces premières observations. Une étude a établi le profil de 2'-O-me pour 110 sites de l'ARNr de 17 échantillons de patients atteints de lymphome diffus à grandes cellules-B et de trois donneurs sains [341]. Alors que certains échantillons de patients malades présentaient une augmentation significative du taux de méthylation pour 12 sites, la majorité des différences correspondait à une diminution de la 2'-O-me pour 65 sites. Une seconde étude la même année a été réalisée sur des tumeurs du cancer du sein chez plus de 200 patientes (Figure 32) [331]. Elle a également permis de mettre en évidence des altérations spécifiques du taux de 2'-O-me, avec notamment une diminution significative pour le site Am576 de l'ARNr 18S dans une série de 10 tumeurs mammaires malignes par rapport à 8 tumeurs bénignes, première mise en lumière d'une variation de la méthylation en fonction du degré pathologique de l'échantillon. Cette étude souligne également que même si les sites de méthylation varient naturellement d'un patient à un autre, 40% des sites présentent une forte variation alors que 60% varient très faiblement.



De manière intéressante, les sites de 2'-O-me "stables" correspondent à des sites conservés tout au long de l'évolution et sont concentrés directement dans les principaux domaines fonctionnels du ribosome alors que les sites "variables" apparaissent au cours de l'évolution et se localisent à la fois dans la deuxième couche des nucléotides formant les domaines fonctionnels principaux du ribosome et à des positions d'interaction directe avec les protéines ribosomiques ou les facteurs de traduction centraux. Différentes hypothèses permettent d'expliquer ces variations du taux de 2'-O-me de l'ARNr. Premièrement, la diminution ou l'augmentation de l'expression de SNORD induirait une variation du taux de méthylation à certains sites spécifiques [340–343]. Deuxièmement, la conservation de certains sites spécifiques pourrait être la conséquence du caractère indispensable de ces sites de méthylation pour l'intégrité du ribosome, ce qui provoquerait la perte des ribosomes hypométhylés au niveau de ces sites. Une troisième hypothèse plus indirecte, est basée sur l'hyperprolifération des cellules cancéreuses : la synthèse de l'ARNr est hyperactivée dans les cellules cancéreuses [344,345], alors que l'expression des facteurs impliqués dans le dépôt de 2'-O-me est limitée, provoquant une diminution globale du taux de méthylation [341,346]. Une dernière hypothèse repose sur le taux d'expression de la méthyltransférase FBL.

Le lien entre FBL et le cancer a été renforcé par l'étude de Marcel et al. en 2013, qui met en évidence que la surexpression de FBL favorise la prolifération cellulaire et la résistance à la chimiothérapie des cellules de cancer du sein MCF-7 [338]. Ils ont notamment retrouvé que p53 réprime l'expression de FBL en se liant directement au gène FBL. Dans le cas du cancer du sein, la suppression de l'expression de p53 dans les cellules épithéliales mammaires s'accompagne alors d'une augmentation du taux de 2'-O-me de l'ARNr. Ce phénotype s'accompagne d'une altération de la fidélité de la traduction et d'une augmentation de la traduction dépendante de l'IRES de gènes clés liés au cancer, tels que IGF-1R, C-MYC, VEGF-A et FGF1/2. Il a été également mis en avant que l'oncogène C-MYC favorise l'expression de FBL en se liant dans son 5'UTR [347–349]. Afin de mieux comprendre les altérations que subit le ribosome en fonction de l'expression de FBL, l'équipe de JJ. Diaz en 2017 publiait la première analyse globale comparative du profil de 2'-O-me dans des cellules humaines sous-exprimant la protéine FBL [350]. Ce travail a mis en évidence qu'une diminution de l'expression de FBL d'un facteur de 5 à 10, inhibe la maturation de l'ARNr 18S mais pas celui des ARNr 5.8S et 28S, et n'a pas d'impact sur la stœchiométrie finale des protéines ribosomiques. Dans ces conditions la biogenèse des ribosomes est suffisante pour maintenir la production de ribosomes à ~80% par rapport à la lignée témoin. Tout comme les analyses réalisées sur les tumeurs de patients, la diminution de la méthylation était très variable selon les sites, allant de 0,2 à 57%.



Les sites hypométhylés sont présents dans toute la structure du ribosome, y compris dans la région la plus conservée par rapport aux ribosomes bactériens (ce qui est contraire aux résultats issus des tissus tumoraux [331]), et au niveau de sites clés pour la traduction proches des sites A et P, des ponts inter-sous unités et du tunnel de sortie des peptides. En revanche, les sites les moins affectés par la sous expression de FBL sont présents au niveau du PTC et du centre de décodage de l'ARNr 18S.

L'hyper ou hypo-méthylation de l'ARNr semble être un facteur décisif dans le caractère tumoral des cellules humaines, induisant une modulation de la fidélité de la traduction. Alors que la composition en ARN et protéines du ribosome ne semble pas être affectée par le taux de méthylation de l'ARNr [350], la récente structure du ribosome humain à 2.5 Å implique la 2'-O-me dans la stabilisation locale et globale de la structure de l'ARNr [351]. La manière dont le niveau de méthylation du ribosome module la fidélité de la traduction reste encore toutefois à être élucidée.

## IV. Projets et objectifs de thèse

Mes travaux de thèse ont porté sur deux projets centrés sur l'impact des modifications de l'ARN dans la fidélité de la traduction eucaryote. L'objectif est d'apporter une meilleure compréhension des mécanismes moléculaires impliqués dans les événements de recodage et de contribuer au développement et à l'amélioration des stratégies thérapeutiques dans le traitement des maladies à codon stop ou des cancers.

Mon premier projet consiste à étudier le rôle des modifications de la boucle anticodon des ARNt, lorsqu'ils sont utilisés en tant qu'ARNt proche cognat. Le modèle mécanistique le plus adapté pour cette étude est la translecture, qui ne met pas en compétition cette incorporation avec celle d'un ARNt cognat. Alors qu'une étude a déjà été réalisée par mon laboratoire chez la levure [306], la transposition au modèle humain a été compliquée par le manque d'informations sur les modifications concernées, le caractère essentiel de protéines impliquées dans ces modifications et les limites d'approches génétiques réalisables dans ce modèle. C'est pour ces raisons que le projet s'est focalisé sur les deux modifications de l'anticodon de l'ARNt<sub>Tyr</sub>, l'ARNt supprimeur naturel le plus incorporé au niveau de UAA et UAG [229]. Ces deux modifications sont GalQ<sub>34</sub>, une modification absente de la levure et dépendante de la présence de queuine dans le milieu cellulaire, et  $\Psi_{35}$ , une modification également présente chez la levure et publiée chez les eucaryotes pour être produite par l'enzyme Pus7 qui est non essentielle. Afin d'étudier le rôle de ces deux modifications dans les règles d'incorporation de l'ARNt<sub>Tyr</sub> par translecture, j'ai mis au point les conditions de culture en concentration de queuine adaptées à mon étude, et j'ai réalisé un knock-out du gène PUS7 par le système CRISPR/Cas9. J'ai également réalisé un knock-out du gène PUS7L, dont la protéine Pus7L est publiée comme orthologue de Pus7 chez l'Homme, et pour qui aucune caractérisation fonctionnelle n'avait été réalisée. J'ai mis en évidence la présence/absence des modifications et des protéines concernées dans mes lignées par différentes méthodes avant de réaliser des études plus globales sur leur impact traductionnel global. Afin de mener à bien mon objectif j'ai réalisé une caractérisation quantitative et qualitative de la translecture en présence ou absence des modifications, par des mesures du taux de translecture ainsi que du niveau d'incorporation des différents acides aminés incorporés au niveau des codons stop UAA, UAG et UGA par spectrométrie de masse. Ces résultats permettent d'apporter une meilleure compréhension des mécanismes d'appariements codon-anticodon chez l'Homme, et représentent également une nouvelle avancée dans le traitement des patients atteints de maladies génétiques à codon stop prématuré.

Le deuxième projet a été réalisé sous la co-direction de Karen Perronet (Optique et microtechnologies appliquées au vivant, Laboratoire Lumière, Matière et Interfaces). Ce projet concerne la mise au point d'un système de traduction *in vitro* permettant de suivre la traduction eucaryote en molécule unique. Comme je l'ai présenté précédemment, le taux de 2'-O-me affecte la traduction des ARN à initiation dépendante d'IRES sans moduler la composition en protéines/ARN du ribosome. Mon objectif a donc été de déterminer si la capacité de reconnaissance des IRES en fonction du taux de méthylation pouvait s'expliquer via une modulation de la cinétique du ribosome. Pour cela j'ai eu pour but d'adapter le système d'étude de la cinétique du ribosome *in vitro* mis au point par nos deux laboratoires [352], à l'analyse individuelle de ribosomes humains hypométhylés, afin d'obtenir des données de cinétique spécifiques en fonction du taux de méthylation du ribosome. Dans cet objectif, j'ai tenté de purifier des ribosomes fonctionnels issus de cellules Hela dans lesquelles FBL est sous-exprimée. En parallèle, j'ai analysé la cinétique de traduction d'ARNm dont l'initiation est dépendante de l'IRES EMCV, IRES qui n'avait pas été étudié avec notre système jusque-là. Ainsi, j'ai analysé la vitesse de traduction de ribosomes de lapin, et les premières étapes d'élongation de la traduction spécifique d'une initiation dépendante de l'IRES EMCV, que j'ai pu comparer aux données issues de l'analyse cinétique des initiation dépendantes des IRES HCV et CrPV obtenues avec le même système. Bien que je n'aie pu obtenir que des données préliminaires, il représente un véritable challenge dans la compréhension de l'impact des 2'-O-me de l'ARNr sur l'activité ribosomale. Il pourrait ainsi contribuer à mieux comprendre les mécanismes impliqués dans la dérégulation globale observée dans les cellules cancéreuses.

## V. Résultats

Le premier axe de mon projet de thèse porte sur l'impact des modifications de la boucle anticodon des ARNt humains sur la fidélité de la traduction. Dans ce chapitre, je débiterai par la présentation d'un article (non soumis) sur les résultats obtenus au cours de mon doctorat. Ensuite je reprendrai l'ensemble des données supplémentaires obtenues mais ne rentrant pas dans le cadre de l'article en cours d'écriture.

### 1. Les modifications de la boucle anticodon des ARNt

#### Résumé de l'article "Stop codon readthrough in human: natural suppressors and impact of tRNA anticodon loop modifications"

La fidélité de la traduction est assurée par la reconnaissance des ARNt par des aminoacyl-ARNt synthétases spécifiques et par l'appariement spécifique des ARNt avec les codons correspondants. Elle est essentielle afin d'éviter l'incorporation erronée d'acides aminés et une terminaison prématurée. Cependant, 15 codons n'ont pas d'anticodon complémentaire dans le génome humain [42]. On parle alors de dégénérescence du code génétique, avec la tolérance d'un mésappariement entre la base 34 de l'ARNt et la position 3 du codon, conduisant à l'incorporation d'un acide aminé par un ARNt proche cognat. Il est acquis à présent que la discrimination par le ribosome entre les ARNt cognats, proches cognats et non-cognats n'est pas conférée uniquement par la stabilité du complexe codon-anticodon, et que le ribosome joue lui-même un rôle actif dans l'incorporation des ARNt proches cognats. En effet, les appariements U-G en première ou deuxième position du codon sont tolérés car ils interagissent avec les bases A1492, A1493 et G530 de l'hélice 44 du ribosome comme des appariements canoniques Watson-Crick. De plus, l'hélice 44 ne vérifierait avec précision que l'appariement des bases 1 et 2 du codon lors de l'accommodation de l'ARNt, tolérant un mauvais appariement à la troisième position du codon tant qu'il n'y a pas d'encombrement stérique. Plus de 100 modifications d'ARN sont connues chez l'Homme, avec un taux moyen de 13 modifications par ARNt. Les modifications des ARNt peuvent avoir deux fonctions principales : (1) la stabilisation de l'ARNt; (2) le décodage de l'ARNm. Les modifications connues pour jouer un rôle actif dans la fidélité de la traduction sont principalement présentes au sein de la boucle de l'anticodon, en particulier en position 34, la position "wobble", et 37, située immédiatement 3' de l'anticodon (positions 34-35-36). Ces modifications ont un impact direct sur le décodage des ARNm par les ARNt, en augmentant la spécificité de la reconnaissance codon-anticodon, ou inversement en favorisant l'incorporation d'ARNt proches cognats. Certaines de ces modifications ont été bien caractérisées chez les bactéries et différents eucaryotes comme *S. cerevisiae* ou *D. melanogaster*. Cependant, les données ne sont pas facilement transposables d'un organisme à l'autre car il existe des différences notables dans les stratégies de décodage d'organismes très éloignés dans

l'évolution. De nombreuses connaissances sont encore manquantes chez l'Homme en raison de la complexité de l'analyse dans ce modèle. Durant ma thèse j'ai utilisé la translecture des codons stop comme outil pour étudier l'impact de ces modifications sur la fidélité de la traduction sans entrer en compétition avec des ARNt cognats. J'ai commencé par caractériser de manière qualitative et quantitative la translecture du codon stop situé dans le gène *OPRL1*, que j'ai par la suite utilisée comme séquence modèle pour tester l'incorporation des ARNt sur les trois codons stop. Je me suis ensuite concentrée sur le tRNA<sub>Tyr</sub>, suppresseur naturel des codons UAA et UAG. Chez l'Homme, cet ARNt présente deux modifications au niveau de l'anticodon : GalQ<sub>34</sub>, une modification spécifique des animaux qui dépend de l'apport en queuine par le microbiote intestinal et l'alimentation, et qui n'a encore jamais été étudiée pour son rôle dans la reconnaissance des codons ; et  $\Psi_{35}$ , une pseudouridine catalysée par l'enzyme non essentielle Pus7, et connue pour augmenter l'incorporation proche cognat de l'ARNt<sub>Tyr</sub> *S. cerevisiae*. Ainsi, l'apport de queuine dans le milieu de culture de cellules humaines Expi293F a été contrôlé afin d'étudier le rôle de la modification GalQ<sub>34</sub> et une lignée de cellules Expi293F avec knockout du gène *PUS7* a été réalisée par CRISPR/Cas9 pour étudier le rôle de  $\Psi_{35}$ . Dans un premier temps, les lignées cellulaires ont été caractérisées par différentes techniques, dont le Northern Blot acide, le séquençage global des ARNt, le profil des polysomes et le séquençage de ribosomes. Dans un second temps, j'ai étudié l'impact des deux modifications de l'anticodon de l'ARNt<sub>Tyr</sub> sur le taux de translecture et le taux d'incorporation de la Tyrosine au niveau des codons stop UAA et UAG. Les résultats de cette étude ont notamment permis de mettre en évidence que la modification GalQ<sub>34</sub> permet de maintenir la fidélité de la traduction en diminuant l'incorporation de l'ARNt<sub>Tyr</sub> par translecture, alors que la modification  $\Psi_{35}$  est importante pour la reconnaissance de l'ARNt proche cognat. La double délétion des deux modifications a démontré un effet épistatique de la délétion de  $\Psi_{35}$  sur celle de GalQ<sub>34</sub>. Dans l'ensemble, ces résultats fournissent de nouvelles informations sur l'appariement codon-anticodon chez l'Homme, avec une meilleure compréhension des règles régissant le décodage.

## Stop codon readthrough in human: natural suppressors and impact of tRNA anticodon loop modifications

Charlène Valadon<sup>1</sup>, David Cornu<sup>1</sup>, Isabelle Hatin<sup>1</sup>, Stéphane Demais<sup>1</sup>, Olivier Namy<sup>1</sup>

Affiliations: (1) Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Université Paris-Saclay, 91198 Gif-sur-Yvette, France

### Abstract

The entry of a stop codon into the ribosomal A-site usually signals the end of translation and the release of the synthesized peptide. However, it is possible that a tRNA is incorporated at the stop codon, allowing the ribosome to continue translation in the same phase until the next stop codon. This is called stop codon readthrough. During this particular situation it is not a cognate tRNA that is incorporated at the stop codon, but a near-cognate tRNA making only 2 of 3 pairings with the codon. Stop codon readthrough may be favored by either nucleotides surrounding the stop codon, drugs, or by RNAs modifications. While the identity of the tRNAs used during stop codon readthrough is now well known in the yeast *S. cerevisiae*, their identification in humans remains poorly studied. Moreover, there is only sparse evidence for the role of tRNA anticodon loop modifications on the incorporation efficiency of near-cognate tRNAs in human cells. In this study, using quantitative mass spectrometry we revealed that Cys, Trp, Arg and Ser are found at UGA, whereas Gln and Tyr are found at UAA and UAG codons. Then, we investigated the role of two modifications, galQ<sub>34</sub> and Ψ<sub>35</sub>, found in the anticodon loop of the Tyrosine tRNA. Our findings revealed the importance of anticodon modifications in the incorporation of the Tyrosine tRNA by readthrough of UAA and UAG stop codons, enhancing our knowledges about the flexibility of the genetic code.

### Introduction

Translation is the decoding of a mRNA by the ribosome which uses tRNAs to incorporate specific amino acids in front the corresponding codons. Thus, that is safeguarded by the recognition of tRNA by specific aminoacyl-tRNA synthetases and the codon-anticodon pairing by cognate tRNA. Translation fidelity is essential in all domains of life to prevent the misincorporation of amino acids and premature termination. However, 15 codons, translated into amino acids, do not have a complementary anticodon in human genome [42]. Thanks to the flexibility of the genetic code, a mismatch between base 34 of the tRNA and the position 3 of the codon can be tolerated and leads to the incorporation of an amino acid by a near-cognate tRNA. It was long thought that ribosome discrimination between cognate, near-cognate and non-cognate tRNAs was conferred by the stability of the codon-anticodon complex. However, numerous studies have proved that the ribosome plays itself an active role in the incorporation of near-cognate tRNAs by forcing mismatched

base pairs to adopt a Watson–Crick geometry [203]. Indeed, the proofreading activity of the ribosome allows U-G pairing at the first or second codon positions because they interact with bases A1492, A1493 and G530 of helix 44 as canonical Watson-Crick pairings. In addition, helix 44 would accurately check the pairing of codon bases 1 and 2 during tRNA accommodation, tolerating a mismatch at the third position of the codon if it does not trigger steric hindrance.

More than 100 modifications of RNAs are known in human cells, with an average rate of 13 modifications per tRNA. Modifications of tRNAs can serve two main functions: (1) stabilization of the tRNA and its overall L-shaped structure [353]; (2) mRNA decoding and fine-tuning of the translation process [306]. The modifications known to play an active role in translation fidelity are mainly present within the anticodon loop, in particular at positions 34, the “wobble” position, and 37, located immediately 3' to the anticodon. These modifications have a direct impact on the decoding of mRNA by the tRNAs, either by increasing the specificity for codon-anticodon recognition or by promoting the incorporation of near-cognate tRNAs. Some of these modifications have been well characterized in bacteria and in different eucaryotes such as *S. cerevisiae* or *D. melanogaster* [293,306]. However, data are not easily transposable from one organism to another because notable differences exist in the decoding strategies of distantly evolutionary related organisms and a lot of knowledges are still missing in human. Here, we used the readthrough of stop codons as a tool to study the impact of these modifications on translation fidelity without being hindered by cognate tRNAs. We selected the stop codon found in *OPRL1* gene because this stop codon has been described to undergo a high level of natural readthrough in mammals. We first measured the basal readthrough efficiency of the three stop codons in the *OPRL1* context using a dual reporter system. Then we used quantitative mass-spectrometry to identify and quantify the proportion of each amino acids inserted at the stop codon location. Among the various amino acids found, we focused on the incorporation of tyrosine because there is only one possible tRNA that can incorporate tyrosine at both UAA and UAG, tRNA<sub>Tyr</sub> (GUA). In human, the anticodon loop of this tRNA carries two specific modifications: galQ<sub>34</sub>, a modification specific of animals that depends on the intake of queuine from gut microbiota and diet and  $\Psi_{35}$ , a pseudouridine catalyzed by the non-essential Pus7 enzyme and found also in *S. cerevisiae*. Our results showed that while galQ<sub>34</sub> prevents the incorporation of tRNA<sub>Tyr</sub> at stop codons,  $\Psi_{35}$  plays an essential role for stabilizing the tRNA<sub>Tyr</sub> at the stop codons. Altogether, our results highlight the specificity of the decoding process in the human cells and provide new information about codon–anticodon pairing in human cells, with a better understanding of the rules governing decoding.

## Materials and Methods

### Cell Lines and media.

All the cell lines used in this study were derived from the Expi293F cells (Gibco), which are human cells derived from HEK293 cultivated in suspension. This cell line is referred as the WT cell line.  $\Delta PUS7$  cell line has been obtained from WT cells using a CrisPR/CAS9 strategy from Origene KN403902 (GFP-P2A-Puro cassette). Cells were cultivated in a synthetic medium (Expi293™ Expression Medium, Thermofisher) in the absence of queuine (CAS 69565-92-0, Santa Cruz), or with its addition directly in the media with a final concentration of 200nM. Cells were transfected using ExpiFectamine (Thermofisher), and with the addition of enhancers 1 and 2 (Thermofisher) 18h post transfection.

### Knockout of *PUS7* gene by CrisPR/CAS9

WT cells were transfected with the GFP-P2A-Puro cassette and the RNA guide 1 from the kit KN403902 (Origene). Cells were grown for 2 weeks in the synthetic medium, before the selection of the cells expressing GFP by FACS cell sorter. Then, GFP positive cells were grown on plate to isolate clones with Puromycin at 10µg/ml in the medium. The expression of the protein Pus7 has been tested in each clone by western blot (Abcam ab226257 antibody) before DNA extraction and sequencing of the mutated region. Finally, the  $\Delta PUS7$  cell line selected for the study is a monoclonal population with alleles carrying either a deletion of 8 bases (+43 from the AUG) or an insertion of the cassette GFP-P2A-Puro, without any expression of the Pus7 protein.

### Plasmids

Readthrough proteins were produced from pCDNA3.4-GST. This vector was constructed by adapting codon frequencies from the GST gene from pYX24-GST to human, and by the addition of a linker sequence of 171nt (H4) [354] at the AfeI immediately downstream the start codon. Stop codon readthrough was quantified with the dual *LacZ/F-Luc* reporter system from pAC18. This vector was constructed from pAC99 plasmid [355] by the addition of two F2A cleavage sequences in each side of the readthrough cassette. The three stop codons in the OPRL1 context (-6nt and +6nt) have been inserted either in the GST or the dual *LacZ/F-Luc* reporters, using the AfeI or MscI sites respectively. PUS7 rescue expression was done by transient expression of pCDNA3.4-PUS7. This vector was constructed by the insertion of PUS7 cDNA downstream the CMV promoter of pCDNA3.4 plasmid.

### Quantification of readthrough efficiency

Quantification of readthrough efficiency has been done using the pAC18 reporter system. Luciferase and  $\beta$ -galactosidase activities were assayed 42 hours post

transfection, in the same cell extract as previously described [251]. Values are reported as box plots, with median values for five independent measurements. The significance of differences was evaluated using Man-Whitney nonparametric statistical test.

### **Purification of readthrough proteins**

Expi293F cells were transfected by pCDNA3.4-GST to express a readthrough GST protein. Cells' extract was done 42 hours post transfection and GST purified on a 1 ml GSTrap HP column (GE Healthcare) with an AKTA purifier 10 FPLC system as previously described [229]. The product of purification was then migrated on SDS-PAGE, allowing the isolation of *S. japonicum* GST from endogenous GST for mass spectrometry analyses.

### **Mass Spectrometry**

Mass spectrometry analysis was performed on Coomassie-stained gel bands corresponding to the readthrough GST proteins [229]. Briefly, bands were subjected to in-gel LysC digestion as previously described and readthrough peptides were further identified and quantified by nanoLC–MSMS using a nanoElute liquid chromatography system (Bruker) coupled to a timsTOF Pro mass spectrometer (Bruker). Before MS analysis, peptides were cleaned up with ZipTip C18 (millipore), dried under vacuum and then resuspended in 0.1% n-octylglucopyranoside (Sigma–Aldrich) to prevent the loss of hydrophobic peptides. Additionally, synthetic heavy isotope labeled peptides (AQUA, Thermo) were spiked in the mixture of their endogenous counterparts to take the relative ionization efficiencies of peptides and matrix effects into account. Peptides were loaded with solvent A and separated on an Aurora analytical column (ION OPTIK, 25 cm × 75 μm, C18, 1.6 μm) with a gradient of 0–35% of solvent B for 30 minutes. Solvent A was 0.1% formic acid and 2% acetonitrile in water and solvent B was acetonitrile with 0.1% formic acid. MS and MS/MS spectra were recorded from m/z 100 to 1700 with a mobility scan range from 0.6 to 1.4 V s/cm<sup>2</sup>. MS /MS spectra were acquired with the PASEF (Parallel Accumulation Serial Fragmentation) ion mobility-based acquisition mode using a number of PASEF MS/MS scans set as 10. MS and MSMS raw data were processed and converted into mgf files with DataAnalysis software (Bruker). Peptide and protein identifications were performed using the MASCOT search engine (Matrix Science, London, UK) against a homemade database (the unknown amino acid at stop codon is represented by « X » in the readthrough GST sequence) (Figure S1). Database searches were performed using LysC cleavage specificity with two possible missed cleavages. Carbamidomethylation of cysteines was set as fixed modification and oxidation of methionines and tryptophan as variable modifications. Peptide and fragment tolerances were set at 10 ppm and 0.05 Da, respectively. Only ions with a score higher than the identity threshold and a false-positive discovery rate of less than 1% (Mascot decoy option) were considered. The XIC-based relative quantification of

readthrough peptides was corrected by dividing the XIC peak areas of endogenous peptides by their heavy isotope labeled counterparts.

### **tRNA-seq**

Total RNA extractions were done in triplicates for each condition. tRNA are extracted by gel purification from 5 $\mu$ g of total RNA by triplicate and sequenced by ArrayStar Inc. (USA). Bioinformatics analysis has been performed by ArrayStar Inc. by aligning tRNA reads onto high confidence tRNA sequences from GtRNAdb. Due the highly redundant nature of tRNA sequences, multiple alignments are allowed. Importantly among the 12 sequences of tRNA<sup>Tyr</sup>, with a tRNAscanScore > 60, present in GtRNAdb, 5 tRNAs generated from different genomic loci have the same sequence. As they are indistinguishable they count for only 1 tRNA. This why we count only 8 different Tyrosine tRNA. All raw data will be deposited at GEO database.

### **Ribo-seq**

Total RNA was extracted in triplicates for each condition from 610<sup>7</sup> cells by triplicate, without addition of cycloheximide. Ribosome profiling has been done as previous described [356] and analyzed using RiboDoc package [357]. All sequencing data will be deposited at GEO database. Ribosome occupancy is calculated as previously described [306], with the following adjustments. The first nucleotide in the ribosome A site was defined as the +16 position relative to the 5' end of the 30 nucleotides-long RPF. Based on hg38 genome annotation, only reads mapping to coding sequences (CDS) for which the nucleotide in the A site was in-frame with the CDS were collected. We then calculated codon coverage that was normalized by the database size. The P value was calculated using student test with  $\alpha = 0.05$ .

### **Polysome profile**

Polysome profile was performed as previously described [306]. Analyses were done in triplicates from 2.10<sup>7</sup> cells each. Cells were incubated with 100 $\mu$ g/ml of cycloheximide during 15mins before 5min of centrifugation at 300g, 4°C. Pellets were resuspended in 500 $\mu$ l of buffer (Tris-HCl 50 mM, pH 7.4, K-acetate 50 mM, Mg-acetate 10 mM, DTT 1 mM, RNAsine 200 units/ml) and cells broken with needle before 5min of centrifugation at 800g, 4°C. OD<sub>260nm</sub> was measured with nano-spectrophotometer and 15 OD were ultracentrifugated through a gradient of sucrose 7- 47%, at 39000 rpm, 4°C during 2h30. Fractions were analyzed with ISCO.

### **Northern Blotting**

The presence of the galQ<sub>34</sub> on the tRNA<sup>Tyr</sup> was assessed by Northern Blotting. The protocol was based on a previous article [358], with some adjustments. Firstly, the Running buffer used was 1 $\times$  TBE, 0.1 M NaOAc/HOAc, pH 4.8. Secondly, the transfer of RNAs from the gel to the membrane has been done by transfer semi-dry with the

Trans-Blot SD sem-dry electrophoretic transfer cell (Biorad), using 3mA per cm<sup>2</sup>, for 45 mins in 0.5X TBE. Then, RNAs were crosslinking to the membrane by UV-crosslinking two times at 254 nm, 120 mJ. Finally, a second hybridization (5S rRNA) was done as a control of the concentration of total RNA.

Probes sequences (5'-biot):

tRNA<sub>Tyr</sub>:

TCCTTCGAGCCGGASTCGAACCAGCGACCTAAGGATCTACAGTCCTCCGCTCTACCARCTG  
AGCTATCGAAGG

5S rRNA:

ACCAGGCCCGACCCTGCTTAGCTTCCGAGATCAGACGAGATCGGGCGCGTTCAGGGTGGT  
ATGGCCGTAGAC

## Results

### Identification of natural suppressor tRNAs incorporated at the three stop codons in the *OPRL1* context

The first step of our study was to characterize the readthrough of the three stop codons in the *OPRL1* context in the Expi293F cells. *OPRL1* is a mammalian gene known to have a high rate of stop codon readthrough [236]. We began by measuring the rate of readthrough of UAA, UAG and UGA stop codons with the dual reporter system (Figure 33.A). We found an average of rate of readthrough of 0.17%, 0.53% and 7.30% for UAA, UAG and UGA respectively. Although lower than previously reported [236], these values confirm that *OPRL1* stop codon undergoes a high level of stop codon readthrough. Then, we investigated the identity of natural tRNAs suppressors of each stop codon (Figure 33.B). After the purification of readthrough GST proteins, peptides obtained were analyzed by LC-MS/MS and systematically normalized with the corresponding synthetic peptide. We found that the same two amino acids, Glutamine and Tyrosine were inserted at the UAA and UAG codons although at different ratios. In parallel, we observed at UGA stop codon the insertion of four amino acids, Tryptophan, Cysteine, Arginine, and Serine, although at a very low level (0.3%).

Interestingly, we observed that while Tyrosine is the main amino-acid incorporated at both UAA and UAG (54% and 92% respectively) in yeast [229], Tyrosine is significantly less incorporated at these codons in human (13% and 40% respectively). We reasoned that this could be due to of the modifications found in tRNA<sub>Tyr</sub> in human cells. Indeed, the human tRNA<sub>Tyr</sub> has two modifications, the  $\Psi_{35}$ , shared with *S. cerevisiae*, and GalQ<sub>34</sub> which is specific of animals. We decided to explore the role of these modifications (Figure 33.C).

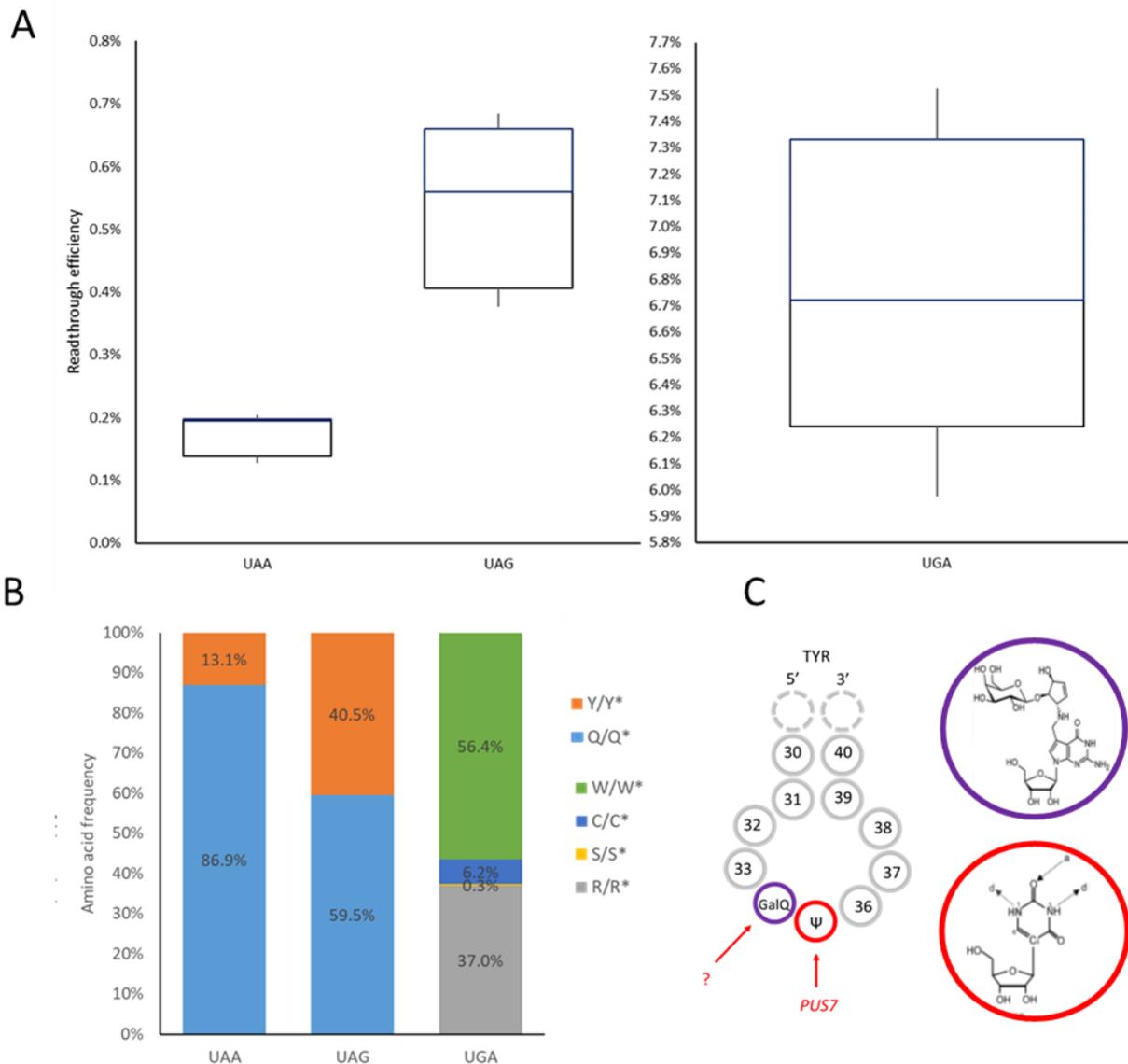


Figure 33 : Readthrough of UAA, UAG and UGA stop codons in the OPRL1 context. (A) Readthrough efficiency of UAA, UAG and UGA stop codons in the OPRL1 context in the Expi293F human cell line. The lines in the center of the boxes indicate the median values; the box limits indicate the 25th and 75th percentiles; the whiskers extend to the minimum and maximum values (n=9). The y-axis scale is different between the left and the right plots to make the left one readable. (B) Relative frequencies of the amino acids incorporated at UAA, UAG and UGA stop codons in the OPRL1 context. Y/Y\*: ratio between the peptide purified from human cells (Y) and the corresponding AQUA peptide (Y\*). n=3 sample points (C) Anticodon modifications of human tRNA<sub>Tyr</sub> and their chemical representation. Purple circle: modification of the position 34. Red circle: modification of the central position of the anticodon, at position 35. The names of the enzymes responsible for the modification are indicated in red. GalQ : galactosyl queuosine .  $\Psi$ : pseudouridine.

## Queuine starvation affects suppressors tRNAs populations but does not impact polysomes

Queuine is the precursor of queuosine (Q) that is found in the anticodon of Tyr, Asp, His, and Asn tRNAs at position 34 where it replaces guanosine [283]. This Q is further modified by the addition of a mannose (ManQ) or a galactose (GalQ) in tRNA Asp and Tyr respectively [285,286]. Queuosine starvation is achieved by omitting

queuine in the culture medium whereas to ensure full tRNA modification with GalQ<sub>34</sub> queuine is added at a final concentration of 200nM. We checked the presence of GalQ<sub>34</sub> modification in these conditions. We performed an acid northern blotting that has been previously shown (ref) to discriminate between tRNA<sub>Tyr</sub> with and without GalQ<sub>34</sub> modification (Figure 34.A). As expected, tRNA<sub>Tyr</sub> does not migrate at the same position depending on the presence of the absence of queuine, whereas the 5S RNA migration is identical. The delay of migration observed under 200nM queuine condition confirms the presence of GalQ<sub>34</sub> which is not present in the 0nM condition. The northern blot only shows a single band in each case, confirming the total modification of the tRNA<sub>Tyr</sub> by GalQ<sub>34</sub> in presence of 200nM queuine. It is important to note that in such conditions an unmodified G<sub>34</sub> and Q<sub>34</sub> are indistinguishable.

Queuine being an essential nutrient we wondered how its absence impacts the tRNA population. To answer this question, we performed a tRNA-seq to quantify iso-decoders expression in presence or absence of queuine (Figure 34.B). The results show a high abundance of all the 8 tRNA<sub>Tyr</sub>(GUA) (in red), in absence of queuine. More surprisingly we noticed that abundance of all tRNA<sub>Gln</sub> (in red) is also increased in presence of queuine, regardless the identity of the anticodon. Then we checked if these variation in tRNA abundance impact translation efficiency (Figure 34.C). To analyse translational status of the cells we performed a polysome profile to analyse the distribution of ribosomes and their subunits on mRNAs. The polysome profile does not show any difference in presence or absence of queuine, suggesting that queuine starvation has no or only a very limited effect on global translation, despite significative changes in tRNAs abundance.

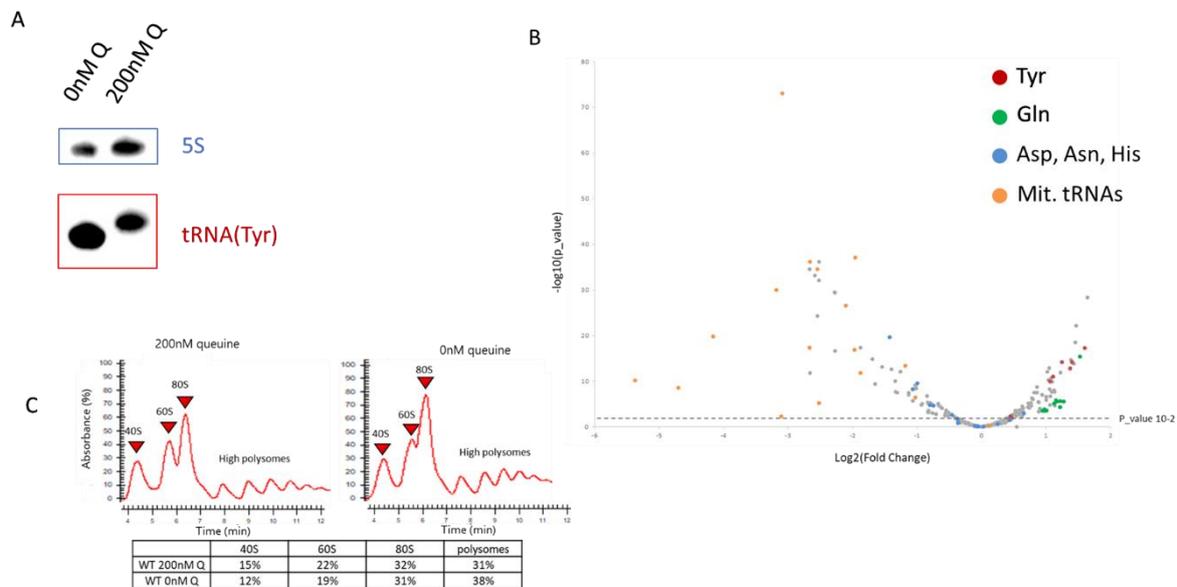


Figure 34 : Global observations about Expi293F queuine starvation. (A) Acid Northern Blotting of total RNA extracted from Expi293F cells cultured without queuine or with 200nM queuine. RNAs of interest were revealed by HRP activity. Blue: 5S rRNA. Red: Tyrosine tRNA. (B) Volcano plot of variations in tRNAs isodecoders pools between Expi293F cells cultured with 0 or 200nM queuine. The Volcano plot was performed from the differential expression calculated using DESeq2 using multimapped reads. All tRNA isodecoders of Tyrosine and Glutamine are represented by red and green dots respectively. The remaining Q-modified tRNAs are indicated in blue and mitochondrial tRNAs are indicated in orange with n=3 samples (C) Polysomes profiles obtained for Expi293F cultured with or without queuine. The table resumes the quantification of the various fractions of the polysome profiles obtained. n=2 samples.

### GalQ<sub>34</sub> prevents UAA and UAG readthrough by tRNA<sub>Tyr</sub>

We assessed the influence of GalQ<sub>34</sub> modification on stop codon decoding by tRNA<sub>Tyr</sub> by comparing stop codon readthrough at UAA and UAG codons in the presence or absence of queuine. We used a dual reporter system  $\beta$ -galactosidase/luciferase to analyze the effect of the absence of GalQ<sub>34</sub> on the levels of UAA and UAG suppression (Figure 35.A). Stop codon readthrough efficiency increases 2.6 fold for both stop codons in the absence of GalQ<sub>34</sub>, which is compatible with the increased abundance of both Tyr and Gln tRNAs previously observed. To determine whether the presence of GalQ<sub>34</sub> impact the ratio between Glutamine and Tyrosine found at stop codons, we performed mass spectrometry on the GST generated by a readthrough event (Figure 35.B). We quantified the frequency of incorporation of Tyrosine and Glutamine at UAA and UAG stop codons in presence or absence of Q. The analysis revealed that Tyrosine incorporation is much more efficient in the absence of queuine at UAA and UAG with a fold change of 4.3 and 1.8 respectively compared to Glutamine, although both tRNA families are more abundant. These results indicate that the presence of GalQ<sub>34</sub> on tRNA<sub>Tyr</sub> limits the incorporation of this tRNA at UAA and UAG stop codons.

Interestingly the anticodon of the tRNA<sub>Tyr</sub> is also modified at position 35, where the U is replaced by a pseudouridine. Pseudouridines are well-known to increase RNA-RNA interactions and stacking to stabilize near-cognate tRNA within the ribosome. It

was interesting to determine whether these results obtained in yeast, where  $\text{tRNA}_{\text{Tyr}}$  lacks the GalQ<sub>34</sub> modification, is also true in human cells.

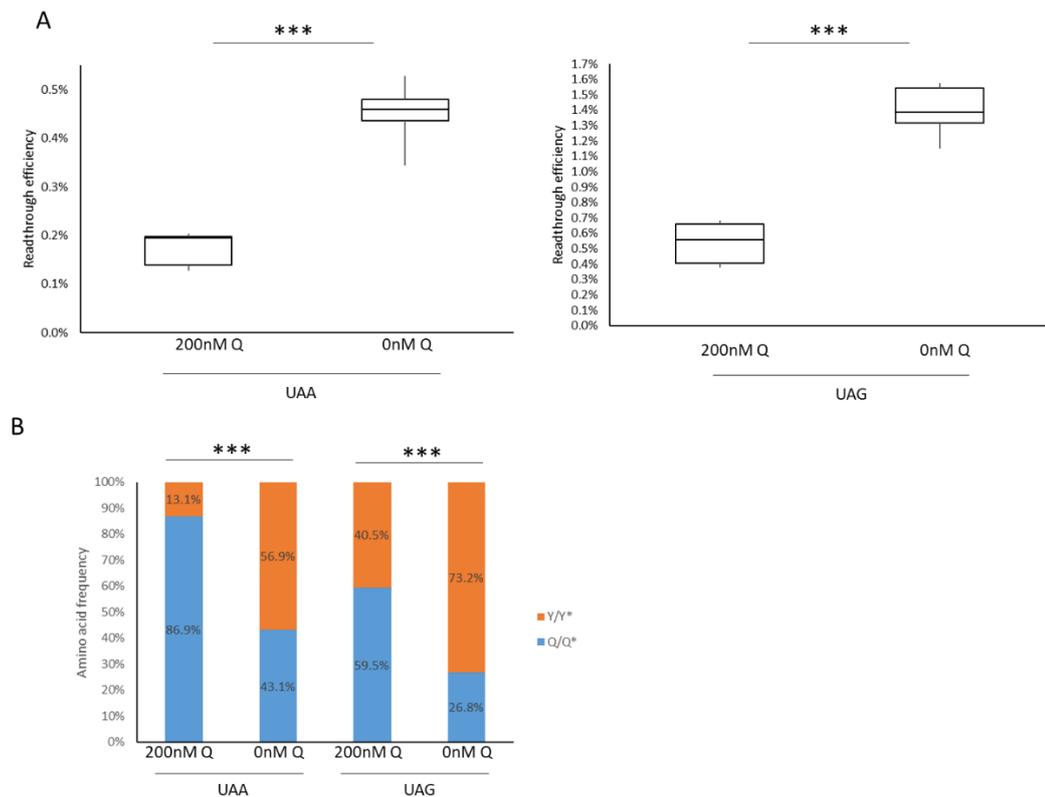


Figure 35 : Readthrough of UAA and UAG stop codons in Expi293F cells cultured with 200nM or no queuine. (A) Readthrough efficiency of UAA and UAG stop codons in Expi293F with or without queuine in the medium. The lines in the center of the boxes indicate the median values; the box limits indicate the 25th and 75th percentiles; the whiskers extend to the minimum and maximum values.  $n=11$  sample points.  $p$  values were obtained by Wilcoxon Test :  $>0.05$  (\*);  $> 0.01$  (\*\*);  $> 0.001$  (\*\*\*). The y-axis scale is different between the left and the right plots. (B) Relative frequencies of the readthrough amino acids incorporated at UAA and UAG stop codons in the WT cells cultured with 200nM or no queuine. Y/Y\*: ratio between the peptide purified from human cells (Y) and the corresponding AQUA peptide (Y\*).  $n=3$  sample points.  $p$  values were obtained by Student Test :  $>0.05$  (\*);  $> 0.01$  (\*\*);  $> 0.001$  (\*\*\*).

## Construction of a PUS7-knockout cell line

In human cells  $\Psi_{35}$  is catalysed by PUS7 [296,308], to remove this modification, we employed CRISPR/Cas9 to generate a PUS7-Knockout EXPI293F cell line, denoted  $\Delta$ PUS7 (Figure S2). The selected clone is totally depleted of the enzyme Pus7p as observed by western-blot (Figure S2A). Moreover, sequencing of the PUS7 locus indicates an homogeneous 8nt depletion at the position of the guide RNA used during the CRISPR/Cas9 (Figure S2B), suggesting the clonal nature of the cells. As previously reported [296] the  $\Delta$ PUS7 cells do not show any difference in cell proliferation. However it was reported that the absence of PUS7 can significantly alter the translation of many genes in stem cells [296]. So, we decided to analyze translational status of the  $\Delta$ PUS7 cell line to identify potential translational defects.

## Pus7 depletion has no impact on tRNAs populations but stimulates global translation

We first assessed the impact of Pus7 deletion on tRNAs populations. A tRNA seq was performed with WT and  $\Delta$ PUS7 cell lines and tRNA populations were compared. As observed in Figure S3.A a very few tRNA abundance significantly change between both cell lines, and none of the tRNA<sub>Tyr</sub> isodecoders are affected (red dots) (Figure S3.A). Previous works reported that global translation is increased in absence of PUS7 [296], so we decided to confirm this observation by performing a polysome profile (Figure S3.B). We indeed observed different profiles with an increase of the polysomes and a decrease of the 80S population in the  $\Delta$ PUS7 cell line compared to the WT comparable to thus previously published [296].

## The presence of $\Psi_{35}$ is important for the incorporation of tRNA<sub>Tyr</sub> at UAA and UAG by readthrough

While  $\Psi_{35}$  is known to increase the ability of stop codon readthrough by tRNA<sub>Tyr</sub> in *S. cerevisiae* [306], we assessed its role in human in the presence of its neighboring modification, GalQ<sub>34</sub>. We first ask whether there is an interdependency between the modifications at positions 34 and 35 of Tyrosine tRNA. Indeed, it is well-known that some modification enzymes require prior modifications at other positions for efficient modification [359]. We checked that GalQ<sub>34</sub> modification is still present in tRNA<sub>Tyr</sub> in  $\Delta$ PUS7 cells (Figure 36.A). The migration of RNAs clearly indicates that tRNA<sub>Tyr</sub> is still fully modified by GalQ<sub>34</sub> in absence of  $\Psi_{35}$ . We then quantified readthrough efficiency in  $\Delta$ PUS7 cells to analyze the effect of the absence of  $\Psi_{35}$  (Figure 36.B). We observed a slight but significant increase of readthrough efficiency at both of stop codons in the  $\Delta$ PUS7 cell line compared to the WT. This effect disappears when Pus7 is expressed from a cDNA in the  $\Delta$ PUS7 cells, indicating that the effect is due to the absence of  $\Psi_{35}$ . Then, we used the GST-readthrough system to quantify the frequency of incorporation of Tyrosine and Glutamine at UAA and UAG stop codons in the  $\Delta$ PUS7 cell line (Figure

36.C). The mass spectrometry analysis shows a decrease of the insertion of Tyrosine in the  $\Delta PUS7$  cells at both terminating codons. This effect is more pronounced for the UAG stop codon (40.5% to 14.7%) that has a higher natural rate of Tyrosine incorporation than the UAA codon. The level of Tyrosine incorporation is restored in the control cell line expressing PUS7 from a cDNA cloned into a plasmid. Our results indicates that  $\Psi_{35}$  plays an essential role to allow tRNA<sub>Tyr</sub> to decode UAA and UAG stop codons.

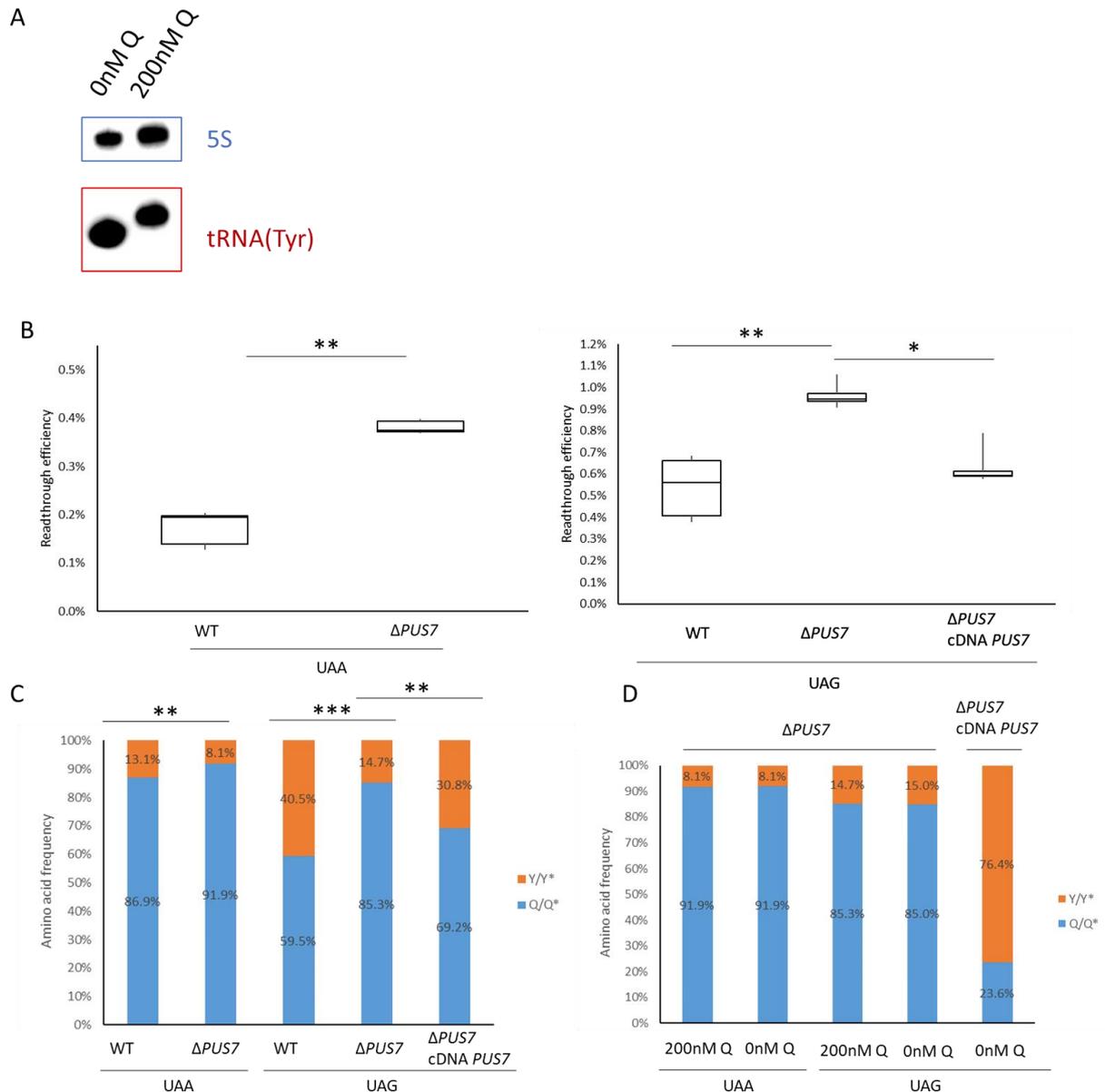


Figure 36 : Impact of *Pus7* deletion on the readthrough of UAA and UAG stop codons (A) Acid Northern Blotting of total RNA extracted from  $\Delta PUS7$  cells cultured without queuine or with 200nM queuine. RNAs of interest were revealed by HRP activity. Blue : 5S rRNA. Red : Tyrosine tRNA. (B) Readthrough efficiency of UAA and UAG stop codons in WT cells,  $\Delta PUS7$  cells or  $\Delta PUS7$  cells transfected with the *PUS7* cDNA, and cultured with 200nM queuine. The lines in the center of the boxes indicate the median values; the box limits indicate the 25th and 75th percentiles; the whiskers extend to the minimum and maximum values.  $n=5$  sample points.  $p$  values were obtained by Wilcoxon Test :  $>0.05$  (\*);  $>0.01$  (\*\*);  $>0.001$  (\*\*\*). The y-axis scale is different between the left and the right plots. (C-D) Relative frequencies of the readthrough amino acids incorporated at UAA

and UAG stop codons.  $Y/Y^*$ : ratio between the peptide purified from human cells (Y) and the corresponding AQUA peptide ( $Y^*$ ).  $n=3$  sample points.  $p$  values were obtained by Student Test :  $>0.05$  (\*);  $> 0.01$  (\*\*);  $> 0.001$  (\*\*\*). (C) Amino acids frequencies in WT cells,  $\Delta PUS7$  cells or  $\Delta PUS7$  cells transfected with the PUS7 cDNA, and cultured with 200nM queuine. (D) Amino acids frequencies in  $\Delta PUS7$  cells or  $\Delta PUS7$  cells transfected with the PUS7 cDNA, and cultured with 200nM or no queuine.

### **The stabilizing role of $\Psi_{35}$ is epistatic on GalQ<sub>34</sub> within the tRNA<sub>Tyr</sub>**

Next, we tested the effect of the loss of GalQ<sub>34</sub> and  $\Psi_{35}$  on the decoding of UAA and UAG by tRNA<sub>Tyr</sub> in the  $\Delta PUS7$  cell line cultured without queuine. We directly assessed the frequency of incorporation of Tyrosine and Glutamine at UAA and UAG stop codons by mass spectrometry with the GST reporter system (Figure 36.D). We observe no differences in Tyrosine incorporation in presence or in absence of queuine in  $\Delta PUS7$  cells for both stop codons. The expression of PUS7 from a cDNA plasmid restores the level of Tyrosine incorporation observed in the WT cells in absence of queuine (compare with Figure 3.B). These results demonstrate that when  $\Psi_{35}$  is absent, the modification status of the position 34 does not influence anymore the efficiency of tRNA<sub>Tyr</sub> incorporation. In other words, the effect of  $\Psi_{35}$  is epistatic on the effect of GalQ<sub>34</sub> when tRNA<sub>Tyr</sub> is used as a near-cognate tRNA.

### **Genome-wide translation in $\Delta PUS7$ cells**

We observe a clear effect of the absence of  $\Psi_{35}$  into the ability of the tRNA<sub>Tyr</sub> to decode both UAA and UAG stop codons. We wondered if this effect was also true when tRNA<sub>Tyr</sub> is used as a cognate tRNA to decode Tyrosine codons. To answer this question, we performed a ribosome profiling experiment in  $\Delta PUS7$  cells to quantify in-vivo translation speed through the calculation of the genome-wide ribosome occupancy at each codon [360]. We first checked the periodicity of the signal to demonstrate that signals obtained correspond to actively translating ribosomes (Figure 37.A). We compared for both the WT and the  $\Delta PUS7$  cells the density of each codon within the A-site of the ribosome (see material and methods)(Figure 37). We observed that ribosomes slow down on Tyrosine codons UAU and UAC (log2FC 0.6  $p_{adj}$   $1.9 \cdot 10^{-6}$  and Log2FC 0.5  $p_{adj}$   $2.6 \cdot 10^{-5}$  respectively), in the  $\Delta PUS7$  cells (Figure 37. B). We also observed that the CGG codon, decoded by tRNA<sub>Arg</sub>(CCG) which is known to be modified by Pus7 at the position 50 [308], showed a decrease of ribosome occupancy (Log2FC - 0.5  $p_{adj}$   $2 \cdot 10^{-4}$ ), suggesting an increase speed of translation of these codons in the  $\Delta PUS7$  cells (Figure 37. C). These results clearly indicates that  $\Psi_{35}$  plays an important role in the decoding by tRNA<sub>Tyr</sub> whether it is used as a cognate or a near-cognate tRNA.

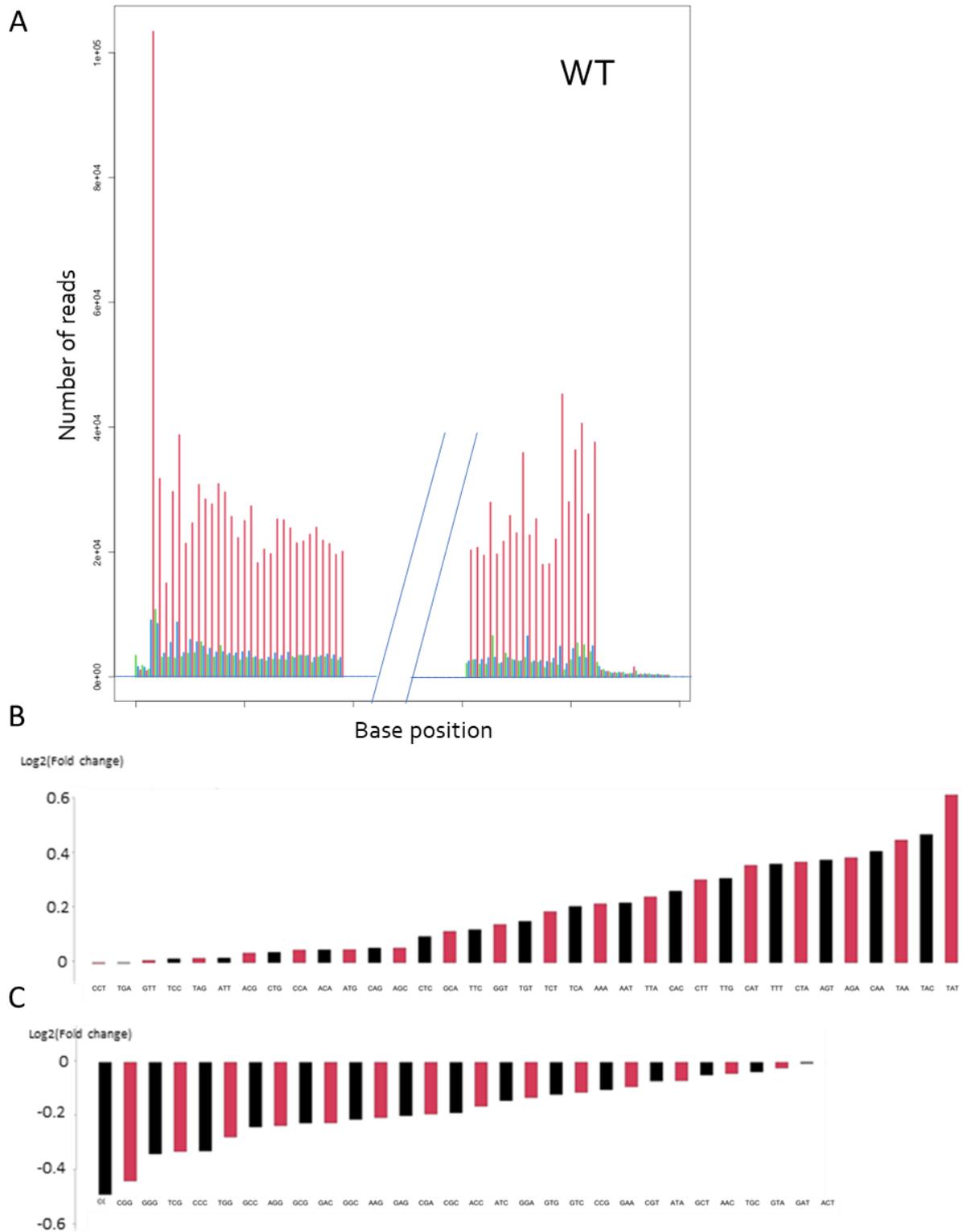


Figure 37 : Impact of *Pus7* deletion on global translation observed by Ribo-seq analyses. (A) Metagene periodicities obtained for WT and  $\Delta$ PUS7 cells cultured with 200nM queuine. Each graph shows the coverage of the CDS and UTR regions of the metagene for a specific read length (here 30 nucleotides), determined by representing all the reads starting at a specific relative position. Each of the three possible reading frames is represented in a different color (blue, green and red). (B-C) Codon bias in the  $\Delta$ PUS7 cell line compared to the WT. (B) In-phase codons with a higher ribosome occupancy in the  $\Delta$ PUS7 cells compared to the WT cells. (C) In-phase codons with a lower ribosome occupancy in the  $\Delta$ PUS7 cells compared to the WT cells.

## Discussion

This work represents the first systematic analysis of the amino-acids incorporated by stop codon readthrough in human cells (Figure 38). To perform this analysis, we used the recoded *OPRL1* stop codon that is submitted to a high level of readthrough in mammals [236]. Although we found a readthrough efficiency lower than previously reported, probably due to the use of different reporter systems, the level of readthrough observed is high enough to obtain correct amount of protein to perform the mass spectrometry analysis. This favorable nucleotide context allowed us to unambiguously identify amino acids using the three stop codons. It is important to note that replacing the natural UGA codon by UAA or UAG has a profound effect on readthrough efficiency as already published (7.3%, 0.17% and 0.53% respectively)[236].

Mass spectrometry analysis revealed that the same amino acids are inserted at the stop codon in human and in yeast, although at different ratio. However, we observed the incorporation of a serine at the UGA codon that was not observed in yeast. It is highly probable that both the nucleotide context around the stop codon and the nature of tRNA modifications play an important role in the selection of the near-cognate tRNAs during stop codon readthrough. Serine can be delivered at the UGA codon by tRNA<sub>Ser</sub>(UGA). This tRNA would be incorporated with a central mismatch which could explain the low abundance of Serine found in our assay. It has been previously published that among the 19 tRNA<sub>Ser</sub> isodecoders only half have the ability to decode stop codon (once their anticodon has been changed) [361] suggesting that in absence of change in the anticodon the situation is probably worst. Alternatively, tRNA<sub>Ser</sub>(IGA) has the ability to recognize all Ser codons due to the presence of the inosine [362] and could be an efficient tRNA suppressor. Our study also highlights a change in the ratio of natural suppressors found in the Human compared to the yeast. For example, the Tyrosine was systematically the major natural suppressor of both UAA and UAG codons in the yeast whereas in human cells we found Glutamine as the major one. The presence of different tRNA modifications explain these observations. Indeed, the efficacy of the incorporation of Tyrosine by readthrough of UAA and UAG codon was similar between the human cells and yeast in the absence of the GalQ<sub>34</sub> modification (which is naturally absent in yeast).

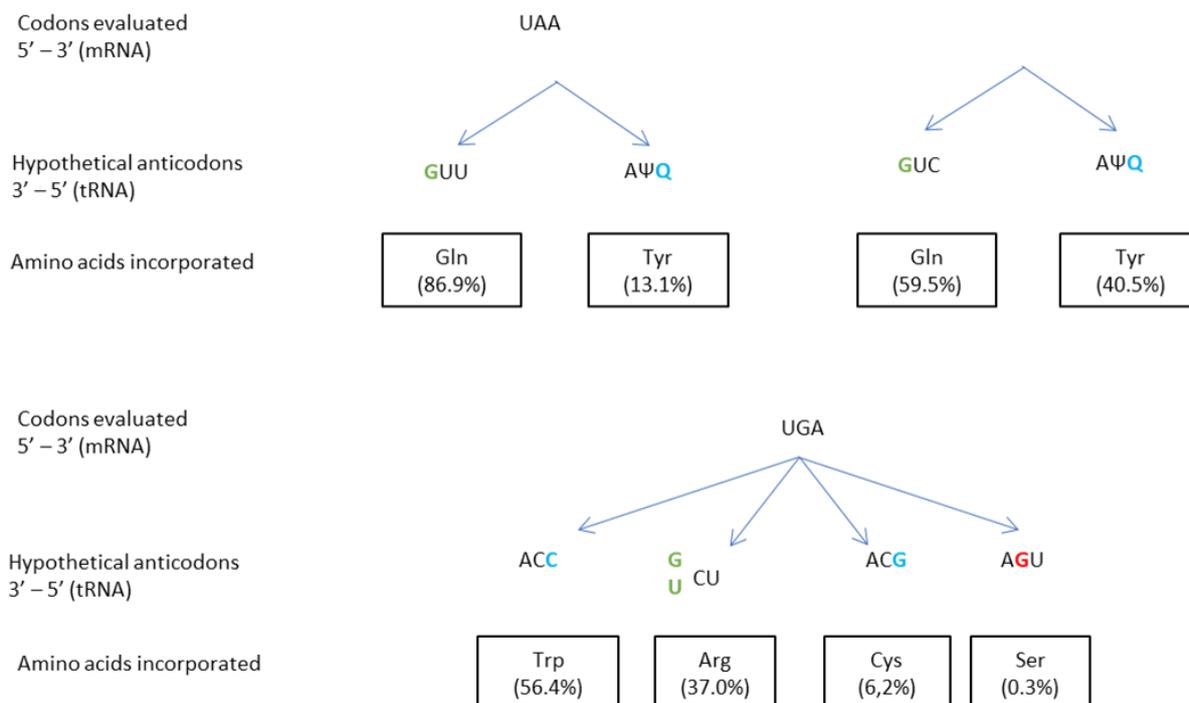


Figure 38 : Prediction of anticodons substitution in front of stop codons based on the amino acid mis-incorporated in our study. The mismatching base is indicated in blue (wobble position), green (36<sup>th</sup> position) or red (central position) in the anticodon. Codon-anticodon pairings evaluated with more than one mismatch are not represented. The ratio of incorporation found by mass spectrometry in WT Expi293F are indicated in parentheses. Ψ: pseudouridine. Q : Galactosyl-queuosine.

Here we addressed the role of  $\Psi_{35}$  found in the anticodon of the tRNA<sub>Tyr</sub> for the accurate decoding of the tRNA. This modification is catalyzed in yeast and in human cells by the protein Pus7 [296,308]. Pus7 has been described to modify a large set of tRNAs and to be essential for the generation of tRFs during differentiation. Using two specific reporter systems we were able to identify the essential role played by  $\Psi_{35}$ . Actually, this modification is essential to allow the tRNA to establish sufficiently stable interactions with the codon to allow the incorporation of the tRNA into the ribosome. When the tRNA<sub>Tyr</sub> is not efficiently incorporated we observed that it is replaced by the tRNA<sub>Gln</sub>. As we observed a slight increase in stop codon readthrough efficiency in this situation, we propose that tRNA<sub>Gln</sub> is kinetically more efficient to be accommodated by the ribosome than the tRNA<sub>Tyr</sub>. Due to the importance of  $\Psi_{35}$  for the incorporation of the tRNA<sub>Tyr</sub> at stop codons, we asked if it was also the case for Tyrosine codons. To answer this question we used ribosome profiling, with high-quality phasing we were able to calculate the ribosome A-site occupancy. This revealed that both Tyrosine codons (UAU and UAC) are overrepresented in ribosome A-site, indicating that ribosomes struggle to decode these two codons. We also observe an underrepresentation of the Arg codon within the ribosomal A-site, indicating that this codon is read faster in absence of Pus7. The tRNA<sub>Arg</sub> does not carry a pseudouridine

within the anticodon loop. However, it is a substrate of Pus7 at the U<sub>50</sub> position [308], which could impact the ability of the tRNA to decode the CCG codon.

Our study is also the first one to investigate the role of GalQ<sub>34</sub> modification on translation fidelity. We observed that GalQ<sub>34</sub> modification decreases the incorporation of tRNA<sub>Tyr</sub> at UAA and UAG codons. The simplest explanation is that the presence of GalQ<sub>34</sub> destabilizes the codon/anticodon pairing when the tRNA<sub>Tyr</sub> is used as a near cognate tRNA (i.e in presence of the stop codon). This observation is interesting because the queuosine modification alone (without galactosyl) is known to be more permissive for wobble mismatches in the drosophila [294]. So the addition of galactose has an important effect on the behavior of the tRNA<sub>Tyr</sub> and its decoding capacities. We next addressed the question of the decoding of sense codons UAU and UAC by tRNA<sub>Tyr</sub> (GUA). First, we did not notice any growth defect in absence of queuine, indicating that Tyrosine codons are decoded whatever is the modification status of the G<sub>34</sub>. However, there is another tRNA<sub>Tyr</sub> predicted in human genome, with the anticodon AUA (<http://gtrnadb.ucsc.edu/>). Although this tRNA could be considered as a UAA and UAG natural suppressor with an A<sub>1</sub>:A<sub>3</sub> or A<sub>1</sub>:G<sub>3</sub> wobble mismatch respectively, we recently demonstrated that this tRNA has no suppression activity [363]. Interestingly, this tRNA is only detected by tRNA-seq (uniquely mapped reads) in the WT cell line when queuine is present in the medium although it does not carry any queuosine modification. We could hypothesize that tRNA<sub>Tyr</sub>(AUA) is expressed in the presence of queuine to compensate for the reduced efficiency of the galactosylated tRNA<sub>Tyr</sub>(GUA) to act as near-cognate tRNA on Tyrosine UAU codons. More studies are needed to better understand the role of GalQ<sub>34</sub> modification on tRNA<sub>Tyr</sub>. Moreover, the increase of stop codon readthrough in the absence of queuine could give opportunities of readthrough therapy for patients with nonsense mutation, with an adjustment of queuine import in their diet.

Supplementary Data are available on page 183.

## **2. Résultats supplémentaires sur les modifications de la boucle anticodon des ARNt**

Le projet portant sur les modifications de la boucle anticodon des ARNt a nécessité un certain nombre de mise au point tout au long de mon doctorat. Pour commencer, j'ai dû mettre au point un système rapporteur de l'incorporation des acides aminés en lignées humaines. Concernant la modification GalQ<sub>34</sub>, il était nécessaire de déterminer la concentration optimale de queue avec laquelle travailler afin d'étudier l'impact sur la translecture. Ensuite, j'ai dû obtenir une lignée cellulaire n'exprimant plus la protéine Pus7 dans le but d'étudier l'impact de la  $\Psi_{35}$ . Il est apparu nécessaire également d'étudier l'impact de la délétion du gène *PUS7L* codant la protéine Pus7L, considérée comme paralogue de la protéine Pus7, absente de la levure et n'ayant pas été caractérisée pour sa potentielle redondance avec Pus7.

### **A. Mise au point de l'analyse des acides aminés incorporés par translecture**

#### **a. La GST**

L'analyse quantitative de l'incorporation des acides aminés par translecture par spectrométrie de masse a tout d'abord nécessité la mise au point d'un système rapporteur adapté aux cellules humaines. Un système rapporteur exprimant la protéine Glutathion S-transferase (GST), issue de l'organisme *S. japonicum*, dont la traduction complète dépend de la translecture d'un codon stop prématuré en début de séquence était déjà présent au sein du laboratoire, mais avait été mis au point pour une expression chez la levure.

Ainsi, la séquence du gène codant la GST a été adaptée à l'Homme en modulant la fréquence des différents codons, et a été clonée dans le plasmide pCDNA3.4 sous le contrôle du promoteur CMV (**Error! Reference source not found..A**). Le site de restriction AfeI était le site d'insertion des séquences de translecture à étudier. Afin de déterminer la nature des acides aminés incorporés, j'ai débuté l'expression de la GST avec différentes séquences de translecture en cellules humaines Expi293F, des dérivées non adhérentes des cellules HEK293, avant de la purifier sur colonne de glutathion (**Error! Reference source not found..B**). Après purification, la fraction contenant la GST est déposée sur un gel d'acrylamide. La révélation des protéines avec le bleu de Coomassie met en évidence deux bandes distinctes, alors que la révélation par anticorps spécifique de la GST en Western Blot permet de confirmer la présence spécifique de la protéine d'intérêt dans la bande supérieure. Les protéines ont ensuite été analysées par spectrométrie de masse par David Cornu. L'analyse quantitative du taux d'incorporation des différents acides aminés a alors été compromise par une dégradation protéolytique de l'extrémité N-terminale de la protéine. Ceci rend impossible la quantification des différents peptides. Il a donc été nécessaire de changer

de système rapporteur, puisque malgré toutes nos tentatives nous n'avons pas réussi à limiter cette dégradation.

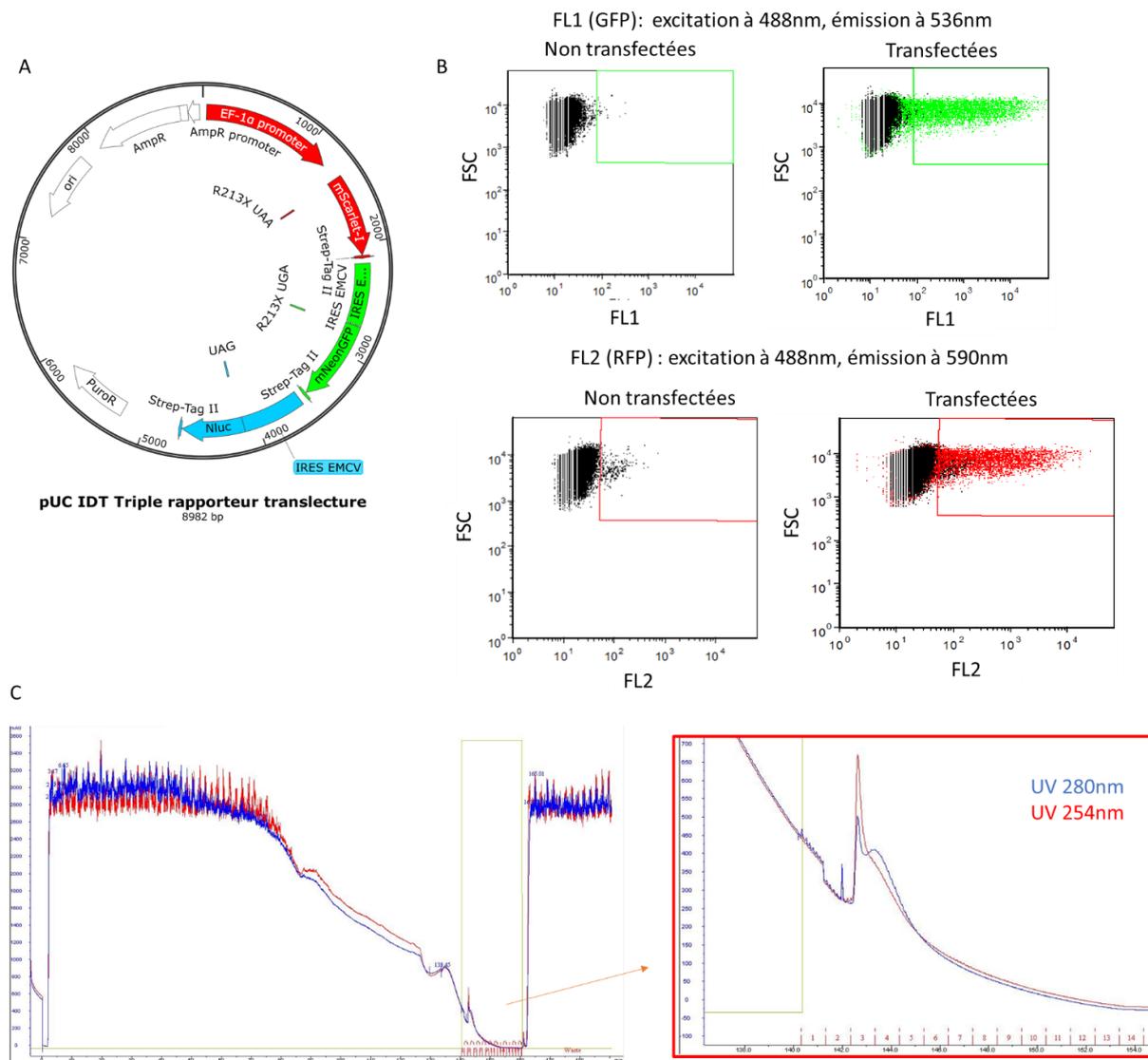


Figure 39 : Mise au point du système triple rapporteur de la translecture en cellules humaines. (A) Carte plasmidique du système triple rapporteur. Les flèches mScarlet, mNeonGFP et Niuc représentent respectivement les gènes codants pour les protéines Scarlet, Néon et NanoLuc. PuroR : gène de résistance à la Puromycine. AmpR : gène de résistance à l'Ampicilline. (B) Répartition des cellules humaines HeLa en fonction de leur fluorescence en cytométrie de flux. FSC : taille des cellules déterminée par la mesure de la lumière diffractée par les cellules en face du rayon laser. Les canaux FL1 et FL2 représentent la mesure de l'émission fluorescente par les cellules respectivement à 536 et 590nm après excitation à 488nm. (C) Chromatogramme de la purification des trois protéines Scarlet, Néon et NanoLuc par fixation du Strep-Tag en colonne Strep-Tactin. La courbe bleue représente la mesure de l'absorbance à 280nm (protéines) et la courbe rouge celle à 254nm (acides nucléiques). L'encadré vert représente la fenêtre d'élution.

Nous avons décidé de construire un système triple rapporteur.

### **b. Le système triple rapporteur**

Le but de mon projet étant d'étudier l'impact de différentes modifications de la boucle anticodon des ARNt, nous avons pensé à un plasmide unique permettant de purifier en une seule étape trois protéines issues d'un événement de translecture des trois codons stop respectifs, dans un contexte nucléotidique identique (Figure 39.A).

Le choix du contexte nucléotidique s'était porté sur la séquence R213X issue d'une mutation du gène p53 chez des patients atteints de cancer, qui présente un des taux de translecture le plus élevé de l'ensemble des mutations à codon stop prématuré [364]. Les trois protéines issues des événements de translecture ont donc été choisies en fonction de la possibilité de détecter leur expression par simple mesure de leur activité. Les gènes rapporteurs choisis ont été ceux codants les protéines Scarlet (longueur d'onde d'excitation à 569nm et émission à 594nm), Néon (longueur d'onde d'excitation à 506nm et émission 517nm), et NanoLuc (luciférase), interrompus respectivement par les codons stop UAA, UGA et UAG. Les trois gènes étaient sous le contrôle du même promoteur et ainsi traduits à partir du même ARNm. La traduction des deux dernières protéines Néon et NanoLuc est réalisée grâce à l'IRES EMCV. Les trois protéines possèdent un tag de purification Strep-Tag II afin de les purifier ensemble sur une colonne Strep-Tactin.

Une fois le plasmide construit, j'ai testé l'expression des différentes protéines par transfection transitoire en cellules humaines Hela. Mes premières mesures mettent en évidence une luminosité relative moyenne plus de 3000 fois supérieure dans la population transfectée. L'activité des protéines Scarlet et Néon a été mesurée en cytométrie de flux avec des cellules Hela transfectées ou non avec le triple rapporteur (Figure 39.B). Le canal FL1, correspondant à la détection de GFP, permet de mettre en évidence les cellules avec de la protéine Néon active et le canal FL2, correspondant à la détection de RFP, permet de détecter les cellules exprimant une protéine Scarlet active. L'interprétation des résultats a été limitée par le fait que la longueur d'onde d'excitation était la même dans les deux canaux et que les longueurs d'ondes d'émissions étaient très proches entre les deux protéines. Il n'a donc pas été possible d'être certains que le cytomètre distinguait les deux activités protéiques. En revanche, la proportion de cellules exprimant une protéine Néon ou Scarlet active étant d'environ 10% au sein de l'ensemble des populations transfectées, j'ai considéré ce taux comme l'efficacité de transfection des cellules Hela avec le triple rapporteur. L'efficacité de transfection étant très basse en cellules Hela, j'ai décidé d'utiliser un autre modèle cellulaire, les cellules humaines Expi293F.

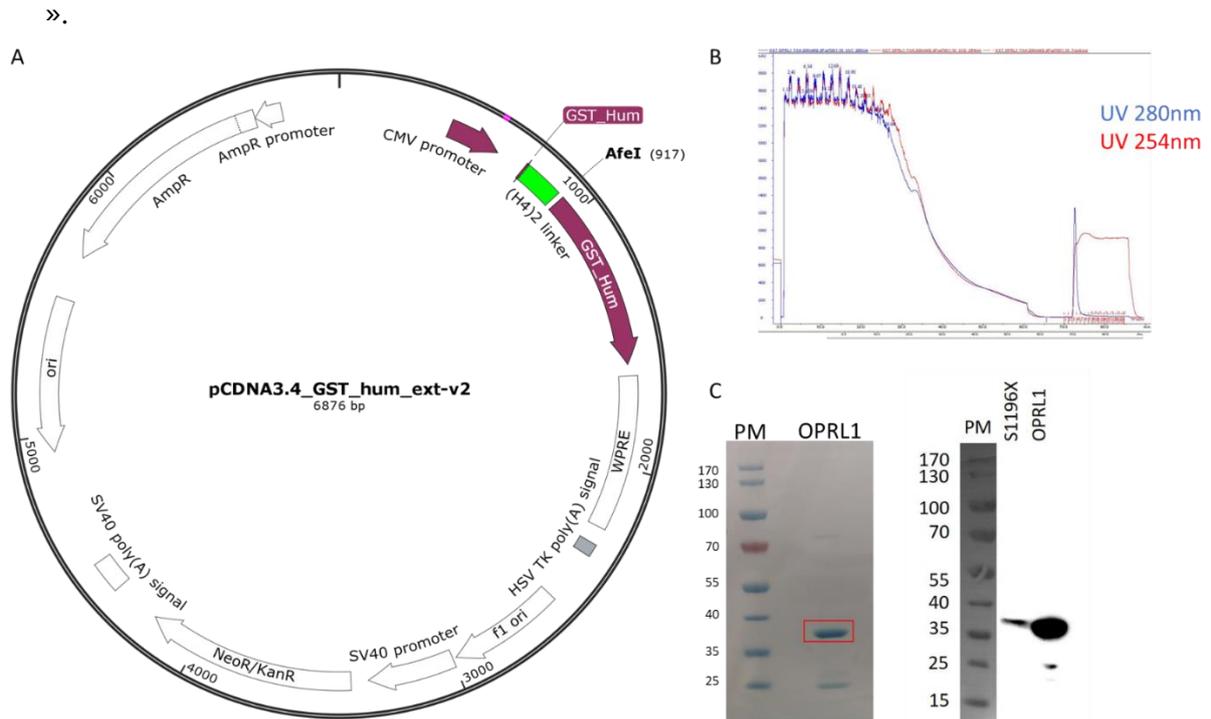


Figure 40 : Le système rapporteur GST avec extension N-terminale. (A) Carte plasmidique du système rapporteur GST avec extension N-terminale. (H4)2 linker : extension N-terminale. AfeI : site de clonage des séquences de transfecture. (B) Chromatogramme de la purification de la GST par fixation sur le glutathione. La courbe bleue représente la mesure de l'absorbance à 280nm (protéines) et la courbe rouge celle à 254nm (acides nucléiques). (C) Migration sur gel acrylamide de la protéine GST avec extension N-terminale. OPRL1 et S1196X : GST avec insertion de séquences de transfecture au niveau du site AfeI. PM : échelle de poids moléculaires. Gauche : révélation au bleu de Coomassie. Encadré rouge : bande spécifique de la GST issue d'un événement de transfecture d'OPRL1. Droite : Western Blot avec anticorps dirigé contre la GST.

Les Expi293F sont issues de la lignée des cellules HEK293 et sont non adhérentes, permettant d'exprimer une grande quantité de protéines dans de petits volumes de culture ce qui était idéal pour l'expression de protéines issues d'évènements de translecture.

Afin de réaliser les premières mesures de l'incorporation des acides aminés par translecture des trois codons stop, j'ai transfecté le plasmide triple rapporteur dans les cellules Expi293F. J'ai ensuite purifié les protéines sur une colonne Strep-Tactin (Figure 39.C) avant de les faire analyser par spectrométrie de masse par David Cornu. L'analyse par spectrométrie de masse met en évidence l'incorporation des acides aminés Tyrosine et Glutamine au niveau du codon UAA (Scarlet), des acides aminés Cystéine, Tryptophane et Arginine au niveau du codon UGA (Néon), mais aucun peptide issu de la NanoLuc n'a été malheureusement détecté.

### **c. La GST avec extension N-terminale**

En parallèle de la mise au point du système triple rapporteur, le système rapporteur avec la protéine GST a été finalisé par Sabrina Karri, une autre doctorante de l'équipe. Ne pouvant pas mesurer efficacement l'incorporation des acides aminés au niveau des trois codons stop avec le système triple rapporteur, j'ai ainsi eu la possibilité d'utiliser le système rapporteur GST (Figure 40.A). Il s'agit de la séquence hélicoïdale (H4)<sub>2</sub> linker, publiée pour favoriser la stabilité des protéines de fusion [354]. La séquence de translecture R213X ayant un taux de translecture très bas, nous avons décidé d'étudier la translecture des trois codons stop dans un nouveau contexte de translecture nommé OPRL1. J'ai cloné les trois codons stop dans le contexte OPRL1 au niveau du site AfeI de la GST avec l'extension N-terminale. J'ai purifié les différentes protéines GST sur colonnes glutathion (Figure 40.B) avant de les analyser sur gel d'acrylamide (Figure 40.C). Le western blot présenté en Figure 40.C est un parfait exemple de l'avantage d'utiliser un contexte efficace comme OPRL1 contrairement à un contexte avec une efficacité plus faible (S1196X). Dans cet exemple, les deux contextes de translecture comportent un codon stop UGA et les protéines ont été purifiées à partir des mêmes conditions de cultures. La révélation de la GST par western-blot montre une quantité de protéine supérieure avec le contexte OPRL1 par rapport au contexte S1196X.

À la suite de la mise au point de ce nouveau système rapporteur, j'ai eu l'opportunité d'être formée par David Cornu à l'analyse des séquences peptidiques par spectrométrie de masse. J'ai ainsi pu effectuer l'ensemble des étapes nécessaires à l'analyse quantitative de l'incorporation des acides aminés par translecture pour mon projet, de l'expression en cellules humaines à l'analyse par spectrométrie de masse dont les résultats ont été présentés précédemment dans l'article de résultats ci-dessus.

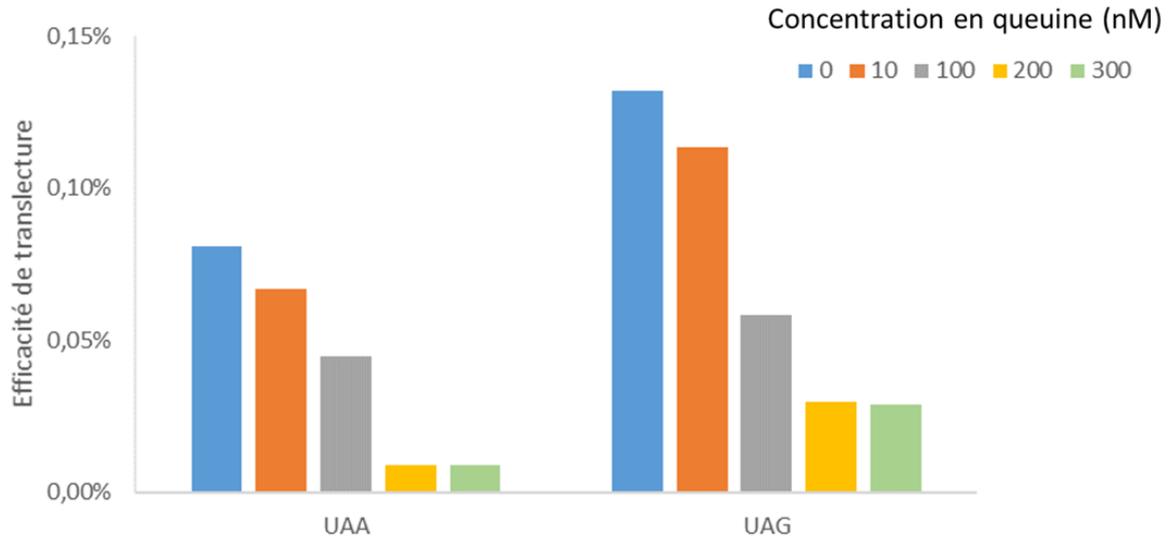


Figure 41 : Efficacité de transfection des codons stop UAA et UAG dans le contexte R213X en fonction de la concentration en quinine dans le milieu de culture des cellules Expi293F. Les mesures ont été effectuées avec le système double rapporteur  $\beta$ -Galactosidase / Luciférase. n=5

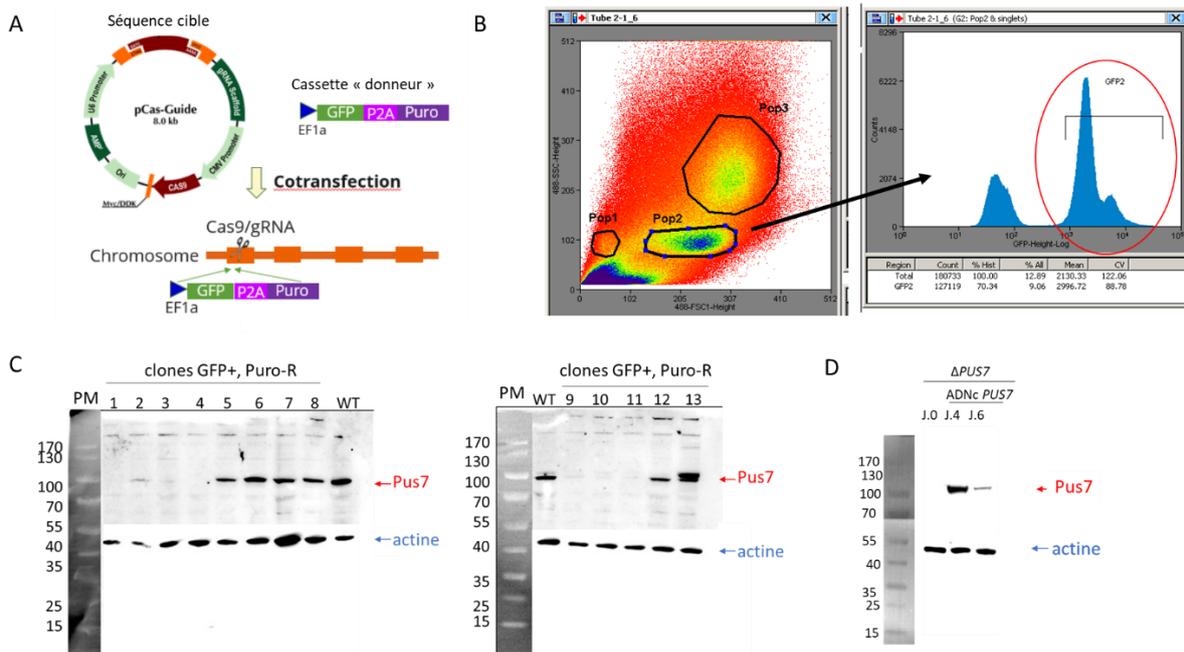


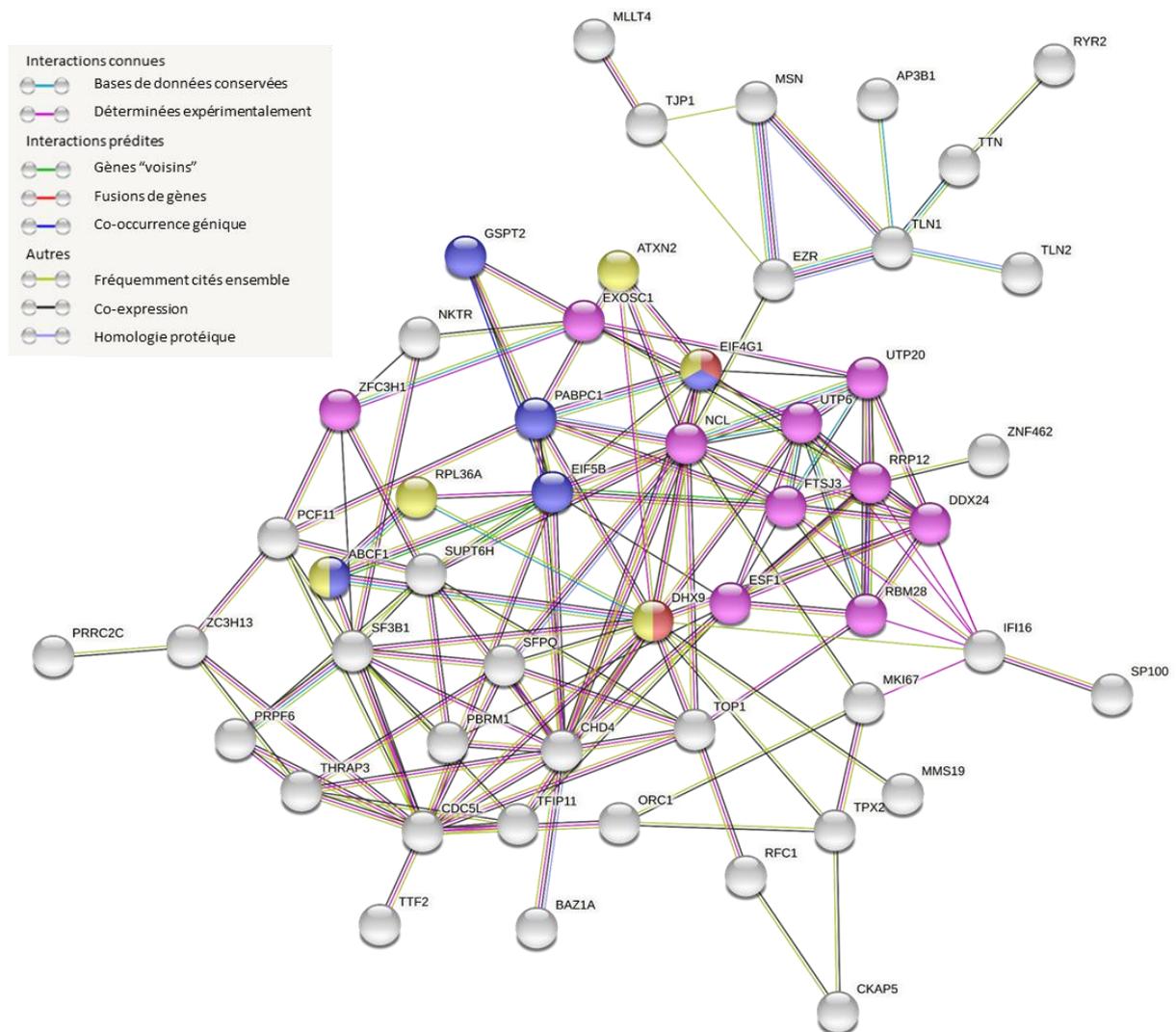
Figure 42 : Obtention de la lignée cellulaire Expi293F  $\Delta$ PUS7. (A) Principe du knockout du gène PUS7 par le kit KN2.0 non-homology-mediated CRISPR. Puro : gène de résistance à la Puromycine. P2A : séquence de clivage peptidique. (B) Sélection des cellules exprimant la GFP par cytométrie de flux trieuse. Gauche : Population cellulaires en fonction de leur taille (FSC) et de la granulométrie (SSC). Droite : Sous-groupe cellulaires issues de la population 2 en fonction de leur fluorescence. Encadré rouge : sélection des cellules exprimant la GFP. (C-D) Révélation par Western Blot des protéines Pus7 et actine. PM : échelle de poids moléculaires. (C) WT : extrait protéique de cellules WT. 1 – 13 : extrait protéique issu d'un clone cellulaire sélectionné pour sa résistance à la Puromycine et l'expression de la GFP. (D) ADNc PUS7 : extraits protéiques de cellules Expi293F $\Delta$ PUS7 dans lesquelles l'ADNc de PUS7 a été transfecté depuis 4 (J.4) ou 6 (J.6) jours.

## **B. Impact des concentrations en queuine dans le milieu de culture sur la translecture**

La modification GalQ<sub>34</sub> de l'ARNt<sub>Tyr</sub> est dépendante de la présence de queuine dans le milieu. La concentration de queuine dans le sérum humain est estimée à 1-10nM [365]. Cependant, selon les articles, les expérimentations sont réalisées en présence de queuine allant de 10 à 1000nM dans le milieu de culture pour des cellules humaines. Mon objectif était de mesurer le niveau d'incorporation des acides aminés par translecture en présence ou en absence de queuine. J'ai mesuré l'efficacité de translecture des codons UAA et UAG en fonction de la concentration en queuine dans le milieu de culture des cellules Expi293F (Figure 41). Ces cellules sont usuellement cultivées dans un milieu synthétique ne comportant pas de queuine, ce qui m'a permis de contrôler précisément l'apport de queuine aux cellules. Les résultats montrent que l'efficacité de translecture diminue inversement à la concentration en queuine dans le milieu de culture, en arrivant à un pallier à 200nM de queuine pour les deux codons stop. Ainsi, les deux conditions retenues pour la poursuite de l'étude ont été 0nM et 200nM de queuine, afin de pouvoir comparer efficacement l'impact de la présence ou absence de GalQ<sub>34</sub>.

## **C. Obtention de la lignée ΔPUS7**

La protéine Pus7 est la pseudouridine synthase responsable de la modification Ψ<sub>35</sub> de l'ARNt<sub>Tyr</sub>. Pus7 n'étant pas essentielle à la survie cellulaire, j'ai réalisé un knockout de son gène PUS7 avec le kit KN2.0 non-homology-mediated CRISPR (OriGene). Le principe de ce kit est d'insérer une cassette en début du gène PUS7 grâce au système CRISPR/Cas9 et son ARN guide (Figure 42.A). La cassette insérée contient le gène codant la GFP et le gène de résistance à la Puromycine pour la sélection de la lignée cellulaire. J'ai transfecté ce système en cellules Expi293F WT avant de les trier en fonction de l'expression de la GFP en cytométrie de flux (Figure 42.B). Les cellules Expi293F non adhérentes ne peuvent pas être cultivées à une faible densité cellulaire, j'ai donc mis au point les conditions de culture pour les rendre adhérentes. Ainsi, j'ai mis les cellules Expi293F triées pour l'expression de la GFP en culture adhérente en présence de Puromycine (10μg/ml). Cette dernière sélection m'a permis d'obtenir 13 clones pour lesquels j'ai mesuré l'expression de la protéine Pus7 par Western Blot (Figure 42.C). J'ai réalisé le Western Blot en présence d'anticorps spécifiques des protéines Pus7 ou actine (témoin positif). La protéine Pus7 n'est pas retrouvée dans certains clones, notamment le 4 qui présente la plus grande différence d'intensité entre les bandes spécifiques de Pus7 et de l'actine. J'ai donc sélectionné le clone 4 et réalisé un séquençage génomique du gène PUS7. Ce séquençage a démontré que le clone 4 était une population monoclonale, avec des copies du gène PUS7 avec une insertion de la cassette ou une délétion de 8 bases au niveau de la séquence de l'ARN guide.



	Description	Comptés parmi le réseau	Force	Significativité
<span style="color: red;">●</span>	Processus biologique (Gene Ontology) Régulation de la liaison des polysomes	2/2	2,49	0,0342
<span style="color: blue;">●</span>	Fonction moléculaire (Gene Ontology) Régulation de la traduction, liaison à un acide nucléique	5/101	1,19	0,0171
<span style="color: yellow;">●</span>	Composant cellulaire (Gene Ontology) Polysome	5/68	1,36	0,00048
<span style="color: pink;">●</span>	Cluster (STRING) Biogenèse des ribosome, activité ARN-hélicase	10/200	1,19	0,0061
<span style="color: grey;">●</span>	Autres			

Figure 43 : Représentation des réseaux génétiques/protéiques mis en évidence avec les protéines hyper-exprimées en cellules  $\Delta PUS7$  par rapport aux WT. Les protéines considérées sont celles présentant un  $\text{Log}_2\text{Foldchange} > 1,5$  d'après l'analyse du ribo-seq. Ne sont pas présentées les protéines sans interaction connues, prédites ou autres avec une autre protéine du réseau. Les traits représentent les associations protéine-protéine. Compté parmi le réseau : le premier nombre indique combien de protéines du réseau sont annotées avec un terme particulier et le second nombre indique combien de protéines ont ce terme assigné dans la base de données (indiquée entre parenthèses). Force :  $\text{Log}_{10}(\text{observé}/\text{attendu})$ . Significativité : pvalue corrigées pour les tests multiples dans chaque catégorie en utilisant la procédure de Benjamini-Hochberg. Données générées avec le logiciel STRING

J'ai choisi ce clone pour mener l'ensemble de l'étude sur l'impact de la modification  $\Psi_{35}$  sur l'incorporation de l'ARN<sub>t</sub><sub>Tyr</sub> par translecture. Pour la suite des expérimentations, j'ai remis en culture non adhérente la population Expi293F  $\Delta$ PUS7, après détachement des cellules par traitement à la trypsine et mise en culture en erlen dans du milieu synthétique.

L'étude de l'impact du knockout de PUS7 a ensuite nécessité la mise au point du sauvetage de l'expression du gène PUS7. Pour cela, j'ai cloné le gène PUS7 sans intron (ADNc) dans le plasmide pCDNA3.4 sous le contrôle du promoteur CMV. J'ai transfecté l'ADNcPUS7 dans les cellules Expi293F  $\Delta$ PUS7 et contrôlé l'expression de la protéine Pus7 par Western Blot sur différents intervalles (Figure 42.D) : J.0 : jour de transfection de l'ADNcPUS7 ; J.4 : 4<sup>ème</sup> jour post-transfection et jour de transfection du système rapporteur d'intérêt ; J.6 : 6<sup>ème</sup> jour post-transfection de l'ADNcPUS7 et jour de mesure/purification du système rapporteur d'intérêt. Alors que la protéine actine est présente dans l'ensemble des extraits cellulaires testés, la protéine Pus7 n'est pas présente à J.0, puis est présente à J.4 et J.6.

#### **D. Analyse fonctionnelle des gènes différentiellement exprimés en lignée $\Delta$ PUS7**

Après l'obtention de la lignée  $\Delta$ PUS7, j'ai réalisé un ribo-seq comparatif entre cette lignée et la WT. Un duplicat de la lignée  $\Delta$ PUS7 et un triplicat de la lignée WT ont été réalisés, avec l'obtention d'une profondeur de séquençage, une reproductibilité et un phasage adéquates à l'analyse poussée des données. Comme présenté dans l'article ci-dessus, j'ai obtenu une expression différentielle de 190 gènes, dont 63 sont hyper-exprimés ( $\log(\text{FC}) > 1,5$ ) et 127 sous-exprimés ( $< -1,5$ ). En absence de RNA-seq il n'était pas possible de conclure sur le niveau de régulation transcriptionnel ou traductionnel de ces gènes. En revanche, j'ai effectué une analyse fonctionnelle des protéines codées par ces gènes afin de comprendre le lien entre leur régulation et la protéine Pus7. Pour cela, j'ai utilisé le logiciel STRING (<https://string-db.org>). J'y ai renseigné le nom des protéines codées par les gènes hyper ou sous exprimés avant de demander une analyse des réseaux protéiques en fonction de différentes bases de données disponible dans le logiciel.

L'analyse des protéines codées par des gènes présentant un  $\text{Log}(\text{FC}) > 1,5$  a permis de mettre en évidence que les protéines codées par des gènes hyper-exprimés en condition  $\Delta$ PUS7 par rapport aux WT formaient un réseau très dense avec une émergence de protéines régulatrices du ribosome (Figure 43). Le logiciel STRING a permis de mettre en avant d'autres clusters protéiques qui concernent en grande partie des protéines de liaison à l'ARN.

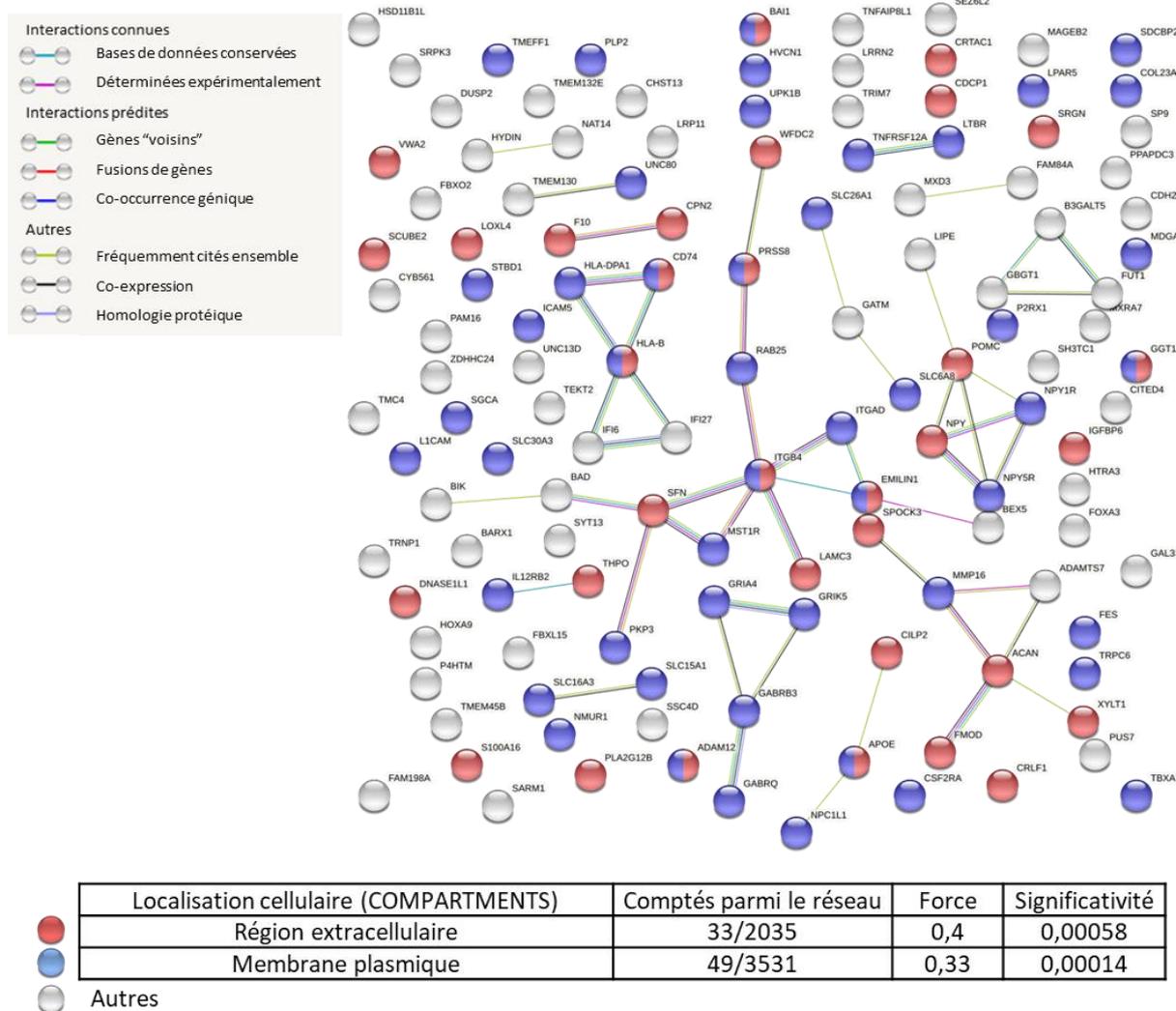


Figure 44 : Représentation des réseaux génétiques/protéiques mis en évidence avec les protéines sous-exprimées en cellules  $\Delta$ PUS7 par rapport aux WT. Les protéines considérées sont celles présentant un  $\text{Log}_2\text{Foldchange} < -1,5$  d'après l'analyse du ribo-seq. Les traits représentent les associations protéine-protéine. Compté parmi le réseau : le premier nombre indique combien de protéines du réseau sont annotées avec un terme particulier et le second nombre indique combien de protéines ont ce terme assigné dans la base de données (indiquée entre parenthèses). Force :  $\text{Log}_{10}(\text{observé}/\text{attendu})$ . Significativité : pvalue corrigées pour les tests multiples dans chaque catégorie en utilisant la procédure de Benjamini-Hochberg. Données générées avec le logiciel STRING

Sachant que la pseudouridine Pus7 n'est pas spécifique des ARNt [296], et que la modification  $\Psi$  est associée à la structure de l'ARN et aux interactions protéine-ARN [300–302], nous pouvons supposer que ces protéines codées par des gènes hyper-exprimés en absence de la protéine Pus7 ont perdu leur motif de liaison à l'ARN. En effet, la perte de la protéine Pus7 provoquerait une baisse du taux de sites  $\Psi$  des ARN cellulaires, ce qui pourrait altérer la conformation des sites de reconnaissance de protéines se liant à l'ARN. La diminution de la liaison de ces protéines à l'ARN provoquerait alors une diminution de leur activité et en réponse leur expression serait augmentée. Finalement, l'augmentation du nombre de polysomes en lignée  $\Delta$ PUS7 est sûrement le résultat d'un double phénomène entre l'augmentation de l'expression de certaines protéines régulatrices du ribosomes et la modulation de l'initiation de la traduction par le biais des dérivés d'ARNt tRF [296].

L'analyse des protéines codées par des gènes présentant un  $\text{Log}(\text{FC}) < -1,5$  a permis de mettre en évidence que les protéines codées par des gènes sous-exprimés en condition  $\Delta$ PUS7 par rapport aux WT ne formaient aucun réseau fonctionnel évident à partir des bases de données présentes dans STRING (Figure 44). En revanche, la base de données COMPARTMENTS a permis de mettre en avant deux clusters hyper-représentés parmi les 127 protéines codées par des gènes sous exprimés, avec des localisations cellulaires au niveau de la membrane plasmique ou alors du milieu extracellulaire. Ces données permettent de formuler deux hypothèses principales :

- L'adressage des protéines à la membrane ou au milieu extracellulaire est facilité en absence de la protéine Pus7, augmentant artificiellement l'activité de ces protéines et provoquant une baisse globale de l'expression de leur gène. Cette hypothèse est envisageable mais ne présente malheureusement aucune piste moléculaire dans la littérature actuelle.
- Les protéines adressées à la membrane et au milieu extracellulaire sont spécifiquement traduites par les ribosomes du réticulum endoplasmique. Si l'absence de la protéine Pus7 a un impact spécifique sur ces ribosomes, le niveau d'expression des gènes codants ces protéines pourraient diminuer fortement. Alors que Pus7 est connue pour modifier l'ARNr 5S chez la levure [366], aucune méthode d'analyse actuelle ne pourrait permettre de détecter des différences dans les profils de modifications entre l'ARNr des ribosomes cytoplasmiques et ceux du réticulum endoplasmique. Cette hypothèse reste toutefois très excitante, notamment vis-à-vis de l'émergence des études sur les ribosomes spécialisés par la modulation du schéma de modifications de l'ARNr.

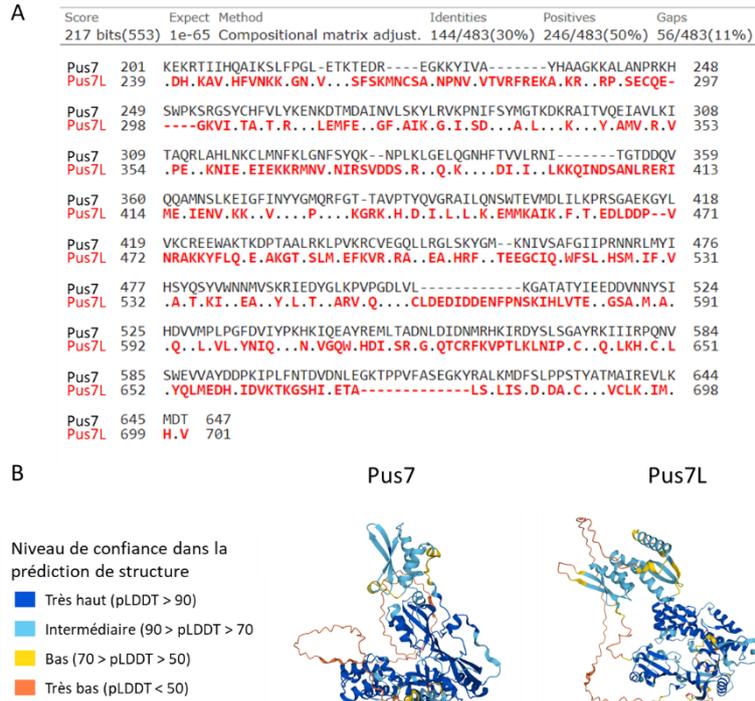


Figure 45 : La protéine Pus7L, paralogue de Pus7. (A) Alignement de la séquence protéique Pus7L avec celle de Pus7. Les points représentent des acides aminés identiques, les traits des délétions, les lettres rouges des substitutions. (B) Structures protéiques prédites sous AlphaFold. Les structures de Pus7 (gauche) et Pus7L (droite) sont colorées selon le niveau de confiance par résidu dans la prédiction de structure (pLDDT).

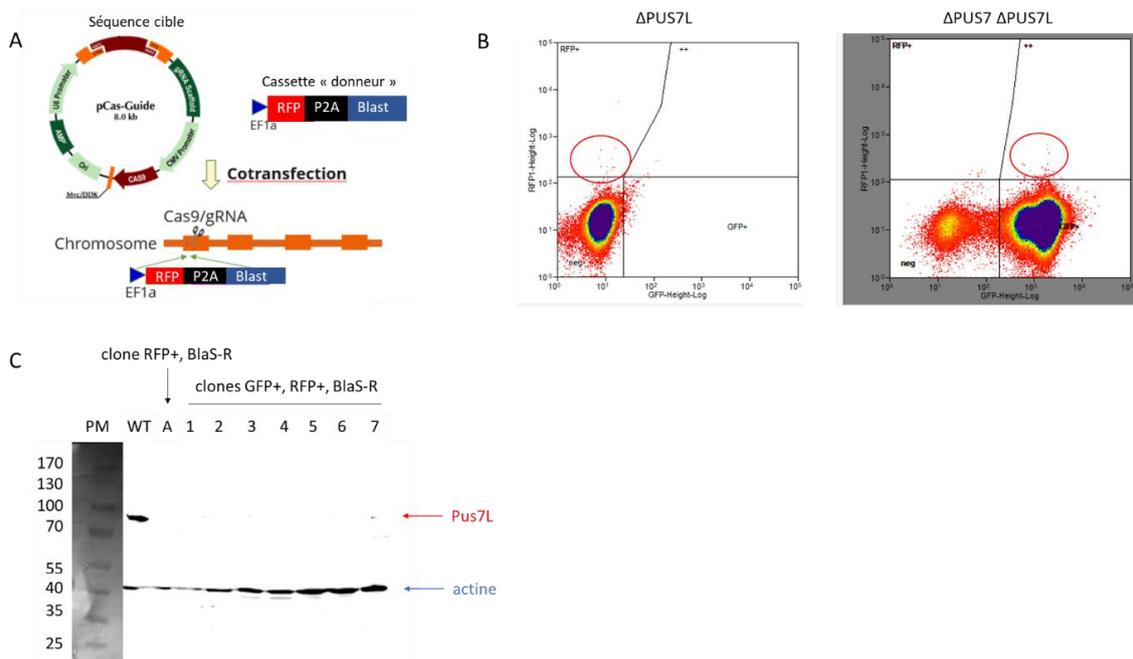


Figure 46 : Obtention des lignées cellulaires Expi293F  $\Delta$ PUS7L et  $\Delta$ PUS7L  $\Delta$ PUS7L. (A) Principe du knockout du gène PUS7L par le kit KN2.0 non-homology-mediated CRISPR. Blast : gène de résistance à la Blastidicine. P2A : séquence de clivage peptidique. (B) Sélection des cellules exprimant la RFP par cytométrie de flux trieuse. Les cellules sont représentées en fonction de leur fluorescence, avec la détection de la GFP en abscisses et la RFP en ordonnées. Gauche : Expi293F WT mutées avec le kit de délétion PUS7L. Encadré rouge : sélection des cellules exprimant la RFP. Droite : Expi293F  $\Delta$ PUS7L mutées avec le kit de délétion PUS7L. Encadré rouge : sélection des cellules exprimant la RFP et la GFP. (C) Révélation par Western Blot des protéines Pus7 et actine. PM : échelle de poids moléculaires. (C) WT : extrait protéique de cellules WT. A : Extrait protéique d'un clone cellulaire sélectionné pour être  $\Delta$ PUS7L. 1-7 : Extraits protéiques de clones cellulaires sélectionnés pour être  $\Delta$ PUS7L  $\Delta$ PUS7L.

## **E. Obtention des lignées $\Delta$ PUS7L**

D'après les bases de données il existe deux protéines Pus7 chez l'homme. La protéine Pus7 et la protéine Pus7L. Les homologies de séquences entre les deux protéines sont faibles (Figure 45.A), mais les prédictions structurales réalisées sur la séquence de la protéine Pus7L prédisent une activité de pseudouridine synthase (Figure 45.B). Il y a donc un risque de ne pas observer de phénotype dans une lignée  $\Delta$ PUS7 si la protéine Pus7L est capable de compléter d'un point de vue fonctionnel l'absence de la protéine Pus7. J'ai alors construit deux nouvelles lignées cellulaires :  $\Delta$ PUS7L et  $\Delta$ PUS7 $\Delta$ PUS7L. La lignée simple délétante  $\Delta$ PUS7L permet de déterminer le rôle isolé de la protéine Pus7, tandis que la lignée double délétante  $\Delta$ PUS7 $\Delta$ PUS7L permet d'éviter toute redondance fonctionnelle entre les deux protéines.

J'ai construit les lignées  $\Delta$ PUS7L et  $\Delta$ PUS7 $\Delta$ PUS7L à partir de cellules Expi293F WT et Expi293F  $\Delta$ PUS7 respectivement. J'ai utilisé la même stratégie de délétion que décrit précédemment avec le système CRISPR/Cas9. La lignée  $\Delta$ PUS7 exprimant déjà la cassette GFP-Puromycine, j'ai utilisé une nouvelle cassette destinée à être insérée dans le gène PUS7L, et permettant l'expression de la protéine RFP et d'un gène de résistance à la Blastidicine (Figure 46.A). J'ai transfecté le système CRISPR/Cas9 en cellules Expi293F WT (groupe 1) ou  $\Delta$ PUS7 (groupe 2) avec de cultiver les cellules de manière adhérente afin de les sélectionner pour leur résistance à la Blastidicine (30 $\mu$ g/ml). J'ai ensuite trié les cellules en fonction de l'expression de la protéine RFP en cytométrie de flux (Figure 46.B). Malgré la sélection préliminaire avec l'antibiotique, très peu de cellules exprimaient efficacement la RFP. J'ai remis en culture adhérente les cellules issues de ce dernier tri et réalisé des extraits protéiques afin de vérifier l'expression de la protéine Pus7L par Western Blot (Figure 46.C). J'ai obtenu un clone pour la lignée issue des cellules WT et 7 clones pour la lignée issue des cellules  $\Delta$ PUS7. Le Western Blot a permis de détecter la protéine Pus7L dans l'extrait des cellules WT non transfectée et dans l'extrait du clone 7 tandis que l'actine a été révélée dans l'ensemble des extraits. L'ADN génomique de deux clones, le clone unique issu de la population 1 et le clone 6 issu de la population 2, a été séquencé afin de vérifier la mutation des copies du gène PUS7L dans lesquelles la cassette ne serait pas insérée. Les deux populations présentaient des copies de PUS7L mutées avec un décalage de lecture en début de séquences. Ces deux clones n'exprimant plus la protéine Pus7L, je les ai remis en culture non adhérente en milieu synthétique et considérés comme les lignées  $\Delta$ PUS7L (population 1) et  $\Delta$ PUS7 $\Delta$ PUS7L (population 2).

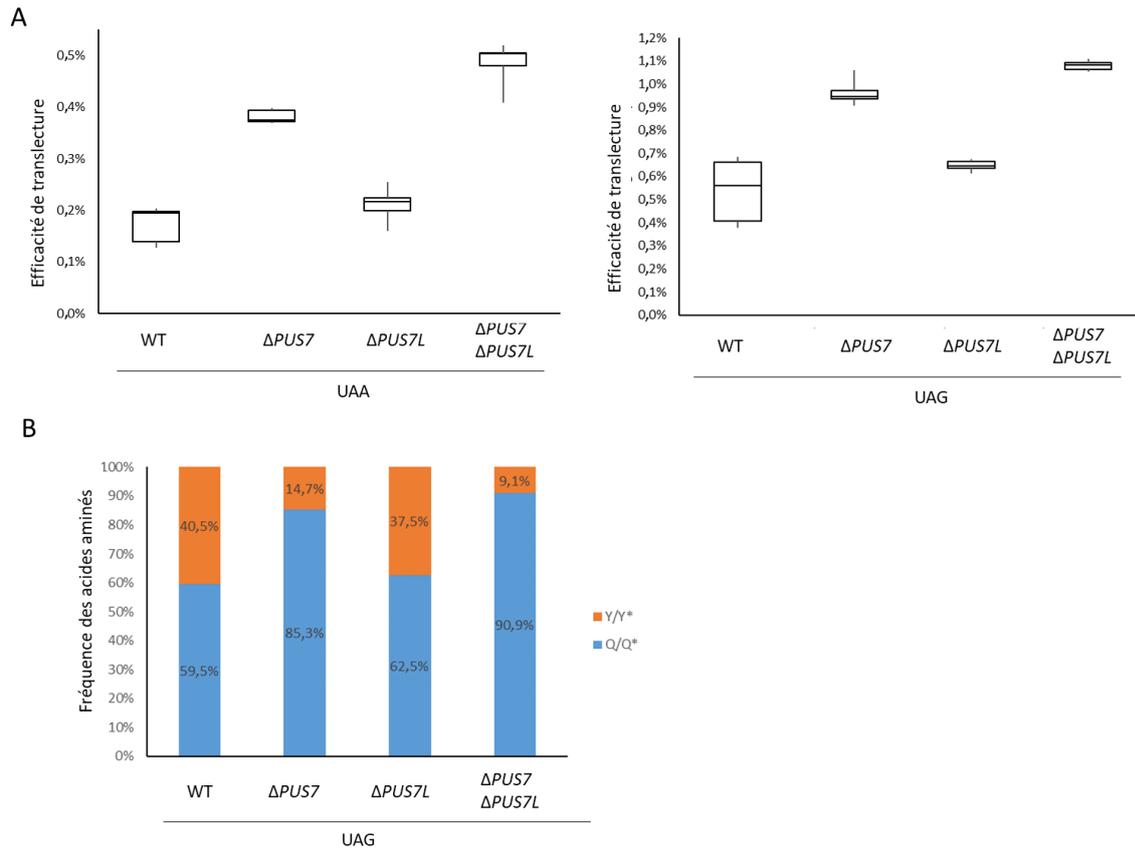


Figure 47 : Impact de la délétion de la protéine Pus7L sur la transfecture des codons UAA et UAG (A) Efficacité de lecture des codons stop UAA et UAG dans les cellules WT, les cellules  $\Delta PUS7$ , les cellules  $\Delta PUS7L$  ou les cellules  $\Delta PUS7 \Delta PUS7L$ , et cultivées avec 200nM de queuine. Les lignes au centre des boîtes indiquent les valeurs médianes ; les limites des boîtes indiquent les 25e et 75e percentiles ; les moustaches s'étendent jusqu'aux valeurs minimales et maximales. n=5 points d'échantillon. L'échelle de l'axe des ordonnées est différente entre les graphiques de gauche et de droite. (B) Fréquences relatives des acides aminés incorporés au niveau du codon stop UAG dans les cellules WT, les cellules  $\Delta PUS7$ , les cellules  $\Delta PUS7L$  ou les cellules  $\Delta PUS7 \Delta PUS7L$ , et cultivées avec 200nM de queuine. Y/Y\* : rapport entre le peptide purifié à partir de cellules humaines (Y) et le peptide AQUA correspondant (Y\*). n=3 points d'échantillon.

## F. Analyse de la translecture dans les lignées $\Delta$ PUS7L et $\Delta$ PUS7 $\Delta$ PUS7L

Dans le but de mettre en évidence une éventuelle redondance d'activité entre les protéines Pus7 et Pus7L, j'ai testé le rôle de Pus7L dans la fidélité de la traduction dans les lignées  $\Delta$ PUS7L et  $\Delta$ PUS7 $\Delta$ PUS7L. Dans un premier temps, j'ai mesuré l'impact de la délétion de Pus7L sur l'efficacité de translecture avec le système double rapporteur  $\beta$ -Galactosidase / Luciférase (Figure 47.A). Les taux de translecture des codons stop UAA et UAG étaient similaires entre les lignées WT et  $\Delta$ PUS7L, et entre les lignées  $\Delta$ PUS7 et  $\Delta$ PUS7 $\Delta$ PUS7L. Ensuite, j'ai mesuré la fréquence des acides aminés incorporés par translecture du codon UAG avec le système rapporteur GST par spectrométrie de masse (Figure 47.B). Comme pour l'efficacité de translecture, la fréquence des acides aminés incorporés est la même entre les lignées WT et  $\Delta$ PUS7L, et entre les lignées  $\Delta$ PUS7 et  $\Delta$ PUS7 $\Delta$ PUS7L. D'après l'ensemble de ces résultats, le knockout de la protéine Pus7L n'a pas d'impact sur l'incorporation de l'ARN<sub>Tyr</sub> par translecture des codons UAA et UAG. J'ai donc conclu qu'il n'y avait pas de redondance fonctionnelle entre les protéines PUS7 et PUS7L.

Afin de tenter d'identifier les sites de pseudouridinations liés à PUS7L, les ARN totaux issus des lignées  $\Delta$ PUS7 et  $\Delta$ PUS7 $\Delta$ PUS7L ont été envoyés en Hydro- $\Psi$ -séquençage sur la plateforme IBSLor (Virginie Marchand). Les résultats sont en cours.

## 3. Modification de l'ARNr

Comme je l'ai présenté dans l'introduction, les modifications 2'-O-me de l'ARNr sont connues pour jouer un rôle important dans la régulation de la traduction en particulier dans les cellules cancéreuses. Le taux de 2'-O-me du ribosome affecte notamment la traduction des ARN dont l'initiation dépend d'une IRES. En effet, il a été observé dans des cellules humaines en culture que la baisse d'expression de la protéine Fibrillarine (FBL) est associée à une diminution du taux de traduction des ARNm dont l'expression est sous contrôle de l'IRES de type II du virus de l'encéphalomyocardite (EMCV). Deux hypothèses peuvent permettre d'expliquer cette observation : (1) une diminution du recrutement du ribosome par l'IRES (le ribosome hypométhylé serait moins efficacement recruté) ; (2) une baisse de l'efficacité du ribosome à initier la traduction (le ribosome hypométhylé serait moins efficace pour quitter l'IRES). Ainsi, le but du second axe de ma thèse a été de tester ces hypothèses en caractérisant d'un point de vue cinétique l'initiation dépendante d'IRES en présence de ribosomes sauvages ou hypométhylés. Pour cela j'ai travaillé en collaboration avec Karen Perronet (ENS Paris-Saclay) afin d'adapter le système d'étude de la cinétique du ribosome *in vitro* en molécule unique (sm) par microscopie de fluorescence par réflexion totale interne (TIRF) aux ribosomes humains, l'objectif étant d'obtenir des données de cinétique en fonction du taux de méthylation du ribosome.

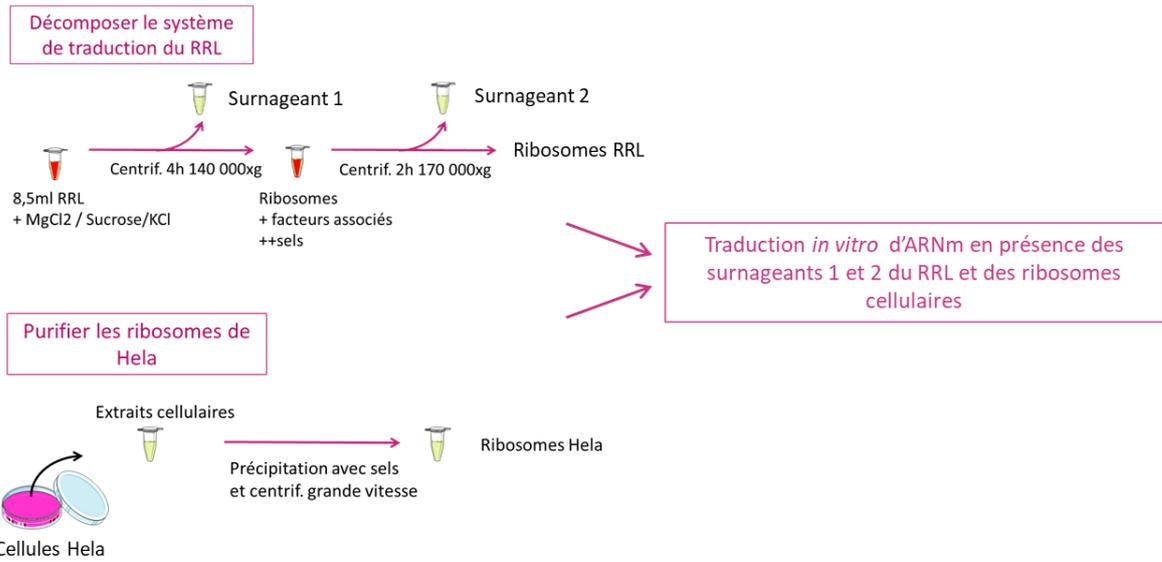


Figure 48 : Principe de la reconstitution du système de traduction eucaryote. Haut : fractionnement du système de traduction du RRL (Lysat de réticulocytes de lapin). Surnageant 1 : Fraction S140. Surnageant 2 : Salt Wash. Bas : extraction des ribosomes issus de cellules Hela.

Pour réaliser cela, mon objectif était double : 1) arriver à compléter le système de traduction *in vitro*, déplété en ribosomes de lapin, avec des ribosomes fonctionnels issus de cellules Hela dans lesquelles FBL est sous-exprimée ; 2) Analyser la cinétique de traduction d'ARNm dont l'initiation est dépendante de l'IRES EMCV, IRES qui n'avait pas été étudiée avec notre système en molécule unique.

## **A. Reconstitution du système de traduction *in vitro***

### **a. Principe de la reconstitution du système de traduction *in vitro***

Une étude de 2016 de Marianna Penzo et.al a publié le protocole nécessaire à la reconstitution d'un système de traduction « hybride », combinant l'extrait de réticulocytes de lapin (RRL) déplété en ribosome, et des ribosomes issus de cellules humaines [367] (Figure 48). Ce protocole présentait deux avantages majeurs pour mon étude. Tout d'abord, le RRL est un système de traduction très efficace qui est classiquement utilisé dans les analyses TIRF que nous menons au laboratoire. Ensuite, il fallait pouvoir obtenir des ribosomes hypométhylés. Le seul moyen d'y parvenir était de les purifier à partir de cellules humaines sous exprimant la protéine fibrillarine. Ainsi le protocole publié était parfaitement adapté à nos besoins.

Le principe de reconstitution du système de traduction *in vitro* reposait donc sur trois étapes principales :

#### 1. Le fractionnement du RRL, avec la séparation des ribosomes du reste du système de traduction

Pour fractionner le RRL, les ribosomes de lapin sont d'abord éliminés par centrifugation, ce qui permet d'obtenir un surnageant post-ribosomique (S140) à utiliser comme source de facteurs impliqués dans le contrôle de la traduction. Ensuite, les facteurs non-ribosomiaux liés aux ribosomes sont détachés par une procédure en deux étapes : les ribosomes de lapin sont remis en suspension dans un tampon à faible force ionique, puis lavés dans une solution à haute force ionique. Le surnageant résultant (surnageant 2) est composé principalement des facteurs fortement associés aux ribosomes alors que le culot contient les ribosomes.

#### 2. La sédimentation des ribosomes de cellules humaines Hela

Les extraits de cellules sont passés à travers deux phases de saccharose par centrifugation. La phase supérieure est très concentrée en sels afin de favoriser la rupture des liaisons ioniques, tandis que la phase inférieure, peu concentrée en sels, permet le lavage des ribosomes en favorisant la rupture des liaisons hydrogène, des forces hydrophobes et de Van der Waals. Le culot contient alors les ribosomes cellulaires, dépourvus de contaminants tels que les facteurs d'initiation et d'élongation, les protéines régulatrices, ainsi que les acides nucléiques (ARNm et ARNt).

### 3. La traduction d'ARNm *in vitro* en présence des ribosomes humains dans le RRL

La traduction est réalisée par addition du surnageant S140, du surnageant 2, et des ribosomes issus des cellules humaines, à l'ARNm dans un tampon appelé Master Mix, permettant notamment l'apport de sels et des acides aminés nécessaires à la traduction.

#### **b. Expérimentation de la reconstitution du système de traduction *in vitro***

La première étape reposait sur le fractionnement des différents éléments du RRL. La première difficulté rencontrée a été la quantité de RRL à disposition. Le protocole publié était basé sur l'utilisation de 10ml de RRL préparés directement à partir de sang prélevé sur des lapins. N'ayant pas de lapins à disposition, j'étais dans l'obligation de l'acheter directement chez Proméga. Avec un prix supérieur à 200 euros pour 1ml de RRL commercial, nous avons décidé de partir de 8,5 ml de RRL pour optimiser nos chances de sédimenter correctement les différentes molécules tout en diminuant les frais d'achat du RRL. Suite à l'obtention du surnageant S140, la concentration en ribosomes resuspendus mesurée au spectrophotomètre à 260nm était de 12mg/ml (le protocole publié indiquait l'obtention d'une concentration en ribosomes 20mg/ml). Cette donnée m'a permis d'adapter la quantité de surnageant 2 utilisée dans l'essai de traduction *in vitro* en considérant que la quantité des facteurs retenus avec les ribosomes après l'obtention du surnageant S140 variait proportionnellement à la concentration en ribosomes. J'ai également conservé les ribosomes présents dans le culot après obtention du surnageant 2 afin de les utiliser comme contrôle dans mes expériences de traduction. Afin de les conserver, j'ai utilisé la méthodologie indiquée pour les ribosomes cellulaires, en les resuspendant dans la solution « ribosome » et en les conservant à 3,5pmol/ $\mu$ l.

La deuxième étape consistait à sédimenter des ribosomes issus de cultures de cellules Hela. Afin de débiter les essais, j'ai utilisé une lignée sauvage Hela disponible au laboratoire. La resuspension du culot contenant les ribosomes m'a permis d'obtenir une concentration en ribosomes mesurée au spectrophotomètre à 260nm de 0,0865pmol/ $\mu$ l, (le protocole préconisait une conservation par aliquots à 5pmol/ $\mu$ l). Cette solution étant destinée à être diluée au 1/80 dans le protocole initial (0,0625pmol/ $\mu$ l), j'ai décidé de continuer mes essais avec cette préparation de ribosomes cellulaires. Cependant, la conservation à plus faible concentration pouvant diminuer l'activité des ribosomes, j'ai décidé d'adapter la quantité de ribosomes cellulaires dans le mix de traduction en ne les diluant pas. Finalement j'allais utiliser dans le mix de traduction 2 $\mu$ l à 0,0865pmol/ $\mu$ l, soit 0,173pmol de ribosomes issus des cellules Hela.

J'ai réalisé les tests de traduction *in vitro* en effectuant différents mix de traduction, tout en utilisant un ARNm codant la Firefly luciférase comme système rapporteur.

Test 1 (Tableau 1) : Incubation 1h à 30°C. Lecture de l'activité de la luciférase contenue dans les 20µl de produit de traduction avec 100µl de tampon luci.

				RLU
1	dRRL + ribosomes Hela	Master Mix 5X	4,4µl	1445
	ARNm Fluc	S140	8,8µl	
		Surnageant 2	1,5µl	
		ARNm Fluc	300ng	
		Ribosomes	0,173 pmol	
		H2O	Qsp 20µl	
2	dRRL + ribosomes RRL	Master Mix	4,4µl	20
	ARNm Fluc	S140	8,8µl	
		Surnageant 2	1,5µl	
		ARNm Fluc	300ng	
		Ribosomes	0,173 pmol	
		H2O	Qsp 20µl	
3	RRL	Mix RRL	20µl du mix : 200µl RRL promega untreated + 6µl AA (-Cys, -Leu) + 40µl H2O	6
	ARNm Fluc	ARNm Fluc	300ng	
4	RRL	Mix RRL	20µl du mix : 200µl RRL promega untreated + 6µl AA (-Cys, -Leu) + 40µl H2O	9
	H2O	H2O	0,3µl	
blanc	tampon luci			12,5

Tableau 1 : Plan d'expérimentation et résultats de l'activité luciférase du premier test de traduction avec le système de traduction reconstitué et ses contrôles. dRRL : RRL délété en ribosomes.

Lors de ce premier test, mes résultats présentent un bon niveau de traduction avec les ribosomes Hela (1) malgré la conservation à plus faible concentration des ribosomes. Malheureusement, les contrôles positifs ne fonctionnent pas, je n'obtiens aucune traduction avec les ribosomes issus du RRL (2) ou avec le RRL commercial (3). Ce résultat un peu surprenant nécessite alors d'être confirmé. Pour cela j'ai décidé de tester si l'augmentation de la quantité de ribosomes ou d'ARNm dans le mix de traduction avec les ribosomes Hela permettait d'obtenir plus de traduction.

Test 2 (Tableau 2) : Incubation 1h à 30°C. Lecture de l'activité de la luciférase contenue dans les 20µl de produit de traduction avec 100µl de tampon luci.

				RLU
1	dRRL + ribosomes Hela	Master Mix 5X	4,4µl	2909
	ARNm Fluc	S140	8,8µl	
		Surnageant 2	1,5µl	
		ARNm Fluc	600ng	
		Ribosomes	0,173 pmol	
		H2O	Qsp 20µl	
2	dRRL + ribosomes Hela	Master Mix	4,4µl	3402
	ARNm Fluc	S140	8,8µl	
		Surnageant 2	1,5µl	
		ARNm Fluc	300ng	
		Ribosomes	0,173 pmol	
		H2O	Qsp 20µl	
3	dRRL + ribosomes Hela	Master Mix	4,4µl	140
	ARNm Fluc	S140	8,8µl	
		Surnageant 2	1,5µl	
		ARNm Fluc	300ng	
		Ribosomes	0,346 pmol	
		H2O	Qsp 20µl	
4	dRRL + ribosomes RRL	Master Mix	4,4µl	149
	ARNm Fluc	S140	8,8µl	
		Surnageant 2	1,5µl	
		ARNm Fluc	300ng	
		Ribosomes	7 pmol	
		H2O	Qsp 20µl	
5	RRL	Mix RRL	20µl du mix : 200µl RRL promega untreated + 6µl AA (-Cys, -Leu) + 40µl H2O	1229628
	ARNm Fluc	ARNm Fluc	300ng	
6	RRL	Mix RRL	20µl du mix : 200µl RRL promega untreated + 6µl AA (-Cys, -Leu) + 40µl H2O	6
	H2O	H2O	0,3µl	
blanc	tampon luci			66

Tableau 2 : Plan d'expérimentation et résultats de l'activité luciférase du deuxième test de traduction avec le système de traduction reconstitué et ses contrôles. dRRL : RRL délété en ribosomes.

Au cours de cette deuxième expérimentation, l'activité luciférase obtenue avec le mix de traduction reconstitué est deux fois supérieur à celle obtenue lors de la première (2), ce qui peut être expliqué par des variabilités de pipetage ou de mesure. Lors de ce test, je tente d'obtenir davantage de traduction avec le système reconstitué. La traduction obtenue avec le système de traduction recomposé en doublant (1) ou non (2) la quantité d'ARNm est la même, ce qui indique que l'ARNm n'était pas limitant dans nos conditions de traduction. J'observe également des niveaux d'activité luciférase plus de 20 fois inférieurs en doublant la quantité de ribosomes (3), ce qui résulte sûrement d'un trop grand apport en sels dans le mix de traduction. Contrairement au premier test, j'ai obtenu ici de l'activité luciférase en présence des ribosomes issus du fractionnement du RRL (4), bien qu'elle soit 20 fois inférieure à celle observée avec les ribosomes de cellules Hela. Ce niveau de traduction peut refléter une perte d'activité des ribosomes issus du RRL. En raison de l'absence d'activité dans le premier test, j'ai apporté dans le deuxième une plus grande concentration en ribosomes (4 fois supérieure) sans savoir que le mix de traduction pouvait être également sensible à un trop grand apport de ribosomes purifiés (3), ce qui peut également expliquer la faible activité des ribosomes issus du RRL fractionné. Pour finir, la traduction réalisée avec le mix de RRL commercial classique permet d'obtenir dans ce nouveau test une activité luciférase 360 fois supérieure à celle réalisée avec le système reconstitué (5). Son contrôle négatif sans ARNm ne présentant pas d'activité luciférase, l'augmentation de l'activité luciférase observée en RRL commercial ne peut être que le résultat d'une plus forte activité de traduction qu'en système reconstitué. J'ai alors décidé de mener un troisième test afin d'identifier un potentiel élément du système reconstitué qui pourrait être à l'origine de la faible activité du système reconstitué par rapport au système commercial. Pour cela, j'ai conduit plusieurs traductions *in vitro* avec soustraction de chaque élément du mix de traduction.

Test 3 (Tableau 3) : Utilisation uniquement de ribosomes Hela. Incubation 1h à 30°C. Lecture de l'activité de la luciférase contenue dans les 20µl de produit de traduction avec 100µl de tampon luci.

	Ribosomes	S140	Salt wash fraction	ARNm Fluc	Master Mix 5 X	H2O	RLU
1	0,173 pmol	8,8µl	1,5µl	300ng	4,4µl	Qsp 20µl	1107
2	0	8,8µl	1,5µl	300ng	4,4µl		5
3	0,173 pmol	0	1,5µl	300ng	4,4µl		8
4	0,173 pmol	8,8µl	0	300ng	4,4µl		1137
5	0,173 pmol	8,8µl	1,5µl	0	4,4µl		6
6	0,173 pmol	8,8µl	1,5µl	300ng	0		6
Blanc	tampon luci						5,33

Tableau 3 : Plan d'expérimentation et résultats de l'activité luciférase du troisième test de traduction avec le système de traduction reconstitué.

Dans ce nouveau test, j'observe à nouveau une variabilité dans l'activité luciférase obtenue avec le système reconstitué, 3 fois inférieure à celle du test précédent (1), encore le résultat de variabilités expérimentales. Je teste l'impact de la suppression de chaque élément du mix de traduction du système reconstitué sur le taux d'activité luciférase obtenu. Ça me permet de mettre en évidence que contrairement à la suppression du mix de traduction des ribosomes, de la fraction S140, de l'ARNm ou du Master Mix qui engendre une absence totale de traduction (2, 4, 5, 6), la suppression de la fraction salt wash n'a aucune incidence sur l'activité luciférase obtenue (3). Ces résultats montrent que la fraction salt wash issue du fractionnement du RRL n'est pas fonctionnelle et donc que les facteurs associés aux ribosomes ne sont soit pas fonctionnels, soit absents de la fraction en question. Ce résultat est très promoteur sur la possibilité d'obtenir un taux de traduction plus important avec le système de traduction reconstitué, mais cela impose de recommencer un nouveau fractionnement du RRL.

Finalement, ces tests ont permis de mettre en évidence que le système de traduction *in vitro* que j'ai reconstitué à partir de ribosomes issus de cellules Hela et du RRL fractionné est bien fonctionnel. Malheureusement le taux de traduction est trop faible pour permettre une analyse de la cinétique de la traduction en smTIRF. Ces tests de traduction *in vitro* ont également indiqué qu'augmenter la concentration d'ARNm

ou de ribosomes dans le mix de traduction ne permettait pas d'augmenter ou bien inhibait le taux de traduction, ce qui semble indiquer que ces deux éléments n'étaient pas en quantité limitante. En revanche, même si la quantité de ribosomes n'était pas limitante, leur activité pouvait avoir été affectée au cours de l'extraction, de la sédimentation ou de la conservation. Il faut également souligner que l'absence de la fraction salt wash dans le mix de traduction n'avait aucun impact sur l'efficacité de traduction, ce qui peut indiquer que les facteurs de traduction censés être présents dans cette fraction ont perdu leur activité. Ce projet était réalisé en collaboration avec le laboratoire de Jean Jacques Diaz (CRCL, Lyon), qui a eu l'occasion d'obtenir un poste d'ingénieur d'études au moment de mes derniers résultats. Ce poste était dédié à l'optimisation de la technique de reconstitution du système de traduction *in vitro*, et nos collaborateurs avaient accès à de grandes quantités de RRL pour effectuer des mises au point de protocole. Nous avons donc décidé d'un commun accord que j'arrêtais ces expérimentations en attendant que leur équipe produise un protocole permettant d'obtenir une meilleure efficacité de traduction. Finalement, ces optimisations n'ont pas été obtenues durant ma thèse, laissant cette partie du projet à l'état de résultats préliminaires.

En parallèle de la mise au point de la reconstitution du système de traduction *in vitro*, j'ai débuté avec Karen Perronet l'analyse par microscopie TIRF de la cinétique de traduction d'un ARNm possédant l'IRES EMCV.

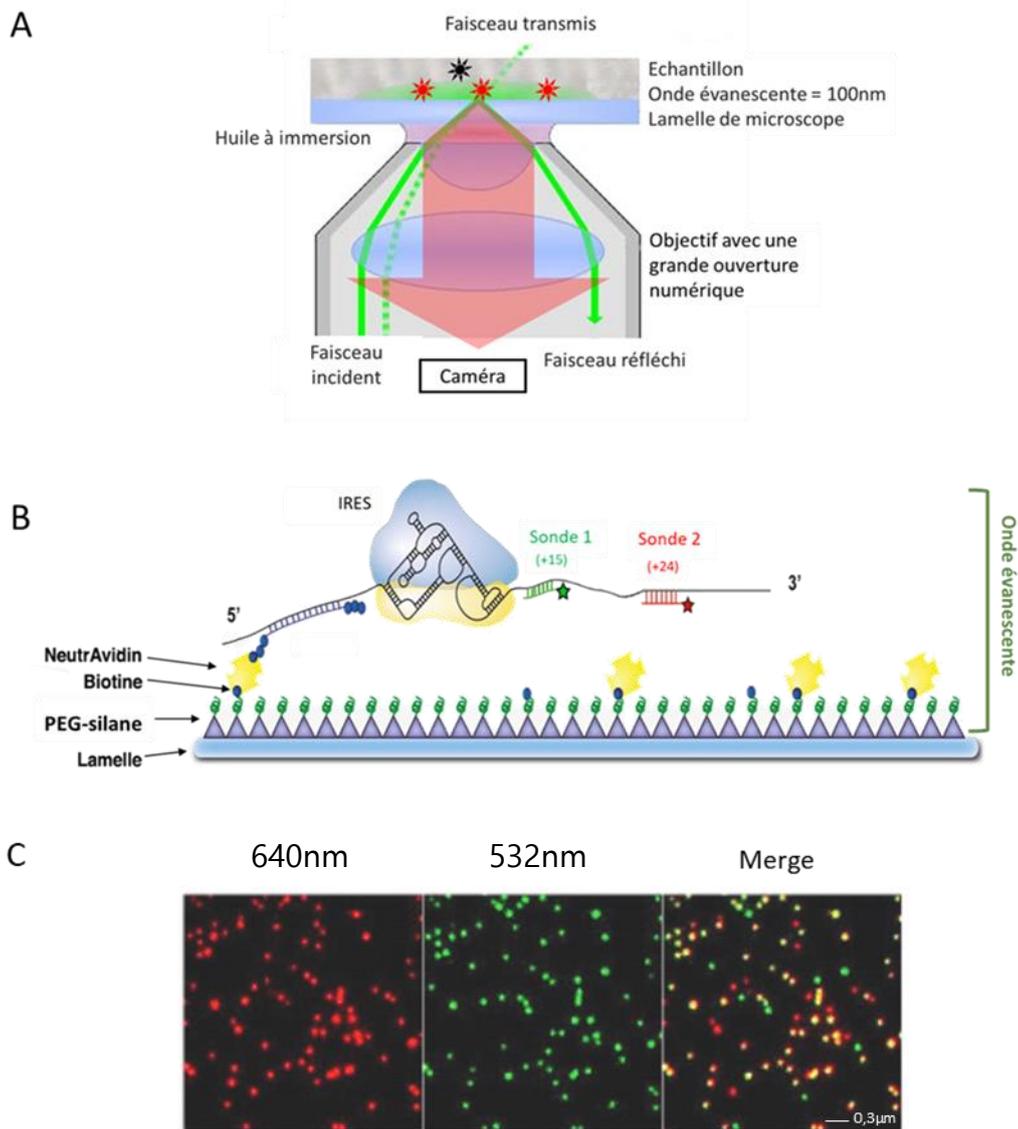


Figure 49 : Principes de la mesure de la cinétique du ribosome eucaryote en molécule unique par TIRF. (A) Principe de la mesure de fluorescence par TIRF. La microscopie TIRF implique la création d'une onde évanescente par le faisceau laser. Le fluorophore est excité (rouge) lorsqu'il est présent dans l'onde évanescente et donc visible via un objectif à grande ouverture numérique. Lorsque le fluorophore sort de l'onde évanescente, il n'est plus excité (noir) et n'est donc pas visible. (B) Modèle moléculaire de la mesure de la cinétique de ribosome en molécule unique par TIRF. Chaque ARNm est maintenu dans l'onde évanescente par couplage à un ARN biotinylé permettant sa fixation à la lamelle du microscope. Deux sondes couplées à un fluorophore spécifique sont hybridées à l'ARNm à des sites définis selon leur distance en codons du codon initiateur de l'IRES. (C) Images réalisées à partir du système de traduction en TIRF. Chaque point correspond à un ARN hybridé à une sonde couplée au fluorophore excité. L'image a été filtrée pour supprimer le bruit de fond. Rouge : sonde avec ATTO647N excité à 640nm ; vert : sonde avec ATTO565 excité à 532nm ; jaune : co-localisation sur le même ARN de deux sondes, visible par superposition de deux images prise en excitation à 640 et 532nm (Merge).

## **B. Analyse cinétique de l'IRES EMCV par smTIRF**

La traduction est un processus dynamique qui a nécessité la mise au point de différentes méthodes d'analyses par molécule unique. L'approche par molécule unique permet d'étudier des processus asynchrones et de mesurer des événements rares sans en moyenniser le comportement individuel. Depuis plus de 10 ans, mon laboratoire a participé au développement d'une analyse de la cinétique de traduction par smTIRF avec l'équipe de Karen Perronet (Figure 49) [352]. Cette méthode permet d'étudier la cinétique d'initiation et d'élongation de la traduction eucaryote en molécule unique avec une machinerie traductionnelle de mammifère non modifiée, sans nécessité de la présence d'épitopes répétés.

### **c. Principe de la mesure de la cinétique de traduction par TIRF**

La microscopie TIRF est une microscopie à fluorescence impliquant la création d'une onde évanescente par la réflexion totale d'un faisceau laser (Figure 49.A). Le faisceau entrant dans la chambre microfluidique excite les fluorophores spécifiques de sa longueur d'onde, à condition qu'ils soient présents dans l'onde évanescente. L'onde évanescente de notre système est confinée sur typiquement 100nm, afin de correspondre à l'échelle du ribosome, tout en limitant le bruit de fond généré par les fluorophores non attachés au support. Une chimie de surface a été développée pour fixer spécifiquement les ARNm étudiés sur la lamelle, afin de les maintenir dans l'onde évanescente (Figure 49.B). Cette chimie est constituée d'une couche de polyéthylène glycol (PEG), électriquement neutre, qui permet de passiver la surface. Les PEG sont couplés à une biotine, sur laquelle se fixe avec une haute affinité la NeutrAvidin. Cette protéine possède 4 sites d'interactions avec les biotines, ce qui permet d'y accrocher une sonde ARN biotinylée (ARN-RNOB), elle-même appariée à l'extrémité 5' de l'ARNm (Figure 49.B). L'ARNm contient une IRES dans sa région 5'UTR, permettant aux ribosomes d'initier la traduction. Deux oligonucléotides de 21 nucléotides couplés chacun à un fluorophore différent (ATTO647N et ATTO565) sont hybridés à l'ARNm (18 nucléotides appariés et 3 A supplémentaires afin d'éloigner le fluorophore et ne pas gêner la formation du complexe double brin). La présence des sondes dans l'onde évanescente est suivie par la prise d'images successives après excitation par les deux lasers correspondants allumés en alternance (Figure 49.C). La machinerie traductionnelle contenue dans le RRL est injectée dans la chambre microfluidique. Les ribosomes initient alors la traduction grâce à l'IRES, puis leur activité hélicase permet de décrocher la première sonde (UP) puis la deuxième sonde (DOWN) après un nombre défini de cycles d'élongation. Lorsqu'une sonde est détachée par le ribosome, elle sort de l'onde évanescente et son fluorophore n'est donc plus excité, ce qui permet de déterminer la cinétique du ribosome.

	ARN traduits (%)		T <sub>long</sub> (s)		T <sub>el</sub> (s)	
	UP	DOWN	UP	DOWN	UP	DOWN
<b>CrPV</b>	11%	8%	20 (±4)	19 (±7)	0,3 (±0,3)	0,2 (±0,1)
<b>HCV</b>	30%	28%	33 (±9)	29 (±5)	0,3 (±0,2)	0,5 (±0,1)
<b>EMCV</b>	38%	31%	11,4 (±2,3)	8,8 (±1,1)	0,1	0,1

Tableau 4 : Tableau récapitulatif des données de cinétique de traduction *in vitro* obtenues en microscopie TIRF.

#### **d. Modèle mathématique de la cinétique du ribosome**

Le modèle mathématique repose sur la mesure indépendante du temps de départ de chaque sonde de l'onde évanescente à partir du moment  $T_{(0)}$  correspondant à l'injection du RRL dans la chambre microfluidique. Des images sont réalisées toutes les 2,5 secondes dans chaque longueur d'onde. Chaque point lumineux isolé correspond à une seule molécule d'ARN. Un point lumineux est considéré comme étant réellement un ARNm à partir du moment qu'il co-localise exactement dans les deux longueurs d'onde d'excitation. Le temps du ribosome à effectuer un cycle d'élongation ( $T_{el}$ ) est déterminé par le nombre d'images entre la fin de l'initiation et la disparition du point, multiplié par le temps passé par image. Le modèle considère que l'activité hélicase permet le détachement d'une sonde à partir du détachement de 15 nucléotides (sur 21 que contient la sonde). Ainsi, les sondes en position +15 et +24 du codon initiateur de l'IRES sont détachées par le ribosome après 15 et 24 cycles d'élongation respectivement. Le modèle mathématique permet également de calculer la proportion d'ARNm traduits sur l'ensemble des ARNm marqués (nommée A). Il est évalué individuellement pour chaque sonde fluorescente et caractérise l'efficacité de reconnaissance de l'ARNm par le ribosome et son efficacité de traduction, deux événements indiscernables par la méthode utilisée. Le modèle mathématique prend aussi en compte le photoblanchiment (perte de fluorescence suite à une excitation) des fluorophores et le départ spontané des sondes fluorescentes non spécifique du passage du ribosome. Pour cela, les valeurs de disparition d'un fluorophore dans des conditions de traduction sont rapportées à celles obtenues avec le témoin négatif, un ARNm ne contenant pas d'IRES ( $\Delta$ IRES) ne pouvant donc pas être traduit par un ribosome. Après ajustement des valeurs obtenues avec le témoin négatif, on obtient la probabilité de départ d'une sonde à moment donné ( $D(\text{trad})$ ).

Des mesures ont été réalisées et publiées avant mon arrivée au laboratoire à partir de deux IRES différentes, CrPV et HCV (Tableau 4). Le modèle mathématique obtenu à partir de ces données de traduction a permis de mettre en évidence la mesure d'un temps long ( $T_{\text{long}}$ ) et d'un temps court, qui correspondent respectivement au mouvement du ribosome en train de quitter l'IRES (mais en y étant toujours associé) et à la phase d'élongation jusqu'au détachement de la sonde fluorescente de l'ARNm. Le temps long dépend de la nature de l'IRES alors que le temps court dépend de la position de la sonde sur l'ARNm.

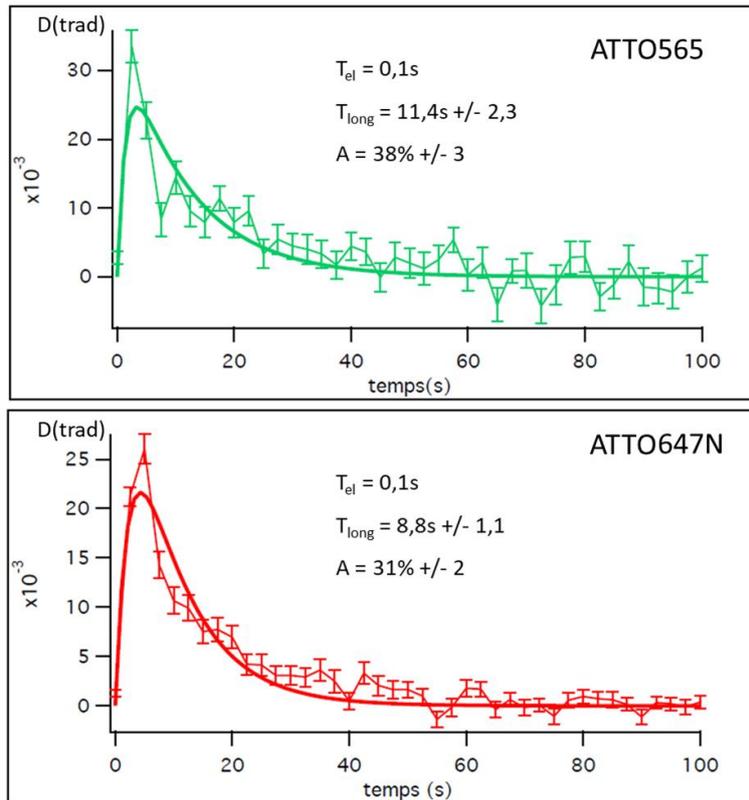


Figure 50 : Probabilité de départ des sondes UP (ATTO 565) ou DOWN (ATTO647N) à un moment donné (courbe avec barres d'erreurs). Courbe continue : courbe avec les valeurs ajustées à l'aide de la fonction :  $Dfit(t) = \frac{A}{t_0 - t_1} (e^{-(t-t_0)} - e^{-(t-t_1)})$

### **e. Mesure de la cinétique de la traduction avec une initiation dépendante de l'IRES EMCV**

A mon arrivée au laboratoire j'ai construit un ARNm possédant l'IRES EMCV (ARNm-EMCV), afin de pouvoir étudier le comportement des ribosomes hypométhylés sur cette IRES. Nous avons effectué le même nombre de mesures de cinétique de traduction en présence de l'ARNm-EMCV ou en présence de l'ARNm témoin  $\Delta$ IRES par smTIRF. Le logiciel de traitement d'images nous a permis d'exploiter environ 3000 points dans le canal d'excitation à 532nm, et 5500 points dans le canal d'excitation à 640nm avec chaque ARNm. Les données ont été traitées à partir du modèle mathématique afin d'obtenir les temps d'initiation et le temps par cycle d'élongation, ainsi que le pourcentage d'ARNm traduits sur l'ensemble des ARNm marqués (Figure 50). Ces résultats mettent en évidence que l'initiation dépendante de l'IRES EMCV est plus efficace d'un point de vue cinétique que celles des IRES CrPV et HCV (Tableau 4). En effet, le  $T_{long}$  calculé est d'environ 10 secondes alors qu'il est deux à trois fois plus long avec les IRES CrPV et HCV. Avec une vitesse d'initiation aussi rapide, le système de prises d'images est devenu très limitant. En effet, avec une image prise toutes les 2,5 secondes, l'initiation dépendante d'EMCV est trop rapide pour obtenir des données fiables sur les temps d'élongation avec une estimation finale autour de 0,1 secondes/cycle, un  $T_{el}$  équivalent à ceux observés avec les autres ARNm. Afin de corriger ce problème, deux solutions sont possibles : éloigner les sondes fluorescentes de l'IRES EMCV, ou diminuer le temps entre chaque image. Nous avons retenu la deuxième solution qui présente l'avantage de ne pas avoir à modifier notre système biologique. Nous avons ainsi débuté des tests de cinétiques en réduisant le temps entre chaque image de 2,5 à 1,4 secondes. Malheureusement, l'augmentation de l'excitation des fluorophores a provoqué une forte augmentation de leur photoblanchiment ne permettant plus une exploitation fiable des données.

Au cours de cette phase de mise au point, nous avons dû faire face à différents problèmes qui ne nous ont pas permis de continuer les expérimentations. Il y a tout d'abord eu la fermeture prolongée du laboratoire de microscopie lors du confinement lié à la COVID19 en 2020, qui a été suivie par le déménagement de ce même laboratoire sur un nouveau site de recherche. A cela se sont additionnés de nombreux soucis techniques qui ont nécessité le rachat de l'objectif de microscope et des deux lasers. Les données de cinétiques de traduction avec un ARNm-EMCV obtenues durant ma thèse sont ainsi restées des données préliminaires. Elles ont toutefois permis de mettre en évidence différents points :

- L'initiation dépendante de l'IRES EMCV est plus efficace que celle dépendante des IRES HCV ou CrPV, avec un pourcentage d'ARNm traduits plus importants et une cinétique d'initiation plus rapide.

- La technique d'analyse de la cinétique de traduction par TIRF permet de calculer des temps d'initiation très rapides.
- Le photoblanchiment induit des limites de détection temporelle, et il est nécessaire d'adapter la puissance des lasers à des cinétiques de traduction rapides.

Afin de continuer les mesures de cinétique de la traduction dépendante de l'IRES EMCV, il faut donc mettre au point le système optique nécessaire à la prise d'images plus fréquentes afin de ne plus être limités par le photoblanchiment. D'un point de vue biologique, il serait très intéressant de comprendre les mécanismes moléculaires à l'origine d'une initiation plus efficace en présence de l'IRES EMCV que des IRES HCV et CrPV. L'IRES EMCV est connue pour recruter différents facteurs de traduction tel que le facteur d'initiation eIF4A [368]. Une possibilité serait donc que le recrutement d'eIF4A permet d'augmenter l'efficacité et notamment la vitesse d'initiation au niveau de l'IRES EMCV. Afin de tester cette hypothèse, il serait possible de mesurer la cinétique d'initiation en présence de la molécule Hippuristanol, une molécule antivirale issue du corail *Isis hippuris* et inhibitrice du facteur eIF4A [369].

Au cours de ma thèse j'ai ainsi obtenu un ensemble de données préliminaires pour le projet portant sur le rôle des modifications 2'-O-me de l'ARNr dans la cinétique de traduction eucaryote. Il est encore indispensable à l'heure actuelle de mettre au point le protocole de reconstitution du système de traduction *in vitro* afin de pouvoir tester la cinétique des ribosomes humains en fonction de leur taux de méthylation en smTIRF. Il faut également obtenir des données de cinétique de traduction de l'ARNm-EMCV plus fiables en augmentant la fenêtre de détection temporelle. La 2'-O-me de l'ARNr est aujourd'hui au cœur de nombreuses études dans le cadre de cancers. L'étude de l'impact l'hypométhylation du ribosome sur la cinétique de traduction permettrait de mieux comprendre les mécanismes impliqués dans le développement des cancers. Je pense notamment qu'à plus long terme, l'étude de l'impact de sites de 2'-O-me spécifiques par ciblage de leur snoRNA pourrait permettre de mettre en avant de nouvelles cibles thérapeutiques dans le traitement des cancers. Cette hypothèse est notamment soutenue par une toute nouvelle étude de Martin D. Jansson et al/ qui a utilisé la méthode de RiboMeth-seq pour démontrer que la variation du niveau de 2'-O-me de l'ARNr représente une source considérable d'hétérogénéité des ribosomes dans les cellules humaines [192]. Ils ont notamment mis en évidence que les niveaux de modification à des sites spécifiques peuvent changer dynamiquement en réponse à des voies de signalisation en amont et que la perte d'un seul site de 2'-O-me pouvait entraîner une modulation de l'expression traductionnelle de certains ARNm. Ceci constitue une nouvelle preuve de l'importance des modifications de l'ARNr dans la régulation de la traduction et de l'importance de comprendre leur mécanisme dans le cadre, par exemple, du cancer.

## VI. Discussion générale

Au cours de ma thèse je me suis intéressée aux modifications des ARNt et ARNr et à leurs rôles dans la fidélité de la traduction.

Mon travail sur les modifications d'ARNt s'est concentré sur deux modifications de l'anticodon de l'ARN<sub>Tyr</sub>, supprimeur naturel des codons stop UAA et UAG. J'ai pu mettre en évidence que la modification GalQ<sub>34</sub> améliore la fidélité de la traduction en diminuant la capacité de l'ARN<sub>Tyr</sub> à être incorporé par translecture des codons stop UAA et UAG. Ces résultats soulèvent de nombreuses interrogations, d'autant plus qu'il s'agit de la seule modification cellulaire avec le ManQ<sub>34</sub> de l'ARN<sub>Asp</sub> à présenter l'ajout d'un sucre sur un ARNt. De nos jours, aucune étude n'a encore révélé la pertinence biologique de ces modifications propres au domaine animal. L'ajout d'un sucre est la modification post-traductionnelle protéique la plus courante [370], avec 50% de protéines modifiées par l'ajout de chaînes polysaccharidiques chez l'Homme. La glycosylation des protéines joue différents rôles cellulaires, notamment dans la structure protéique et l'adressage à la surface cellulaire [371]. En ce qui concerne la glycosylation des ARN, il a été démontré récemment que de petits ARN non codant nommés « glycoRNAs » sont modifiés par ajout de glycanes, sont localisés à la surface cellulaire et jouent un rôle dans la réponse immunitaire chez les mammifères [372]. Alors que mes résultats démontrent une baisse d'incorporation de l'ARN<sub>Tyr</sub> par translecture des codons stop UAA et UAG, il serait intéressant de vérifier la localisation cellulaire de cet ARN<sub>Tyr</sub>(GalQ<sub>34</sub>). Si l'on regarde l'ensemble des méthodes de localisation cellulaire actuelles, la localisation d'un ARNt par sonde nucléique est défavorisée par sa structure 3D ; l'immunolocalisation nécessiterait la production d'un anticorps spécifique de l'ARN<sub>Tyr</sub> alors que cela n'a jamais été réalisé ; le marquage puis transfection d'un ARNt de levure n'est pas non plus envisageable car les modifications GalQ<sub>34</sub> et ManQ<sub>34</sub> sont spécifiques des ARNt des animaux. Une dernière méthode publiée en 2017 pourrait être envisagée avec la production d'un ARNt recombinant. Dans cette étude réalisée chez *E.coli*, un aptamère nommé Spinach a été ajouté dans la boucle variable de l'ARN<sub>Tyr</sub> et permet de localiser l'ARNt par ajout de 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI) dans le milieu de culture. Le DFHBI est un ligand pouvant traverser la membrane cellulaire et est non toxique, permettant d'activer la fluorescence de Spinach similaire à celle de la GFP. Ces premiers résultats prometteurs pourraient être transposés aux cellules humaines, d'autant plus qu'une étude a déjà utilisé la sur-expression d'ARN<sub>Tyr</sub> à partir d'un plasmide pour étudier le rôle de cet ARNt dans la translecture [363]. En plus de la localisation cellulaire de ces ARNt, il serait également intéressant de mener une étude par microscopie cryo-électronique ou cristallographie par rayons X afin de visualiser l'accommodation d'un ARNt glycosylé dans les différents sites du ribosome, sous l'hypothèse qu'il soit

toujours dans la capacité d'être adressé puis d'entrer dans le ribosome. Afin de déterminer le rôle du GalQ34 de l'ARNt<sub>Tyr</sub>, il serait également nécessaire d'identifier l'enzyme Galactosyl transférase. Alors que la mannosyltransférase responsable de la catalyse de ManQ a été isolée chez le rat, aucune galactosyltransférase n'a encore été identifiée pour GalQ [288]. Une étude réalisée en cellules humaines a retrouvé trois formes d'ARNt<sub>Tyr</sub> cytosolique, G<sub>34</sub>, Q<sub>34</sub> et GalQ<sub>34</sub>, alors que l'ARNt<sub>Asp</sub> a toujours été trouvé sous formes G<sub>34</sub> ou ManQ<sub>34</sub> [289]. Dans mon étude je me suis focalisée sur la présence ou l'absence du groupement total GalQ, générés par modulation de la concentration en queuine dans le milieu. En plus des effets non spécifiques générés par l'absence de queuine, je n'ai pas pu élucider le rôle spécifique de l'ajout du sucre sur la queuosine dans la reconnaissance codon-anticodon. Identifier et cibler l'enzyme responsable de l'ajout du galactose semble d'autant plus intéressant qu'une étude a démontré que chez la drosophile, la queuosine présente sans sucre sur l'ARNt augmente la capacité de mis-incorporation [294], ce qui va dans le sens contraire à mes résultats sur GalQ chez l'Homme.

Mes travaux se sont également portés sur la  $\Psi_{35}$  de l'ARNt<sub>Tyr</sub>. J'ai pu mettre en évidence que cette modification augmentait la capacité d'incorporation de l'ARNt<sub>Tyr</sub> au niveau des codons stop UAA et UAG, mais également au niveau de ses codons sens. L'étude de la délétion de la protéine Pus7 a également permis de mettre en évidence par ribosome profiling que le ribosome était beaucoup moins présent sur le codon sens CCG de l'ARNt<sub>Arg</sub> (Log2FC -0,5 padj  $2.10^{-4}$ ), suggérant une incorporation plus efficace cinétiquement de cet ARNt dans le ribosome. L'ARNt<sub>Arg</sub> est connu pour être modifié en position 50, soit dans le bras T, par Pus7 [308]. Les modifications présentes dans cette partie de l'ARNt sont généralement impliquées dans un rôle de structure. Une plus grande stabilité de l'ARNt<sub>Arg</sub>(CCG) en présence du  $\Psi_{50}$  pourrait provoquer une augmentation de la population de cet ARNt et ainsi augmenter la vitesse de traduction de son codon sens. Cette hypothèse pourrait être vérifiée par FL-PCR, une méthode quantitative de RT-PCR mise au point spécifiquement pour les ARNt [373]. Cette modification pourrait également augmenter la capacité d'accommodation de cet ARNt dans le ribosome. Pour cela il serait intéressant de comparer l'accommodation d'un ARNt<sub>Arg</sub>(CCG) avec ou sans la modification dans le site A du ribosome par cryo-microscopie électronique ou cristallographie par rayons X. Au cours de mon étude de la modification  $\Psi_{35}$ , je me suis également intéressée à l'enzyme Pus7L orthologue de Pus7, suggérant une activité pseudouridine synthase similaire. Ainsi j'ai généré des lignées cellulaires humaines exprimant ou non les protéines Pus7 et Pus7L avant de caractériser l'effet de la délétion de leur gène respectif sur la translecture. D'après mes résultats, la délétion du gène PUS7L n'a aucun impact sur la translecture des codons stop UAA et UAG par l'ARNt<sub>Tyr</sub>, contrairement à la délétion du gène PUS7, suggérant

que la protéine Pus7L ne modifie pas l'ARNt<sub>Tyr</sub> au niveau de la position 35. Dans les années à venir, cette étude ouvre la porte à une nouvelle série d'expérience afin de caractériser le rôle de cette potentielle pseudouridine synthase propre à l'Homme. Alors qu'un hydro-Ψ-seq est en cours sur les ARNt et ARNr issues de cellules délétées pour le gène PUS7L, ses résultats pourraient mettre en évidence des sites Ψ spécifiques de cette enzyme. Il serait alors nécessaire de compléter l'étude par une approche iCLIP qui permettrait de confirmer les cibles ARNt/ARNr de cette enzyme, de mettre potentiellement en évidence d'autres cibles issues d'autres espèces ARN, mais également de trouver le motif de liaison de cette protéine à l'ARN.

Mon étude s'est focalisée sur deux modifications de l'anticodon de l'ARNt<sub>Tyr</sub> et leur rôle dans la fidélité de la traduction eucaryote. Si l'on s'intéresse à la boucle anticodon des ARNt humains, de nombreuses autres modifications pourraient être étudiées dans les années à venir telles que les Ψ en position 38 et 39, l'inosine en positions 34 et 37, les méthylations de l'uridine en position 34 de type mcm<sup>5</sup>U et les modifications Q et ManQ en position 34. De nombreux obstacles expliquent le manque de données à ce jour en cellules humaines : les difficultés de manipulation génétique dans les lignées cellulaires humaines, l'identité manquante des enzymes responsables de ces modifications ou bien le caractère essentiel de ces enzymes à la survie cellulaire. L'étude des modifications d'ARNt chez l'Homme, est rendue difficile par l'absence de cartographie complète de l'ensemble des modifications des ARNt humains. Alors que ces cartes sont déjà à disposition pour des organismes tels que *S. cerevisiae* et *E. coli*, aucune n'a encore été réalisée pour les ARNt humains [257]. Heureusement, le développement de technologies de type spectrométrie de masse devrait dans les prochaines années permettre l'émergence de nouvelles données sur les modifications d'ARNt chez l'Homme, permettant de répondre à des problématiques spécifiques comme la fidélité de la traduction, mais également plus globales avec l'importance des modifications d'ARN dans les maladies génétiques [258].

Le second axe de ma thèse reposait sur l'étude de la 2'-O-me de l'ARNr dans la cinétique du ribosome humain. Malgré l'absence de résultats concluants, mon étude a permis d'apporter de nouvelles mises au point pour le protocole d'analyse de cinétique du ribosome par sm-TIRF. La traduction est un processus dynamique et différents laboratoires ont déjà mis au point des méthodes d'analyses par molécule unique afin d'étudier des processus asynchrones et de mesurer des événements rares sans en moyenniser le comportement individuel. Pour la traduction, on peut citer de manière non exhaustive l'analyse par Förster Resonance Energy Transfer (FRET), permettant une analyse fine de la dynamique des différents facteurs de la traduction, mais nécessitant un marquage direct de ces facteurs sans garantie de ne pas altérer le processus étudié [374] ; et l'analyse de la cinétique par mesure de la fluorescence de fluorophores

couplés à des anticorps qui reconnaissent des épitopes apparaissant ou disparaissant au cours de la traduction, comme la méthode d'analyse par apparition des épitopes SunTag et MoonTag [375]. L'inconvénient principal de ces méthodes est la limite de détection : de multiples épitopes sur un même ARN sont nécessaires afin d'obtenir un signal détectable. En dehors de la molécule unique, on retrouve également la mesure de la cinétique de traduction par ribo-seq [376–378], permettant de mesurer l'efficacité et la vitesse d'initiation et d'élongation dans une population globale de ribosomes. Le développement de la technique par sm-TIRF pourrait permettre dans un futur proche d'affiner la spécificité de la mesure de la cinétique des ribosomes. D'autre part, l'application d'une nouvelle technologie comme la pince magnétique, actuellement mise au point pour la mesure de la cinétique de transcription d'ADN accroché à des billes magnétiques, à la mesure de la cinétique des ribosomes en cours de traduction pourrait permettre d'élucider de nouvelles interrogations, comme l'impact des structures d'ARNm dans la cinétique de traduction. Cette technologie serait d'autant plus importante dans la suite de mon étude, que mes résultats préliminaires semblent mettre en évidence un impact de la nature de l'IRES sur la vitesse et/ou l'efficacité d'initiation. Bien que dans l'étude des IRES EMCV, HCV et crPV l'intervention du facteur eIF4A dans l'initiation dépendante de l'IRES EMCV pourrait suffire à expliquer une initiation plus rapide, ce qui sera testé prochainement par l'utilisation d'hippuristanol dans le tampon de traduction, ces IRES diffèrent également de par leurs structures. L'étude des IRES virales représente une réelle opportunité d'améliorer nos connaissances dans l'infection des cellules humaines par des virus, sujet plus que d'actualité depuis le début de la décennie 2020.

## VII. Protocoles

### 1. Transcription in vitro

#### Digestion de l'ADN

Plasmide MU	10µg
Tampon vert 10X	4µl
Enzyme Eco 53KI	2µl
H2O	Qsp 40µl

Le mix de digestion est incubé 2h à 37°C avant d'ajouter 2µl d'enzyme de digestion supplémentaires et d'incuber à nouveau 2h à 37°C.

#### Transcription de l'ARN

SP6 Transcription 5X Buffer	20µl
rNTPs (25mM ATP, CTP, GTP, UTP)	20µl
Plasmide digéré	40µl
Enzyme mix SP6	10µl
H2O	10µl

Le mix de transcription est incubé 3h à 37°C. Ensuite 10µl de RQ1 RNase Free DNase à 1U/µl sont ajoutés avant d'incuber 15min à 37°C.

L'ARN est alors passé sur colonnes Microspin G25 pour se débarrasser des rNTPS : les colonnes sont resuspendues en centrifugeant rapidement, le bouchon puis le bout du bas sont détachés avant de mettre la colonne dans un tube Eppendorf de 1,5ml. La colonne est centrifugée une première fois 1min 1000 RCF avant d'être placée dans un nouveau tube. On charge l'ARN à raison de 500µl/colonne avant de centrifuger 4min 1000 RCF. L'ARN est dans le tube.

### 2. Transfection du système CRISPR Cas9

#### Jour 0

Une culture d'Expi293F est repiquée à  $2 \cdot 10^6$  cellules/ml afin d'être en phase de division cellulaire le jour de la transfection.

#### Jour 1

Une culture de 2,5ml de cellules Expi293F est préparée à une concentration de  $1 \cdot 10^6$  cellules/ml. En parallèle 1µg du vecteur d'ARN guide est dilué dans 150µl OPTIMEM et le mélange est vortexé avant d'y ajouter 1µg de vecteur donneur et d'être

vortexé à nouveau. On dilue 8µl d'expifectamine dans 150µl OPTIMEM avant de vortexer. Les différents mix sont incubés 5min à température ambiante avant d'ajouter le mix ADN dans le mix Expifectamine. Ce dernier mélange est incubé 25min à température ambiante avant d'être ajouté à la culture cellulaire. Les cellules sont cultivées à 37°C, 8%CO<sub>2</sub>, 120RPM.

#### Jour 2

Les enhanceurs de transfection du kit Expifectamine sont ajoutés 11h après transfection : 15µl enhanceur 1 et 150µl enhanceur 2 par erlen.

### **3. Transfection des plasmides des systèmes rapporteurs**

#### Jour 0

Une culture d'Expi293F est repiquée à 2.10<sup>6</sup> cellules/ml afin d'être en phase de division cellulaire le jour de la transfection.

#### Jour 1

Une culture de 204ml de cellules Expi293F est préparée à une concentration de 3.10<sup>6</sup> cellules/ml. En parallèle 240µg du plasmide sont dilués dans 12ml d'Optimem et 640µl d'Expifectamine sont dilués dans 12ml d'Optimem. Les mélanges sont homogénéisés à la pipette avant d'incuber 5 min à température ambiante. Le mix ADN est ajouté dans le mix Expifectamine avant une incubation de 25min à température ambiante. Le mix de transfection est alors ajouté dans l'erlen de culture Les cellules sont cultivées à 37°C, 8%CO<sub>2</sub>, 120RPM.

#### Jour 2

Les enhanceurs de transfection du kit Expifectamine sont ajoutés 11h après transfection : 120µl enhanceur 1 et 12ml enhanceur 2 par erlen.

### **4. Purification GST**

Tampons :

<b>Tampon de cassage</b>			
DTT	Triton 100X	PBS 1X	Antiprotéase 25X (dernier moment)
1mM	1%	Qsp 500ml	1ml dans 25ml

Filtration à 0,22µm avant l'ajout de l'antiprotéase, conservation à 4°C.

<b>Tampon de lavage</b>	
DTT	PBS 1X
1mM	Qsp 500ml

Filtration à 0,22µm, dégazage (ne pas remuer la bouteille), conservation à 4°C.

<b>Tampon d'éluion</b>	
Gluthation (10mM)	Tris HCl pH8 50mM
0,307g	100ml

Filtration à 0,22µm, dégazage (ne pas remuer la bouteille), conservation à 4°C.

Protocole :

Une culture de 200ml de cellules Expi293F 3 jours post-transfection est centrifugée dans 4 tubes de 50ml 5 min à 300 RCF, 4°C. Les culots sont rincés au PBS avant d'être repris chacun dans 1ml de tampon de cassage par pipetage. Les cellules sont alors réparties dans des tubes de 2ml (sur glace) en présence de billes avant d'être vortexées 30min à 4°C. Les cellules sont centrifugées 5 min à 3300 RCF, le surnageant est récupéré et centrifugé 30min à vitesse max. Le surnageant est ensuite filtré sur 0,45µm et le volume final est ajusté afin de correspondre au volume de la boucle de l'Akta (10ml). L'extrait est passé sur colonne de glutathion avec les tampons de lavage/élution indiqués ci-dessus.

## 5. Western Blot

### Jour 1

Les protéines sont dénaturées par ajout de Laemmli et incubation 3 minutes à 95°C. Ensuite, 10µg de protéines dénaturées sont déposées sur un gel de polyacrylamide à gradient 4-12% (SDS-PAGE) et migrent à 50V pendant 1h puis 100V jusqu'à ce que le bleu soit sorti du gel. Le gel est transféré sur membrane de nitrocellulose pendant 1H (30V). La membrane est saturée avec du PBS- lait 5% 1h sous agitation. La membrane est incubée avec l'anticorps primaire une nuit à 4°C sous agitation.

### Jour 2

La membrane est lavée par deux lavages de 10min au PBS, deux lavages de 10min au PBS-tween 0,1% et un lavage de 15min PBS. Elle est saturée pendant 10min dans du PBS lait 5%, avant d'être incubée 1H avec l'anticorps secondaire HRP-conjugate. La membrane est lavée par un lavage de 15min au PBS, deux lavages de 15min au PBS-tween 0,1% et deux lavages 15min PBS. La révélation est effectuée grâce au kit de détection de chimioluminescence ECL prime (Amersham), à 428 nm.

## 6. Extraction d'ARN totaux

Un culot cellulaire de 10ml de cellules Expi293F à 2.10<sup>6</sup> cellules/ml est obtenu après centrifugation 5 min à 300 RCF. Il est ensuite resuspendu dans 460 µL de solution de resuspension (glucose 10%, 12,5 mM Tris pH 7,6, 10 mM EDTA) avant d'ajouter 500µl de phénol acide. Les cellules sont vortexées 15 min à 4°C avant d'être centrifugées 10 min à 13000 rpm à 4°C. La phase aqueuse supérieure est transférée dans un tube

eppendorf stérile (sous la hotte chimique). Ensuite, 1 mL de Trizol est ajouté et l'ensemble est mélangé précautionneusement à la pipette avant d'incuber 5 min à température ambiante. On ajoute 100µL de chloroforme avant de mélanger doucement à la pipette et de laisser incuber 5 min à température ambiante. Le tube est centrifugé 5 min à 13000 rpm à 4°C. Transférer la phase aqueuse dans un nouveau tube eppendorf. On ajoute 200µL de chloroforme avant de mélanger doucement à la pipette et de laisser incuber 5 min à température ambiante. Le tube est centrifugé 5 min à 13000 rpm à 4°C. Transférer la phase aqueuse dans un nouveau tube eppendorf. Il faut ensuite ajouter 500 µL d'isopropanol, agiter en inversant le tube, et laisser 15 minutes dans la glace (un stockage à -20°C est possible à cette étape). Le tube est ensuite centrifugé 15 min à vitesse max à 4°C. Le culot est rincé avec 1 mL d'éthanol à 70% et centrifugation 5 min à 13000 rpm à 4°C. On retire alors le surnageant et on laisse sécher le culot quelques minutes sur la paillasse à température ambiante avant de le reprendre dans minimum 50 µL H<sub>2</sub>O RNase free.

### **7. Extraction d'ADN génomique**

Un culot cellulaire de 10ml de cellules Expi293F à  $2.10^6$  cellules/ml est obtenu après centrifugation 5 min à 300 RCF. 200µl du tampon d'extraction (2% Triton X-100, 1% SDS, 100 mM NaCl, 100 mM Tris-Cl (pH 8,0), 1 mM Na-EDTA) est ajouté sur le culot avant d'ajouter 200µl de phénol/chloroforme/alcool isoamylique (25:24:1). Le mix est alors vortexé 15 min à 4°C avant d'être centrifugé pendant 5 minutes à vitesse max. La couche supérieure contenant l'ADN est alors transférée dans un nouveau tube.

## VIII. Références bibliographiques

1. Ho, B.; Baryshnikova, A.; Brown, G.W. Unification of Protein Abundance Datasets Yields a Quantitative *Saccharomyces cerevisiae* Proteome. *Cell Syst.* **2018**, *6*, 192–205.e3, doi:10.1016/J.CELS.2017.12.004.
2. Palade, G.E. A small particulate component of the cytoplasm. *J. Biophys. Biochem. Cytol.* **1955**, *1*, 59–68, doi:10.1083/jcb.1.1.59.
3. Yusupov, M.M.; Yusupova, G.Z.; Baucom, A.; Lieberman, K.; Earnest, T.N.; Cate, J.H.D.; Noller, H.F. Crystal Structure of the Ribosome at 5.5 Å Resolution. *Science (80-. ).* **2001**, *292*, 883–896, doi:10.1126/SCIENCE.1060089.
4. Ben-Shem, A.; Loubresse, N.G. de; Melnikov, S.; Jenner, L.; Yusupova, G.; Yusupov, M. The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution. *Science (80-. ).* **2011**, *334*, 1524–1529, doi:10.1126/SCIENCE.1212642.
5. Melnikov, S.; Ben-Shem, A.; Garreau De Loubresse, N.; Jenner, L.; Yusupova, G.; Yusupov, M. One core, two shells: bacterial and eukaryotic ribosomes. *Nat. Struct. Mol. Biol.* **2012**, *19*, 560–567, doi:10.1038/NSMB.2313.
6. Lecompte, O.; Ripp, R.; Thierry, J.-C.; Moras, D.; Poch, O. Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. *Nucleic Acids Res.* **2002**, *30*, 5382, doi:10.1093/NAR/GKF693.
7. Zhou, X.; Liao, W.-J.; Liao, J.-M.; Liao, P.; Lu, H. Ribosomal proteins: functions beyond the ribosome. *J. Mol. Cell Biol.* **2015**, *7*, 92, doi:10.1093/JMCB/MJV014.
8. Mailliot, J.; Garreau De Loubresse, N.; Yusupova, G.; Meskauskas, A.; Dinman, J.D.; Yusupov, M. Crystal Structures of the uL3 Mutant Ribosome: Illustration of the Importance of Ribosomal Proteins for Translation Efficiency. *J. Mol. Biol.* **2016**, *428*, 2195–2202, doi:10.1016/J.JMB.2016.02.013.
9. Wilson, D.N.; Nierhaus, K.H. Ribosomal proteins in the spotlight. *Crit. Rev. Biochem. Mol. Biol.* **2005**, *40*, 243–267.
10. Wawiórka, L.; Molestak, E.; Szajwaj, M.; Michalec-Wawiórka, B.; Boguszewska, A.; Borkiewicz, L.; Liudkovska, V.; Kufel, J.; Tchórzewski, M. Functional analysis of the uL11 protein impact on translational machinery. *Cell Cycle* **2016**, *15*, 1060–1072, doi:10.1080/15384101.2016.1154245.
11. Panopoulos, P.; Dresios, J.; Synetos, D. Biochemical evidence of translational infidelity and decreased peptidyltransferase activity by a sarcin/ricin domain mutation of yeast 25S rRNA. *Nucleic Acids Res.* **2004**, *32*, 5398–5408, doi:10.1093/NAR/GKH860.
12. Loenarz, C.; Sekirnik, R.; Thalhammer, A.; Ge, W.; Spivakovsky, E.; Mackeen, M.M.; McDonough, M.A.; Cockman, M.E.; Kessler, B.M.; Ratcliffe, P.J.; et al. Hydroxylation

- of the eukaryotic ribosomal decoding center affects translational accuracy. *Proc. Natl. Acad. Sci.* **2014**, *111*, 4019–4024, doi:10.1073/PNAS.1311750111.
13. Lodish, H.; Berk, A.; Zipursky, S.L.; Matsudaira, P.; Baltimore, D.; Darnell, J. Processing of rRNA and tRNA. **2000**.
  14. Tafforeau, L. Qu'en est-il de la biogenèse des ribosomes chez l'homme? *médecine/sciences* **2015**, *31*, 622–628, doi:10.1051/MEDSCI/20153106015.
  15. Klinge, S.; Voigts-Hoffmann, F.; Leibundgut, M.; Ban, N. Atomic structures of the eukaryotic ribosome. *Trends Biochem. Sci.* **2012**, *37*, 189–198, doi:10.1016/J.TIBS.2012.02.007.
  16. Gilles, A.; Frechin, L.; Natchiar, K.; Biondani, G.; Loeffelholz, O. von; Holvec, S.; Malaval, J.L.; Winum, J.Y.; Klaholz, B.P.; Peyron, J.F. Targeting the Human 80S Ribosome in Cancer: From Structure to Function and Drug Design for Innovative Adjuvant Therapeutic Strategies. *Cells 2020, Vol. 9, Page 629* **2020**, *9*, 629, doi:10.3390/CELLS9030629.
  17. Wimberly, B.T.; Brodersen, D.E.; Clemons, W.M.; Morgan-Warren, R.J.; Carter, A.P.; Vonrhein, C.; Hartsch, T.; Ramakrishnan, V. Structure of the 30S ribosomal subunit. *Nat. 2000 4076802* **2000**, *407*, 327–339, doi:10.1038/35030006.
  18. Rabl, J.; Leibundgut, M.; Ataïde, S.F.; Haag, A.; Ban, N. Crystal Structure of the Eukaryotic 40S Ribosomal Subunit in Complex with Initiation Factor 1. *Science (80-. ).* **2011**, *331*, 730–736, doi:10.1126/SCIENCE.1198308.
  19. Hussain, T.; Llácer, J.L.; Fernández, I.S.; Munoz, A.; Martin-Marcos, P.; Savva, C.G.; Lorsch, J.R.; Hinnebusch, A.G.; Ramakrishnan, V. Structural Changes Enable Start Codon Recognition by the Eukaryotic Translation Initiation Complex. *Cell* **2014**, *159*, 597–607, doi:10.1016/J.CELL.2014.10.001.
  20. Takyar, S.; Hickerson, R.P.; Noller, H.F. mRNA Helicase Activity of the Ribosome. *Cell* **2005**, *120*, 49–58, doi:10.1016/J.CELL.2004.11.042.
  21. Des Georges, A.; Dhote, V.; Kuhn, L.; Hellen, C.U.T.; Pestova, T. V.; Frank, J.; Hashem, Y. Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* **2015**, *525*, 491–495, doi:10.1038/nature14891.
  22. Klinge, S.; Voigts-Hoffmann, F.; Leibundgut, M.; Arpagaus, S.; Ban, N. Crystal Structure of the Eukaryotic 60S Ribosomal Subunit in Complex with Initiation Factor 6. *Science (80-. ).* **2011**, *334*, 941–948, doi:10.1126/SCIENCE.1211204.
  23. Nissen, P.; Hansen, J.; Ban, N.; Moore, P.B.; Steitz, T.A. The structural basis of ribosome activity in peptide bond synthesis. *Science (80-. ).* **2000**, *289*, 920–930, doi:10.1126/science.289.5481.920.
  24. Schmeing, T.M.; Ramakrishnan, V. What recent ribosome structures have

- revealed about the mechanism of translation. *Nat.* 2009 4617268 **2009**, 461, 1234–1242, doi:10.1038/nature08403.
25. González-Porta, M.; Frankish, A.; Rung, J.; Harrow, J.; Brazma, A. Transcriptome analysis of human tissues and cell lines reveals one dominant transcript per gene. *Genome Biol.* 2013 147**2013**, 14, 1–11, doi:10.1186/GB-2013-14-7-R70.
  26. Sakharkar, M.K.; Chow, V.T.K.; Kanguane, P. Distributions of Exons and Introns in the Human Genome. *In Silico Biol.* **2004**, 4, 387–393.
  27. Elkon, R.; Ugalde, A.P.; Agami, R. Alternative cleavage and polyadenylation: extent, regulation and function. *Nat. Rev. Genet.* 2013 147 **2013**, 14, 496–506, doi:10.1038/nrg3482.
  28. Gruber, A.J.; Zavolan, M. Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.* 2019 2010**2019**, 20, 599–614, doi:10.1038/s41576-019-0145-z.
  29. Goss, D.J.; Kleiman, F.E. Poly(A) Binding Proteins—Are they all created equal? *Wiley Interdiscip. Rev. RNA* **2013**, 4, 167, doi:10.1002/WRNA.1151.
  30. Torres, A.G. Enjoy the Silence: Nearly Half of Human tRNA Genes Are Silent. *Bioinform. Biol. Insights* 2019, 13, 1–3.
  31. Kasai, H.; Nakanishi, K.; Macfarlane, R.D.; Torgerson, D.F.; Ohashi, Z.; McCloskey, J.A.; Gross, H.J.; Nishimura, S. The Structure of Q\* Nucleoside Isolated from Rabbit Liver Transfer Ribonucleic Acid. *J. Am. Chem. Soc.* 1976, 98, 5044–5046.
  32. Holley, R.W.; Apgar, J.; Everett, G.A.; Madison, J.T.; Marquisee, M.; Merrill, S.H.; Penswick, J.R.; Zamir, A. Structure of a Ribonucleic Acid. *Science (80-. )*. **1965**, 147, 1462–1465, doi:10.1126/SCIENCE.147.3664.1462.
  33. Goldman, E. Transfer RNA. *Encycl. Life Sci.* **2008**, doi:10.1002/9780470015902.A0000878.PUB2.
  34. Hardt, W.D.; Schlegl, J.; Erdmann, V.A.; Hartmann, R.K. Role of the D arm and the anticodon arm in tRNA recognition by eubacterial and eukaryotic RNase P enzymes. *Biochemistry* **2002**, 32, 13046–13053, doi:10.1021/BI00211A014.
  35. Nguyen, H.A.; Hoffer, E.D.; Dunham, C.M. Importance of a tRNA anticodon loop modification and a conserved, noncanonical anticodon stem pairing in tRNACGGPro for decoding. *J. Biol. Chem.* **2019**, 294, 5281–5291, doi:10.1074/jbc.RA119.007410.
  36. Chan, C.W.; Chetnani, B.; Mondragón, A. Structure and function of the T-loop structural motif in noncoding RNAs. *Wiley Interdiscip. Rev. RNA* **2013**, 4, 507–522, doi:10.1002/WRNA.1175.
  37. Yu, Y.C.; Han, J.M.; Kim, S. Aminoacyl-tRNA synthetases and amino acid signaling.

- Biochim. Biophys. Acta - Mol. Cell Res.* **2021**, *1868*, 118889, doi:10.1016/J.BBAMCR.2020.118889.
38. Grosjean, H.; Westhof, E. An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res.* **2016**, *44*, 8020–8040, doi:10.1093/NAR/GKW608.
  39. Shulgina, Y.; Eddy, S.R. A computational screen for alternative genetic codes in over 250,000 genomes. *Elife* **2021**, *10*, doi:10.7554/ELIFE.71402.
  40. Chan, P.P.; Lowe, T.M. GtRNADB 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* **2016**, *44*, D184–D189, doi:10.1093/NAR/GKV1309.
  41. Phizicky, E.M.; Hopper, A.K. tRNA biology charges to the front. *Genes Dev.* **2010**, *24*, 1832–1860, doi:10.1101/GAD.1956510.
  42. Ou, X.; Cao, J.; Cheng, A.; Peppelenbosch, M.P.; Pan, Q. Errors in translational decoding: tRNA wobbling or misincorporation? *PLoS Genet.* **2019**, *15*, doi:10.1371/JOURNAL.PGEN.1008017.
  43. CRICK, F.H.C.; BARNETT, L.; BRENNER, S.; WATTS-TOBIN, R.J. General Nature of the Genetic Code for Proteins. *Nat. 1961 1924809* **1961**, *192*, 1227–1232, doi:10.1038/1921227a0.
  44. Bernfield, M.R.; Nirenberg, M.W. RNA Codewords and Protein Synthesis. *Science (80-. )*. **1965**, *147*, 479–484, doi:10.1126/SCIENCE.147.3657.479.
  45. Crick, F.H.C. Codon—anticodon pairing: The wobble hypothesis. *J. Mol. Biol.* **1966**, *19*, 548–555, doi:10.1016/S0022-2836(66)80022-0.
  46. Knight, R.D.; Freeland, S.J.; Landweber, L.F. Rewiring the keyboard: evolvability of the genetic code. *Nat. Rev. Genet.* *2001 21* **2001**, *2*, 49–58, doi:10.1038/35047500.
  47. Barrell, B.G.; Bankier, A.T.; Drouin, J. A different genetic code in human mitochondria. *Nat. 1979 2825735* **1979**, *282*, 189–194, doi:10.1038/282189a0.
  48. Merrick, W.C.; Pavitt, G.D. Protein Synthesis Initiation in Eukaryotic Cells. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a033092, doi:10.1101/CSHPERSPECT.A033092.
  49. Kapp, L.D.; Lorsch, J.R. GTP-dependent Recognition of the Methionine Moiety on Initiator tRNA by Translation Factor eIF2. *J. Mol. Biol.* **2004**, *335*, 923–936, doi:10.1016/J.JMB.2003.11.025.
  50. Pisarev, A. V.; Hellen, C.U.T.; Pestova, T. V. Recycling of eukaryotic post-termination ribosomal complexes. *Cell* **2007**, *131*, 286, doi:10.1016/J.CELL.2007.08.041.

51. Sokabe, M.; Fraser, C.S.; Hershey, J.W.B. The human translation initiation multi-factor complex promotes methionyl-tRNA<sup>i</sup> binding to the 40S ribosomal subunit. *Nucleic Acids Res.* **2012**, *40*, 905, doi:10.1093/NAR/GKR772.
52. Asano, K.; Clayton, J.; Shalev, A.; Hinnebusch, A.G. A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA<sup>Met</sup> is an important translation initiation intermediate in vivo. *Genes Dev.* **2000**, *14*, 2534, doi:10.1101/GAD.831800.
53. Kessler, S.H.; Sachs, A.B. RNA Recognition Motif 2 of Yeast Pab1p Is Required for Its Functional Interaction with Eukaryotic Translation Initiation Factor 4G. *Mol. Cell. Biol.* **1998**, *18*, 51, doi:10.1128/MCB.18.1.51.
54. Imataka, H.; Gradi, A.; Sonenberg, N. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* **1998**, *17*, 7480, doi:10.1093/EMBOJ/17.24.7480.
55. Jackson, R.; Hellen, C.; Pestova, T. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 113–127, doi:10.1038/NRM2838.
56. Aitken, C.E.; Lorsch, J.R. A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.* **2012**, *19*, 568–576, doi:10.1038/nsmb.2303.
57. Bohlen, J.; Fenzl, K.; Kramer, G.; Bukau, B.; Teleman, A.A. Selective 40S Footprinting Reveals Cap-Tethered Ribosome Scanning in Human Cells. *Mol. Cell* **2020**, *79*, 561–574.e5, doi:10.1016/J.MOLCEL.2020.06.005.
58. Hinnebusch, A.G. The scanning mechanism of eukaryotic translation initiation. *Annu. Rev. Biochem.* **2014**, *83*, 779–812.
59. Brina, D.; Miluzio, A.; Ricciardi, S.; Biffo, S. eIF6 anti-association activity is required for ribosome biogenesis, translational control and tumor progression. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2015**, *1849*, 830–835, doi:10.1016/J.BBAGRM.2014.09.010.
60. Weis, F.; Giudice, E.; Churcher, M.; Jin, L.; Hilcenko, C.; Wong, C.C.; Traynor, D.; Kay, R.R.; Warren, A.J. Mechanism of eIF6 release from the nascent 60S ribosomal subunit. *Nat. Struct. Mol. Biol.* **2015**, *22*, 914–919, doi:10.1038/nsmb.3112.
61. Muthmann, N.; Špaček, P.; Reichert, D.; van Dülmen, M.; Rentmeister, A. Quantification of mRNA cap-modifications by means of LC-QqQ-MS. *Methods* **2021**, doi:10.1016/J.YMETH.2021.05.018.
62. Akichika, S.; Hirano, S.; Shichino, Y.; Suzuki, T.; Nishimasu, H.; Ishitani, R.; Sugita,

- A.; Hirose, Y.; Iwasaki, S.; Nureki, O.; et al. Cap-specific terminal N 6 -methylation of RNA by an RNA polymerase II-associated methyltransferase. *Science (80-. ).* **2019**, *363*, doi:10.1126/SCIENCE.AAV0080.
63. Sendinc, E.; Valle-Garcia, D.; Dhall, A.; Chen, H.; Henriques, T.; Navarrete-Perea, J.; Sheng, W.; Gygi, S.P.; Adelman, K.; Shi, Y. PCIF1 Catalyzes m6Am mRNA Methylation to Regulate Gene Expression. *Mol. Cell* **2019**, *75*, 620-630.e9, doi:10.1016/J.MOLCEL.2019.05.030.
  64. Pandey, R.R.; Delfino, E.; Homolka, D.; Roithova, A.; Chen, K.M.; Li, L.; Franco, G.; Vågbø, C.B.; Taillebourg, E.; Fauvarque, M.O.; et al. The Mammalian Cap-Specific m6Am RNA Methyltransferase PCIF1 Regulates Transcript Levels in Mouse Tissues. *Cell Rep.* **2020**, *32*, 108038, doi:10.1016/J.CELREP.2020.108038.
  65. Barth-Baus, D.; Bhasker, C.R.; Zoll, W.; Merrick, W.C. Influence of translation factor activities on start site selection in six different mRNAs. <http://dx.doi.org/10.4161/trla.24419> **2013**, *1*, e24419, doi:10.4161/TRLA.24419.
  66. Atger, F.; Gobet, C.; Marquis, J.; Martin, E.; Wang, J.; Weger, B.; Lefebvre, G.; Descombes, P.; Naef, F.; Gachon, F. Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proc. Natl. Acad. Sci.* **2015**, *112*, E6579–E6588, doi:10.1073/PNAS.1515308112.
  67. Gandin, V.; Masvidal, L.; Hulea, L.; Gravel, S.-P.; Cargnello, M.; McLaughlan, S.; Cai, Y.; Balanathan, P.; Morita, M.; Rajakumar, A.; et al. nanoCAGE reveals 5' UTR features that define specific modes of translation of functionally related MTOR-sensitive mRNAs. *Genome Res.* **2016**, *26*, 636–648, doi:10.1101/GR.197566.115.
  68. Elfakess, R.; Sinvani, H.; Haimov, O.; Svitkin, Y.; Sonenberg, N.; Dikstein, R. Unique translation initiation of mRNAs-containing TISU element. *Nucleic Acids Res.* **2011**, *39*, 7598–7609, doi:10.1093/NAR/GKR484.
  69. Sinvani, H.; Haimov, O.; Svitkin, Y.; Sonenberg, N.; Tamarkin-Ben-Harush, A.; Viollet, B.; Dikstein, R. Translational Tolerance of Mitochondrial Genes to Metabolic Energy Stress Involves TISU and eIF1-eIF4GI Cooperation in Start Codon Selection. *Cell Metab.* **2015**, *21*, 479–492, doi:10.1016/J.CMET.2015.02.010.
  70. Zhang, H.; Wang, Y.; Lu, J. Function and Evolution of Upstream ORFs in Eukaryotes. *Trends Biochem. Sci.* **2019**, *44*, 782–794, doi:10.1016/J.TIBS.2019.03.002.
  71. Dever, T.E.; Ivanov, I.P.; Sachs, M.S. Conserved Upstream Open Reading Frame Nascent Peptides That Control Translation. *Annu. Rev. Genet.* **2020**, *54*, 237–264.
  72. Calvo, S.E.; Pagliarini, D.J.; Mootha, V.K. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among

- humans. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 7507, doi:10.1073/PNAS.0810916106.
73. Resch, A.M.; Ogurtsov, A.Y.; Rogozin, I.B.; Shabalina, S.A.; Koonin, E. V Evolution of alternative and constitutive regions of mammalian 5'UTRs. *BMC Genomics* **2009**, *10*, 162, doi:10.1186/1471-2164-10-162.
74. Lawless, C.; Pearson, R.D.; Selley, J.N.; Smirnova, J.B.; Grant, C.M.; Ashe, M.P.; Pavitt, G.D.; Hubbard, S.J. Upstream sequence elements direct post-transcriptional regulation of gene expression under stress conditions in yeast. *BMC Genomics* **2009**, *10*, 7, doi:10.1186/1471-2164-10-7.
75. Pöyry, T.A.A.; Kaminski, A.; Jackson, R.J. What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes Dev.* **2004**, *18*, 62, doi:10.1101/GAD.276504.
76. Loughran, G.; Sachs, M.S.; Atkins, J.F.; Ivanov, I.P. Stringency of start codon selection modulates autoregulation of translation initiation factor eIF5. *Nucleic Acids Res.* **2012**, *40*, 2898, doi:10.1093/NAR/GKR1192.
77. Ivanov, I.P.; Loughran, G.; Sachs, M.S.; Atkins, J.F. Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1). *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18056, doi:10.1073/PNAS.1009269107.
78. Pestova, T. V.; Borukhov, S.I.; Hellen, C.U.T. Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nat.* **1998**, *394*, 854–859, doi:10.1038/29703.
79. Maag, D.; Algire, M.A.; Lorsch, J.R. Communication between Eukaryotic Translation Initiation Factors 5 and 1A within the Ribosomal Pre-initiation Complex Plays a Role in Start Site Selection. *J. Mol. Biol.* **2006**, *356*, 724–737, doi:10.1016/J.JMB.2005.11.083.
80. Palam, L.R.; Baird, T.D.; Wek, R.C. Phosphorylation of eIF2 Facilitates Ribosomal Bypass of an Inhibitory Upstream ORF to Enhance CHOP Translation. *J. Biol. Chem.* **2011**, *286*, 10939, doi:10.1074/JBC.M110.216093.
81. Wang, X.-Q.; Rothnagel, J.A. 5'-Untranslated regions with multiple upstream AUG codons can support low-level translation via leaky scanning and reinitiation. *Nucleic Acids Res.* **2004**, *32*, 1382, doi:10.1093/NAR/GKH305.
82. Kozak, M. Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **2002**, *299*, 1, doi:10.1016/S0378-1119(02)01056-9.
83. Calkhoven, C.F.; Müller, C.; Leutz, A. Translational control of C/EBP $\alpha$  and C/EBP $\beta$  isoform expression. *Genes Dev.* **2000**, *14*, 1920.
84. Lin, F.T.; MacDougald, O.A.; Diehl, A.M.; Lane, M.D. A 30-kDa alternative

- translation product of the CCAAT/enhancer binding protein alpha message: transcriptional activator lacking antimitotic activity. *Proc. Natl. Acad. Sci.* **1993**, *90*, 9606–9610, doi:10.1073/PNAS.90.20.9606.
85. T, P.; BU, M.; P, Z.; HS, R.; S, N.; S, S.; G, B.; W, H.; DG, T. Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia. *Nat. Genet.* **2001**, *27*, 263–270, doi:10.1038/85820.
  86. Yang, Y.; Wang, Z. IRES-mediated cap-independent translation, a path leading to hidden proteome. *J. Mol. Cell Biol.* **2019**, *11*, 911–919, doi:10.1093/JMCB/MJZ091.
  87. Pelletier, J.; Sonenberg, N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nat.* *1988 3346180* **1988**, *334*, 320–325, doi:10.1038/334320A0.
  88. Balvay, L.; Rifo, R.S.; Ricci, E.P.; Decimo, D.; Ohlmann, T. Structural and functional diversity of viral IRESes. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2009**, *1789*, 542–557, doi:10.1016/J.BBAGRM.2009.07.005.
  89. Hellen, C.U.T.; Sarnow, P. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.* **2001**, *15*, 1593–1612, doi:10.1101/GAD.891101.
  90. Komar, A.A.; Hatzoglou, M. Internal Ribosome Entry Sites in Cellular mRNAs: Mystery of Their Existence \*. *J. Biol. Chem.* **2005**, *280*, 23425–23428, doi:10.1074/JBC.R400041200.
  91. Kieft, J.S. Viral IRES RNA structures and ribosome interactions. *Trends Biochem. Sci.* **2008**, *33*, 274–283, doi:10.1016/J.TIBS.2008.04.007.
  92. KANAMORI, Y.; NAKASHIMA, N. A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. *RNA* **2001**, *7*, 266–274, doi:10.1017/S1355838201001741.
  93. Jan, E.; Sarnow, P. Factorless Ribosome Assembly on the Internal Ribosome Entry Site of Cricket Paralysis Virus. *J. Mol. Biol.* **2002**, *324*, 889–902, doi:10.1016/S0022-2836(02)01099-9.
  94. Nishiyama, T.; Yamamoto, H.; Shibuya, N.; Hatakeyama, Y.; Hachimori, A.; Uchiumi, T.; Nakashima, N. Structural elements in the internal ribosome entry site of Plautia stali intestine virus responsible for binding with ribosomes. *Nucleic Acids Res.* **2003**, *31*, 2434–2442, doi:10.1093/NAR/GKG336.
  95. Tsukiyama-Kohara, K.; Iizuka, N.; Kohara, M.; Nomoto, A. Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.* **1992**, *66*, 1476.
  96. Rijnbrand, R.; Straaten, T. van der; Rijn, P.A. van; Spaan, W.J.; Bredenbeek, P.J.

- Internal entry of ribosomes is directed by the 5' noncoding region of classical swine fever virus and is dependent on the presence of an RNA pseudoknot upstream of the initiation codon. *J. Virol.* **1997**, *71*, 451.
97. Pisarev, A. V.; Chard, L.S.; Kaku, Y.; Johns, H.L.; Shatsky, I.N.; Belsham, G.J. Functional and Structural Similarities between the Internal Ribosome Entry Sites of Hepatitis C Virus and Porcine Teschovirus, a Picornavirus. *J. Virol.* **2004**, *78*, 4487–4497, doi:10.1128/JVI.78.9.4487-4497.2004.
  98. Pestova, T. V; Hellen, C.U.; Shatsky, I.N. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol. Cell. Biol.* **1996**, *16*, 6859–6869, doi:10.1128/MCB.16.12.6859.
  99. Pestova, T. V; Shatsky, I.N.; Hellen, C.U. Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol. Cell. Biol.* **1996**, *16*, 6870–6878, doi:10.1128/MCB.16.12.6870.
  100. Lomakin, I.B.; Hellen, C.U.T.; Pestova, T. V. Physical Association of Eukaryotic Initiation Factor 4G (eIF4G) with eIF4A Strongly Enhances Binding of eIF4G to the Internal Ribosomal Entry Site of Encephalomyocarditis Virus and Is Required for Internal Initiation of Translation. *Mol. Cell. Biol.* **2000**, *20*, 6019–6029, doi:10.1128/MCB.20.16.6019-6029.2000.
  101. Sweeney, T.R.; Abaeva, I.S.; Pestova, T. V; Hellen, C.U.T. The mechanism of translation initiation on Type 1 picornavirus IRESs. *EMBO J.* **2014**, *33*, 76–92, doi:10.1002/EMBJ.201386124.
  102. Lulla, V.; Dinan, A.M.; Hosmillo, M.; Chaudhry, Y.; Sherry, L.; Irigoyen, N.; Nayak, K.M.; Stonehouse, N.J.; Zilbauer, M.; Goodfellow, I.; et al. An upstream protein-coding region in enteroviruses modulates virus infection in gut epithelial cells. *Nat. Microbiol.* **2018**, *4*, 280–292, doi:10.1038/S41564-018-0297-1.
  103. Komar, A.A.; Hatzoglou, M. Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states. *Cell Cycle* **2011**, *10*, 229–240.
  104. Meyer, K.D.; Patil, D.P.; Zhou, J.; Zinoviev, A.; Skabkin, M.A.; Elemento, O.; Pestova, T. V.; Qian, S.B.; Jaffrey, S.R. 5' UTR m6A Promotes Cap-Independent Translation. *Cell* **2015**, *163*, 999–1010, doi:10.1016/j.cell.2015.10.012.
  105. Yang, Y.; Fan, X.; Mao, M.; Song, X.; Wu, P.; Zhang, Y.; Jin, Y.; Yang, Y.; Chen, L.-L.; Wang, Y.; et al. Extensive translation of circular RNAs driven by N6-methyladenosine. *Cell Res.* **2017**, *27*, 626–641, doi:10.1038/CR.2017.31.
  106. Dresios, J.; Chappell, S.A.; Zhou, W.; Mauro, V.P. An mRNA-rRNA base-pairing mechanism for translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.* **2005**, *13*, 30–34, doi:10.1038/NSMB1031.

107. Weingarten-Gabbay, S.; Elias-Kirma, S.; Nir, R.; Gritsenko, A.A.; Stern-Ginossar, N.; Yakhini, Z.; Weinberger, A.; Segal, E. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science (80-. )*. **2016**, *351*, doi:10.1126/SCIENCE.AAD4939.
108. Dever, T.; Green, R. The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, 1–16, doi:10.1101/CSHPERSPECT.A013706.
109. Gutierrez, E.; Shin, B.-S.; Woolstenhulme, C.J.; Kim, J.-R.; Saini, P.; Buskirk, A.R.; Dever, T.E. eIF5A Promotes Translation of Polyproline Motifs. *Mol. Cell* **2013**, *51*, 35–45, doi:10.1016/J.MOLCEL.2013.04.021.
110. Melnikov, S.; Mailliot, J.; Shin, B.S.; Rigger, L.; Yusupova, G.; Micura, R.; Dever, T.E.; Yusupov, M. Crystal Structure of Hypusine-Containing Translation Factor eIF5A Bound to a Rotated Eukaryotic Ribosome. *J. Mol. Biol.* **2016**, *428*, 3570–3576, doi:10.1016/J.JMB.2016.05.011.
111. Schmidt, C.; Becker, T.; Heuer, A.; Braunger, K.; Shanmuganathan, V.; Pech, M.; Berninghausen, O.; Wilson, D.N.; Beckmann, R. Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. *Nucleic Acids Res.* **2016**, *44*, 1944–1951, doi:10.1093/NAR/GKV1517.
112. Shin, B.-S.; Katoh, T.; Gutierrez, E.; Kim, J.-R.; Suga, H.; Dever, T.E. Amino acid substrates impose polyamine, eIF5A, or hypusine requirement for peptide synthesis. *Nucleic Acids Res.* **2017**, *45*, 8392–8402, doi:10.1093/NAR/GKX532.
113. Pelechano, V.; Alepuz, P. EIF5A facilitates translation termination globally and promotes the elongation of many non polyproline-specific tripeptide sequences. *Nucleic Acids Res.* **2017**, *45*, 7326–7338, doi:10.1093/nar/gkx479.
114. Moazed, D.; Noller, H.F. Intermediate states in the movement of transfer RNA in the ribosome. *Nat.* **1989**, *342*, 142–148, doi:10.1038/342142A0.
115. Budkevich, T.; Giesebrecht, J.; Altman, R.B.; Munro, J.B.; Mielke, T.; Nierhaus, K.H.; Blanchard, S.C.; Spahn, C.M.T. Structure and Dynamics of the Mammalian Ribosomal Pretranslocation Complex. *Mol. Cell* **2011**, *44*, 214–224, doi:10.1016/J.MOLCEL.2011.07.040.
116. Behrmann, E.; Loerke, J.; Budkevich, T.V.; Yamamoto, K.; Schmidt, A.; Penczek, P.A.; Vos, M.R.; Bürger, J.; Mielke, T.; Scheerer, P.; et al. Structural Snapshots of Actively Translating Human Ribosomes. *Cell* **2015**, *161*, 845–857, doi:10.1016/J.CELL.2015.03.052.
117. Spahn, C.M.; Gomez-Lorenzo, M.G.; Grassucci, R.A.; Jørgensen, R.; Andersen, G.R.; Beckmann, R.; Penczek, P.A.; Ballesta, J.P.; Frank, J. Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA

- translocation. *EMBO J.* **2004**, *23*, 1008–1019, doi:10.1038/SJ.EMBOJ.7600102.
118. Taylor, D.J.; Nilsson, J.; Merrill, A.R.; Andersen, G.R.; Nissen, P.; Frank, J. Structures of modified eEF2-80S ribosome complexes reveal the role of GTP hydrolysis in translocation. *EMBO J.* **2007**, *26*, 2421, doi:10.1038/SJ.EMBOJ.7601677.
119. Ranjan, N.; Pochopien, A.A.; Wu, C.C.; Beckert, B.; Blanchet, S.; Green, R.; Rodnina, M. V.; Wilson, D.N. Yeast translation elongation factor eEF3 promotes late stages of tRNA translocation. *EMBO J.* **2021**, *40*, doi:10.15252/EMBJ.2020106449.
120. Nilsson, O.B.; Hedman, R.; Marino, J.; Wickles, S.; Bischoff, L.; Johansson, M.; Müller-Lucks, A.; Trovato, F.; Puglisi, J.D.; O'Brien, E.P.; et al. Cotranslational Protein Folding inside the Ribosome Exit Tunnel. *Cell Rep.* **2015**, *12*, 1533–1540, doi:10.1016/J.CELREP.2015.07.065.
121. Komar, A.A. Unraveling co-translational protein folding: Concepts and methods. *Methods* **2018**, *137*, 71–81, doi:10.1016/J.YMETH.2017.11.007.
122. Kramer, G.; Shiber, A.; Bukau, B. Mechanisms of cotranslational maturation of newly synthesized proteins. *Annu. Rev. Biochem.* 2019, *88*, 337–364.
123. Hellen, C.U.T. Translation Termination and Ribosome Recycling in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a032656, doi:10.1101/CSHPERSPECT.A032656.
124. Lawson, M.R.; Lessen, L.N.; Wang, J.; Prabhakar, A.; Corsepius, N.C.; Green, R.; Puglisi, J.D. Mechanisms that ensure speed and fidelity in eukaryotic translation termination. *bioRxiv* **2021**, *373*, 2021.04.01.438116, doi:10.1101/2021.04.01.438116.
125. Alkalaeva, E.Z.; Pisarev, A. V.; Frolova, L.Y.; Kisselev, L.L.; Pestova, T. V. In Vitro Reconstitution of Eukaryotic Translation Reveals Cooperativity between Release Factors eRF1 and eRF3. *Cell* **2006**, *125*, 1125–1136, doi:10.1016/j.cell.2006.04.035.
126. Bertram, G.; Bell, H.A.; Ritchie, D.W.; Fullerton, G.; Stansfield, I. Terminating eukaryote translation: domain 1 of release factor eRF1 functions in stop codon recognition. *RNA* **2000**, *6*, 1236, doi:10.1017/S1355838200000777.
127. Frolova, L.Y.; Tsivkovskii, R.Y.; Sivolobova, G.F.; Oparina, N.Y.; Serpinsky, O.I.; Blinov, V.M.; Tatkov, S.I.; Kisselev, L.L. Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. *RNA* **1999**, *5*, 1014, doi:10.1017/S135583829999043X.
128. Haurlyliuk, V.; Zavialov, A.; Kisselev, L.; Ehrenberg, M. Class-1 release factor eRF1 promotes GTP binding by class-2 release factor eRF3. *Biochimie* **2006**, *88*, 747–757, doi:10.1016/J.BIOCHI.2006.06.001.

129. Cheng, Z.; Saito, K.; Pisarev, A. V.; Wada, M.; Pisareva, V.P.; Pestova, T. V.; Gajda, M.; Round, A.; Kong, C.; Lim, M.; et al. Structural insights into eRF3 and stop codon recognition by eRF1. *Genes Dev.* **2009**, *23*, 1106, doi:10.1101/GAD.1770109.
130. Georges, A. des; Hashem, Y.; Unbehauen, A.; Grassucci, R.A.; Taylor, D.; Hellen, C.U.T.; Pestova, T. V.; Frank, J. Structure of the mammalian ribosomal pre-termination complex associated with eRF1•eRF3•GDPNP. *Nucleic Acids Res.* **2014**, *42*, 3409, doi:10.1093/NAR/GKT1279.
131. Preis, A.; Heuer, A.; Barrio-Garcia, C.; Hauser, A.; Eyler, D.E.; Berninghausen, O.; Green, R.; Becker, T.; Beckmann, R. Cryoelectron Microscopic Structures of Eukaryotic Translation Termination Complexes Containing eRF1-eRF3 or eRF1-ABCE1. *Cell Rep.* **2014**, *8*, 59, doi:10.1016/J.CELREP.2014.04.058.
132. Kryuchkova, P.; Grishin, A.; Eliseev, B.; Karyagina, A.; Frolova, L.; Alkalaeva, E. Two-step model of stop codon recognition by eukaryotic release factor eRF1. *Nucleic Acids Res.* **2013**, *41*, 4573–4586, doi:10.1093/NAR/GKT113.
133. Korostelev, A.; Asahara, H.; Lancaster, L.; Laurberg, M.; Hirschi, A.; Zhu, J.; Trakhanov, S.; Scott, W.G.; Noller, H.F. Crystal structure of a translation termination complex formed with release factor RF2. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 19684–19689, doi:10.1073/pnas.0810953105.
134. Weixlbaumer, A.; Jin, H.; Neubauer, C.; Voorhees, R.M.; Petry, S.; Kelley, A.C.; Ramakrishnan, V. Insights into translational termination from the structure of RF2 bound to the ribosome. *Science (80-. ).* **2008**, *322*, 953–956, doi:10.1126/science.1164840.
135. Rees, D.C.; Johnson, E.; Lewinson, O. ABC transporters: The power to change. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 218, doi:10.1038/NRM2646.
136. Becker, T.; Franckenberg, S.; Wickles, S.; Shoemaker, C.J.; Anger, A.M.; Armache, J.-P.; Sieber, H.; Ungewickell, C.; Berninghausen, O.; Daberkow, I.; et al. Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature* **2012**, *482*, 501, doi:10.1038/NATURE10829.
137. Franckenberg, S.; Becker, T.; Beckmann, R. Structural view on recycling of archaeal and eukaryotic ribosomes after canonical termination and ribosome rescue. *Curr. Opin. Struct. Biol.* **2012**, *22*, 786–796, doi:10.1016/J.SBI.2012.08.002.
138. Kiosze-Becker, K.; Ori, A.; Gerovac, M.; Heuer, A.; Nürenberg-Goloub, E.; Rashid, U.J.; Becker, T.; Beckmann, R.; Beck, M.; Tampé, R. Structure of the ribosome post-recycling complex probed by chemical cross-linking and mass spectrometry. *Nat. Commun.* **2016**, *7*, doi:10.1038/NCOMMS13248.
139. Heuer, A.; Gerovac, M.; Schmidt, C.; Trowitzsch, S.; Preis, A.; Kötter, P.; Berninghausen, O.; Becker, T.; Beckmann, R.; Tampé, R. Structure of the 40S–

- ABCE1 post-splitting complex in ribosome recycling and translation initiation. *Nat. Struct. Mol. Biol.* **2017**, *24*, 453–460, doi:10.1038/NSMB.3396.
140. Pisarev, A. V.; Skabkin, M.A.; Pisareva, V.P.; Skabkina, O. V.; Rakotondrafara, A.M.; Hentze, M.W.; Hellen, C.U.T.; Pestova, T. V. The Role of ABCE1 in Eukaryotic Posttermination Ribosomal Recycling. *Mol. Cell* **2010**, *37*, 196–210, doi:10.1016/j.molcel.2009.12.034.
  141. Pisareva, V.P.; Skabkin, M.A.; Hellen, C.U.T.; Pestova, T. V.; Pisarev, A. V. Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO J.* **2011**, *30*, 1804–1817, doi:10.1038/emboj.2011.93.
  142. Shoemaker, C.J.; Green, R. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, doi:10.1073/pnas.1113956108.
  143. Barthelme, D.; Dinkelaker, S.; Albers, S.V.; Londei, P.; Ermler, U.; Tampé, R. Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3228–3233, doi:10.1073/pnas.1015953108.
  144. Lomakin, I.B.; Shirokikh, N.E.; Yusupov, M.M.; Hellen, C.U.T.; Pestova, T. V The fidelity of translation initiation: reciprocal activities of eIF1, IF3 and YciH. *EMBO J.* **2006**, *25*, 196, doi:10.1038/SJ.EMBOJ.7600904.
  145. Skabkin, M.A.; Skabkina, O. V.; Dhote, V.; Komar, A.A.; Hellen, C.U.T.; Pestova, T. V. Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes Dev.* **2010**, *24*, 1787, doi:10.1101/GAD.1957510.
  146. Lomakin, I.B.; Stolboushkina, E.A.; Vaidya, A.T.; Zhao, C.; Garber, M.B.; Dmitriev, S.E.; Steitz, T.A. Crystal structure of the human ribosome in complex with DENR-MCT-1. *Cell Rep.* **2017**, *20*, 521, doi:10.1016/J.CELREP.2017.06.025.
  147. Weisser, M.; Schäfer, T.; Leibundgut, M.; Böhringer, D.; Aylett, C.H.S.; Ban, N. Structural and Functional Insights into Human Re-initiation Complexes. *Mol. Cell* **2017**, *67*, 447–456.e7, doi:10.1016/J.MOLCEL.2017.06.032.
  148. Goldstrohm, A.C.; Wickens, M. Multifunctional deadenylase complexes diversify mRNA control. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 337–344, doi:10.1038/nrm2370.
  149. Kurland, C.; Gallant, J. Errors of heterologous protein expression. *Curr. Opin. Biotechnol.* **1996**, *7*, 489–493, doi:10.1016/S0958-1669(96)80050-4.
  150. Meyer, S.; Temme, C.; Wahle, E. Messenger RNA turnover in eukaryotes: Pathways and enzymes. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 197–216.

151. Tatosyan, K.A.; Ustyantsev, I.G.; Kramerov, D.A. RNA Degradation in Eukaryotic Cells. *Mol. Biol.* **2020**, *54*, 485–502, doi:10.1134/S0026893320040159.
152. Wahle, E.; Winkler, G.S. RNA decay machines: Deadenylation by the Ccr4–Not and Pan2–Pan3 complexes. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2013**, *1829*, 561–570, doi:10.1016/J.BBAGRM.2013.01.003.
153. Ivshina, M.; Lasko, P.; Richter, J.D. Cytoplasmic polyadenylation element binding proteins in development, health, and disease. *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 393–415.
154. Lombardi, S.; Testa, M.F.; Pinotti, M.; Branchini, A. Molecular Insights into Determinants of Translational Readthrough and Implications for Nonsense Suppression Approaches. *Int. J. Mol. Sci.* **2020**, *21*, 1–25, doi:10.3390/IJMS21249449.
155. Wang, M.; Ogé, L.; Perez-Garcia, M.-D.; Hamama, L.; Sakr, S. The PUF Protein Family: Overview on PUF RNA Targets, Biological Functions, and Post Transcriptional Regulation. *Int. J. Mol. Sci.* **2018**, *19*, doi:10.3390/IJMS19020410.
156. Legrand, N.; Dixon, D.A.; Sobolewski, C. AU-rich element-binding proteins in colorectal cancer. *World J. Gastrointest. Oncol.* **2019**, *11*, 71, doi:10.4251/WJGO.V11.I2.71.
157. He, F.; Li, X.; Spatrick, P.; Casillo, R.; Dong, S.; Jacobson, A. Genome-Wide Analysis of mRNAs Regulated by the Nonsense-Mediated and 5' to 3' mRNA Decay Pathways in Yeast. *Mol. Cell* **2003**, *12*, 1439–1452, doi:10.1016/S1097-2765(03)00446-5.
158. Lelivelt, M.J.; Culbertson, M.R. Yeast Upf Proteins Required for RNA Surveillance Affect Global Expression of the Yeast Transcriptome. *Mol. Cell. Biol.* **1999**, *19*, 6710–6719, doi:10.1128/MCB.19.10.6710.
159. Ivanov, P. V; Gehring, N.H.; Kunz, J.B.; Hentze, M.W.; Kulozik, A.E. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J.* **2008**, *27*, 736, doi:10.1038/EMBOJ.2008.17.
160. Kashima, I.; Yamashita, A.; Izumi, N.; Kataoka, N.; Morishita, R.; Hoshino, S.; Ohno, M.; Dreyfuss, G.; Ohno, S. Binding of a novel SMG-1–Upf1–eRF1–eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev.* **2006**, *20*, 355, doi:10.1101/GAD.1389006.
161. Singh, G.; Rebbapragada, I.; Lykke-Andersen, J. A Competition between Stimulators and Antagonists of Upf Complex Recruitment Governs Human Nonsense-Mediated mRNA Decay. *PLoS Biol.* **2008**, *6*, 860–871,

doi:10.1371/JOURNAL.PBIO.0060111.

162. Arias-Palomo, E.; Yamashita, A.; Fernández, I.S.; Núñez-Ramírez, R.; Bamba, Y.; Izumi, N.; Ohno, S.; Llorca, O. The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. *Genes Dev.* **2011**, *25*, 153, doi:10.1101/GAD.606911.
163. Deniaud, A.; Karuppasamy, M.; Bock, T.; Masiulis, S.; Huard, K.; Garzoni, F.; Kerschgens, K.; Hentze, M.W.; Kulozik, A.E.; Beck, M.; et al. A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation. *Nucleic Acids Res.* **2015**, *43*, 7600, doi:10.1093/NAR/GKV668.
164. Melero, R.; Uchiyama, A.; Castaño, R.; Kataoka, N.; Kurosawa, H.; Ohno, S.; Yamashita, A.; Llorca, O. Structures of SMG1-UPFs Complexes: SMG1 Contributes to Regulate UPF2-Dependent Activation of UPF1 in NMD. *Structure* **2014**, *22*, 1105–1119, doi:10.1016/J.STR.2014.05.015.
165. Yamashita, A.; Izumi, N.; Kashima, I.; Ohnishi, T.; Saari, B.; Katsuhata, Y.; Muramatsu, R.; Morita, T.; Iwamatsu, A.; Hachiya, T.; et al. SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. *Genes Dev.* **2009**, *23*, 1091, doi:10.1101/GAD.1767209.
166. Eberle, A.B.; Lykke-Andersen, S.; Mühlemann, O.; Jensen, T.H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat. Struct. Mol. Biol.* **2008**, *16*, 49–55, doi:10.1038/NSMB.1530.
167. Huntzinger, E.; Kashima, I.; Fauser, M.; Saulière, J.; Izaurralde, E. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* **2008**, *14*, 2609, doi:10.1261/RNA.1386208.
168. Loh, B.; Jonas, S.; Izaurralde, E. The SMG5–SMG7 heterodimer directly recruits the CCR4–NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes Dev.* **2013**, *27*, 2125, doi:10.1101/GAD.226951.113.
169. Joazeiro, C.A.P. Mechanisms and functions of ribosome-associated protein quality control. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 368, doi:10.1038/S41580-019-0118-2.
170. Hir, H. Le; Izaurralde, E.; Maquat, L.E.; Moore, M.J. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. *EMBO J.* **2000**, *19*, 6860, doi:10.1093/EMBOJ/19.24.6860.
171. Saulière, J.; Murigneux, V.; Wang, Z.; Marquet, E.; Barbosa, I.; Le Tonquèze, O.; Audic, Y.; Paillard, L.; Crollius, H.R.; Le Hir, H. CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nat. Struct.*

- Mol. Biol.* 2012 1911 **2012**, 19, 1124–1131, doi:10.1038/NSMB.2420.
172. Nagy, E.; Maquat, L.E. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem. Sci.* **1998**, 23, 198–199, doi:10.1016/S0968-0004(98)01208-0.
  173. Fatscher, T.; Boehm, V.; Weiche, B.; Gehring, N.H. The interaction of cytoplasmic poly(A)-binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. *RNA* **2014**, 20, 1579, doi:10.1261/RNA.044933.114.
  174. Peixeiro, I.; Inácio, Â.; Barbosa, C.; Silva, A.L.; Liebhaber, S.A.; Romão, L. Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. *Nucleic Acids Res.* **2012**, 40, 1160, doi:10.1093/NAR/GKR820.
  175. Silva, A.L.; Ribeiro, P.; Inácio, Â.; Liebhaber, S.A.; Romão, L. Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. *RNA* **2008**, 14, 563, doi:10.1261/RNA.815108.
  176. Frischmeyer, P.A.; Van Hoof, A.; O'Donnell, K.; Guerrerio, A.L.; Parker, R.; Dietz, H.C. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science (80-. )*. **2002**, 295, 2258–2261, doi:10.1126/science.1067338.
  177. Van Hoof, A.; Frischmeyer, P.A.; Dietz, H.C.; Parker, R. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science (80-. )*. **2002**, 295, 2262–2264, doi:10.1126/science.1067272.
  178. J, L.; C, D. Electrostatics in the ribosomal tunnel modulate chain elongation rates. *J. Mol. Biol.* **2008**, 384, 73–86, doi:10.1016/J.JMB.2008.08.089.
  179. Ito-Harashima, S.; Kuroha, K.; Tatematsu, T.; Inada, T. Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. *Genes Dev.* **2007**, 21, 519, doi:10.1101/GAD.1490207.
  180. Han, P.; Shichino, Y.; Schneider-Poetsch, T.; Mito, M.; Hashimoto, S.; Udagawa, T.; Kohno, K.; Yoshida, M.; Mishima, Y.; Inada, T.; et al. Genome-wide Survey of Ribosome Collision. *Cell Rep.* **2020**, 31, 107610, doi:10.1016/J.CELREP.2020.107610.
  181. Doma, M.K.; Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nat.* 2006 4407083 **2006**, 440, 561–564, doi:10.1038/NATURE04530.
  182. Passos, D.O.; Doma, M.K.; Shoemaker, C.J.; Muhlrud, D.; Green, R.; Weissman, J.; Hollien, J.; Parker, R. Analysis of Dom34 and Its Function in No-Go Decay.

<https://doi.org/10.1091/mbc.e09-01-0028> **2009**, *20*, 3025–3032, doi:10.1091/MBC.E09-01-0028.

183. Tomecki, R.; Kristiansen, M.S.; Lykke-Andersen, S.; Chlebowski, A.; Larsen, K.M.; Szczesny, R.J.; Drazkowska, K.; Pastula, A.; Andersen, J.S.; Stepien, P.P.; et al. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J.* **2010**, *29*, 2342–2357, doi:10.1038/EMBOJ.2010.121.
184. Shao, S.; Brown, A.; Santhanam, B.; Hegde, R.S. Structure and assembly pathway of the ribosome quality control complex. *Mol. Cell* **2015**, *57*, 433–444, doi:10.1016/j.molcel.2014.12.015.
185. Shen, P.S.; Park, J.; Qin, Y.; Li, X.; Parsawar, K.; Larson, M.H.; Cox, J.; Cheng, Y.; Lambowitz, A.M.; Weissman, J.S.; et al. Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. *Science* **2015**, *347*, 75, doi:10.1126/SCIENCE.1259724.
186. Brandman, O.; Stewart-Ornstein, J.; Wong, D.; Larson, A.; Williams, C.C.; Li, G.-W.; Zhou, S.; King, D.; Shen, P.S.; Weibezahn, J.; et al. A Ribosome-Bound Quality Control Complex Triggers Degradation of Nascent Peptides and Signals Translation Stress. *Cell* **2012**, *151*, 1042–1054, doi:10.1016/J.CELL.2012.10.044.
187. Mauro, V.P.; Edelman, G.M. The ribosome filter hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12031, doi:10.1073/PNAS.192442499.
188. Locati, M.D.; Pagano, J.F.B.; Girard, G.; Ensink, W.A.; Olst, M. van; Leeuwen, S. van; Nehrdich, U.; Spaink, H.P.; Rauwerda, H.; Jonker, M.J.; et al. Expression of distinct maternal and somatic 5.8S, 18S, and 28S rRNA types during zebrafish development. *RNA* **2017**, *23*, 1188–1199, doi:10.1261/RNA.061515.117.
189. Belin, S.; Beghin, A.; Solano-González, E.; Bezin, L.; Brunet-Manquat, S.; Textoris, J.; Prats, A.-C.; Mertani, H.C.; Dumontet, C.; Diaz, J.-J. Dysregulation of Ribosome Biogenesis and Translational Capacity Is Associated with Tumor Progression of Human Breast Cancer Cells. *PLoS One* **2009**, *4*, e7147, doi:10.1371/JOURNAL.PONE.0007147.
190. Bellodi, C.; Krasnykh, O.; Haynes, N.; Theodoropoulou, M.; Peng, G.; Montanaro, L.; Ruggero, D. Loss of Function of the Tumor Suppressor DKC1 Perturbs p27 Translation Control and Contributes to Pituitary Tumorigenesis. *Cancer Res.* **2010**, *70*, 6026–6035, doi:10.1158/0008-5472.CAN-09-4730.
191. Yoon, A.; Peng, G.; Brandenburg, Y.; Zollo, O.; Xu, W.; Rego, E.; Ruggero, D. Impaired Control of IRES-Mediated Translation in X-Linked Dyskeratosis Congenita. *Science (80- )*. **2006**, *312*, 902–906, doi:10.1126/SCIENCE.1123835.
192. Jansson, M.D.; Häfner, S.J.; Altinel, K.; Tehler, D.; Krogh, N.; Jakobsen, E.; Andersen, J. V.; Andersen, K.L.; Schoof, E.M.; Ménard, P.; et al. Regulation of translation by

- site-specific ribosomal RNA methylation. *Nat. Struct. Mol. Biol.* 2021 2811 **2021**, 28, 889–899, doi:10.1038/s41594-021-00669-4.
193. Kondrashov, N.; Pusic, A.; Stumpf, C.R.; Shimizu, K.; Hsieh, A.C.; Xue, S.; Ishijima, J.; Shiroishi, T.; Barna, M. Ribosome-Mediated Specificity in Hox mRNA Translation and Vertebrate Tissue Patterning. *Cell* **2011**, 145, 383–397, doi:10.1016/J.CELL.2011.03.028.
194. Xue, S.; Tian, S.; Fujii, K.; Kladwang, W.; Das, R.; Barna, M. RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation. *Nat.* 2014 5177532 **2014**, 517, 33–38, doi:10.1038/nature14010.
195. Spence, J.; Gali, R.R.; Dittmar, G.; Sherman, F.; Karin, M.; Finley, D. Cell Cycle-Regulated Modification of the Ribosome by a Variant Multiubiquitin Chain. *Cell* **2000**, 102, 67–76, doi:10.1016/S0092-8674(00)00011-8.
196. Fan-Minogue, H.; Bedwell, D.M. Eukaryotic ribosomal RNA determinants of aminoglycoside resistance and their role in translational fidelity. *RNA* **2008**, 14, 148, doi:10.1261/RNA.805208.
197. Garzia, A.; Jafarnejad, S.M.; Meyer, C.; Chapat, C.; Gogakos, T.; Morozov, P.; Amiri, M.; Shapiro, M.; Molina, H.; Tuschl, T.; et al. The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nat. Commun.* 2017 81 **2017**, 8, 1–10, doi:10.1038/ncomms16056.
198. Cate, J.H.; Yusupov, M.M.; Yusupova, G.Z.; Earnest, T.N.; Noller, H.F. X-ray Crystal Structures of 70S Ribosome Functional Complexes. *Science (80-. ).* **1999**, 285, 2095–2104, doi:10.1126/SCIENCE.285.5436.2095.
199. Ogle, J.M.; Brodersen, D.E.; Clemons, W.M.; Tarry, M.J.; Carter, A.P.; Ramakrishnan, V. Recognition of Cognate Transfer RNA by the 30S Ribosomal Subunit. *Science (80-. ).* **2001**, 292, 897–902, doi:10.1126/SCIENCE.1060612.
200. Shao, S.; Murray, J.; Brown, A.; Taunton, J.; Ramakrishnan, V.; Hegde, R.S. Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. *Cell* **2016**, 167, 1229, doi:10.1016/J.CELL.2016.10.046.
201. Loveland, A.B.; Demo, G.; Grigorieff, N.; Korostelev, A.A. Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nat.* 2017 5467656 **2017**, 546, 113–117, doi:10.1038/NATURE22397.
202. Rozov, A.; Demeshkina, N.; Westhof, E.; Yusupov, M.; Yusupova, G. Structural insights into the translational infidelity mechanism. *Nat. Commun.* **2015**, 6, doi:10.1038/NCOMMS8251.
203. Rozov, A.; Demeshkina, N.; Khusainov, I.; Westhof, E.; Yusupov, M.; Yusupova, G. Novel base-pairing interactions at the tRNA wobble position crucial for accurate

- reading of the genetic code. **2016**, *7*, 1–10.
204. Kondo, J.; Urzhumtsev, A.; Westhof, E. Two conformational states in the crystal structure of the Homo sapiens cytoplasmic ribosomal decoding A site. *Nucleic Acids Res.* **2006**, *34*, 676, doi:10.1093/NAR/GKJ467.
205. Demeshkina, N.; Jenner, L.; Westhof, E.; Yusupov, M.; Yusupova, G. A new understanding of the decoding principle on the ribosome. *Nat.* **2012**, *484*, 256–259, doi:10.1038/nature10913.
206. Voorhees, R.M.; Ramakrishnan, V. Structural basis of the translational elongation cycle. *Annu. Rev. Biochem.* **2013**, *82*, 203–236.
207. Velichutina, I. V.; Hong, J.Y.; Mesecar, A.D.; Chernoff, Y.O.; Liebman, S.W. Genetic interaction between yeast *Saccharomyces cerevisiae* release factors and the decoding region of 18 S rRNA. *J. Mol. Biol.* **2001**, *305*, 715–727, doi:10.1006/JMBI.2000.4329.
208. O'Connor, M.; Thomas, C.L.; Zimmermann, R.A.; Dahlberg, A.E. Decoding fidelity at the ribosomal A and P sites: influence of mutations in three different regions of the decoding domain in 16S rRNA. *Nucleic Acids Res.* **1997**, *25*, 1185, doi:10.1093/NAR/25.6.1185.
209. Chernoff, Y.O.; Newnam, G.P.; Liebman, S.W. The translational function of nucleotide C1054 in the small subunit rRNA is conserved throughout evolution: genetic evidence in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 2517, doi:10.1073/PNAS.93.6.2517.
210. Sharma, D.; Cukras, A.R.; Rogers, E.J.; Southworth, D.R.; Green, R. Mutational Analysis of S12 Protein and Implications for the Accuracy of Decoding by the Ribosome. *J. Mol. Biol.* **2007**, *374*, 1065–1076, doi:10.1016/J.JMB.2007.10.003.
211. Atkins, J.F.; Elseviers, D.; Gorini, L. Low Activity of  $\beta$ -Galactosidase in Frameshift Mutants of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 1192, doi:10.1073/PNAS.69.5.1192.
212. Ellis, N.; Gallant, J. An estimate of the global error frequency in translation. *Mol. Gen. Genet. MGG 1982 1882* **1982**, *188*, 169–172, doi:10.1007/BF00332670.
213. Jakubowski, H. Quality control in tRNA charging. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 295–310, doi:10.1002/WRNA.122.
214. Sun, L.; Gomes, A.C.; He, W.; Zhou, H.; Wang, X.; Pan, D.W.; Schimmel, P.; Pan, T.; Yang, X.L. Evolutionary gain of alanine mischarging to non-cognate tRNAs with a G4:U69 base pair. *J. Am. Chem. Soc.* **2016**, *138*, 12948, doi:10.1021/JACS.6B07121.
215. Caliskan, N.; Katunin, V.I.; Belardinelli, R.; Peske, F.; Rodnina, M. V. Programmed –

- 1 Frameshifting by Kinetic Partitioning during Impeded Translocation. *Cell* **2014**, *157*, 1619, doi:10.1016/J.CELL.2014.04.041.
216. Clark, M.B.; Jänicke, M.; Gottesbühren, U.; Kleffmann, T.; Legge, M.; Poole, E.S.; Tate, W.P. Mammalian Gene PEG10 Expresses Two Reading Frames by High Efficiency –1 Frameshifting in Embryonic-associated Tissues \*. *J. Biol. Chem.* **2007**, *282*, 37359–37369, doi:10.1074/JBC.M705676200.
217. Wills, N.M.; Moore, B.; Hammer, A.; Gesteland, R.F.; Atkins, J.F. A Functional –1 Ribosomal Frameshift Signal in the Human Paraneoplastic Ma3 Gene \*. *J. Biol. Chem.* **2006**, *281*, 7082–7088, doi:10.1074/JBC.M511629200.
218. Brandt, J.; Veith, A.M.; Volff, J.-N. A family of neofunctionalized Ty3/gypsy retrotransposon genes in mammalian genomes. *Cytogenet. Genome Res.* **2005**, *110*, 307–317, doi:10.1159/000084963.
219. Sundararajan, A.; Michaud, W.A.; Qian, Q.; Stahl, G.; Farabaugh, P.J. Near-cognate peptidyl-tRNAs promote +1 programmed translational frameshifting in yeast. *Mol. Cell* **1999**, *4*, 1005–1015, doi:10.1016/S1097-2765(00)80229-4.
220. Vimaladithan, A.; Farabaugh, P.J. Special peptidyl-tRNA molecules can promote translational frameshifting without slippage. *Mol. Cell. Biol.* **1994**, *14*, 8107, doi:10.1128/MCB.14.12.8107.
221. Taliaferro, D.; Farabaugh, P.J. An mRNA sequence derived from the yeast EST3 gene stimulates programmed +1 translational frameshifting. *RNA* **2007**, *13*, 606, doi:10.1261/RNA.412707.
222. Namy, O.; Galopier, A.; Martini, C.; Matsufuji, S.; Fabret, C.; Rousset, J.-P. Epigenetic control of polyamines by the prion [PSI+]. *Nat. Cell Biol.* **2008**, *10*, 1069–1075, doi:10.1038/ncb1766.
223. Castets, P.; Lescure, A.; Guicheney, P.; Allamand, V. Selenoprotein N in skeletal muscle: from diseases to function. *J. Mol. Med. (Berl)*. **2012**, *90*, 1095–1107, doi:10.1007/s00109-012-0896-x.
224. Franke, K.W. A New Toxicant Occurring Naturally in Certain Samples of Plant FoodstuffsI. Results Obtained in Preliminary Feeding Trials: Eight Figures. *J. Nutr.* **1934**, *8*, 597–608, doi:10.1093/JN/8.5.597.
225. Lammi, M.J.; Qu, C. Selenium-Related Transcriptional Regulation of Gene Expression. *Int. J. Mol. Sci.* **2018**, *19*, doi:10.3390/IJMS19092665.
226. Bulteau, A.L.; Chavatte, L. Update on selenoprotein biosynthesis. *Antioxidants Redox Signal.* **2015**, *23*, 775–794.
227. Gupta, M.; Copeland, P.R. Functional Analysis of the Interplay between Translation Termination, Selenocysteine Codon Context, and Selenocysteine

- Insertion Sequence-binding Protein 2. *J. Biol. Chem.* **2007**, *282*, 36797, doi:10.1074/JBC.M707061200.
228. Shetty, S.P.; Copeland, P.R. Selenocysteine incorporation: A trump card in the game of mRNA decay. *Biochimie* **2015**, *114*, 97, doi:10.1016/J.BIOCHI.2015.01.007.
229. Blanchet, S.; Cornu, D.; Argentini, M.; Namy, O. New insights into the incorporation of natural suppressor tRNAs at stop codons in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **2014**, *42*, 10061–10072, doi:10.1093/nar/gku663.
230. Xue, X.; Mutyam, V.; Thakerar, A.; Mobley, J.; Bridges, R.J.; Rowe, S.M.; Keeling, K.M.; Bedwell, D.M. Identification of the amino acids inserted during suppression of CFTR nonsense mutations and determination of their functional consequences. *Hum. Mol. Genet.* **2017**, *26*, 3116–3129, doi:10.1093/hmg/ddx196.
231. Skuzeski, J.M.; Nichols, L.M.; Gesteland, R.F.; Atkins, J.F. The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. *J. Mol. Biol.* **1991**, *218*, 365–373, doi:10.1016/0022-2836(91)90718-L.
232. Firth, A.E.; Brierley, I. Non-canonical translation in RNA viruses. *J. Gen. Virol.* **2012**, *93*, 1385, doi:10.1099/VIR.0.042499-0.
233. Kramarski, L.; Arbely, E. Translational read-through promotes aggregation and shapes stop codon identity. *Nucleic Acids Res.* **2020**, *48*, 3747, doi:10.1093/NAR/GKAA136.
234. Loughran, G.; Jungreis, I.; Tzani, I.; Power, M.; Dmitriev, R.I.; Ivanov, I.P.; Kellis, M.; Atkins, J.F. Stop codon readthrough generates a C-terminally extended variant of the human vitamin D receptor with reduced calcitriol response. *J. Biol. Chem.* **2018**, *293*, 4434, doi:10.1074/JBC.M117.818526.
235. Stiebler, A.C.; Freitag, J.; Schink, K.O.; Stehlik, T.; Tillmann, B.A.M.; Ast, J.; Bölker, M. Ribosomal Readthrough at a Short UGA Stop Codon Context Triggers Dual Localization of Metabolic Enzymes in Fungi and Animals. *PLoS Genet.* **2014**, *10*, doi:10.1371/JOURNAL.PGEN.1004685.
236. Loughran, G.; Chou, M.Y.; Ivanov, I.P.; Jungreis, I.; Kellis, M.; Kiran, A.M.; Baranov, P. V.; Atkins, J.F. Evidence of efficient stop codon readthrough in four mammalian genes. *Nucleic Acids Res.* **2014**, *42*, 8928–8938, doi:10.1093/nar/gku608.
237. Jungreis, I.; Lin, M.F.; Spokony, R.; Chan, C.S.; Negre, N.; Victorsen, A.; White, K.P.; Kellis, M. Evidence of abundant stop codon readthrough in *Drosophila* and other metazoa. *Genome Res.* **2011**, *21*, 2096, doi:10.1101/GR.119974.110.
238. Lindblad-Toh, K.; Garber, M.; Zuk, O.; Lin, M.F.; Parker, B.J.; Washietl, S.; Kheradpour, P.; Ernst, J.; Jordan, G.; Mauceli, E.; et al. A high-resolution map of

- human evolutionary constraint using 29 mammals. *Nature* **2011**, *478*, 476, doi:10.1038/NATURE10530.
239. Namy, O.; Hatin, I.; Rousset, J.-P. Impact of the six nucleotides downstream of the stop codon on translation termination. *EMBO Rep.* **2001**, *2*, 787, doi:10.1093/EMBO-REPORTS/KVE176.
240. Manuvakhova, M.; Keeling, K.; Bedwell, D.M. *Aminoglycoside antibiotics mediate context-dependent suppression of termination codons in a mammalian translation system*; 2000;
241. Bidou, L.; Hatin, I.; Perez, N.; Allamand, V.; Panthier, J.-J.; Rousset, J.-P. Premature stop codons involved in muscular dystrophies show a broad spectrum of readthrough efficiencies in response to gentamicin treatment. *Gene Ther.* **2004**, *11*, 619–627, doi:10.1038/SJ.GT.3302211.
242. Ban, N.; Nissen, P.; Hansen, J.; Moore, P.B.; Steitz, T.A. The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution. *Science (80-. ).* **2000**, *289*, 905–920, doi:10.1126/SCIENCE.289.5481.905.
243. Conard, S.E.; Buckley, J.; Dang, M.; Bedwell, G.J.; Carter, R.L.; Khass, M.; Bedwell, D.M. Identification of eRF1 residues that play critical and complementary roles in stop codon recognition. *RNA* **2012**, *18*, 1210, doi:10.1261/RNA.031997.111.
244. Tate, W.P.; Cridge, A.G.; Brown, C.M. 'Stop' in protein synthesis is modulated with exquisite subtlety by an extended RNA translation signal. *Biochem. Soc. Trans.* **2018**, *46*, 1615–1625, doi:10.1042/BST20180190.
245. Bonetti, B.; Fu, L.; Moon, J.; Bedwell, D.M. The efficiency of translation termination is determined by a synergistic interplay between upstream and downstream sequences in *Saccharomyces cerevisiae*. *J. Mol. Biol.* **1995**, *251*, 334–345, doi:10.1006/jmbi.1995.0438.
246. Cassan, M.; Rousset, J.-P. UAG readthrough in mammalian cells: Effect of upstream and downstream stop codon contexts reveal different signals. *BMC Mol. Biol.* **2001**, *2*, 3, doi:10.1186/1471-2199-2-3.
247. McCaughan, K.K.; Brown, C.M.; Dalphin, M.E.; Berry, M.J.; Tate, W.P. Translational termination efficiency in mammals is influenced by the base following the stop codon. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 5431, doi:10.1073/PNAS.92.12.5431.
248. Brown, A.; Shao, S.; Murray, J.; Hegde, R.; Ramakrishnan, V. Structural basis for stop codon recognition in eukaryotes. *Nature* **2015**, *524*, 493–496, doi:10.1038/NATURE14896.
249. Matheisl, S.; Berninghausen, O.; Becker, T.; Beckmann, R. Structure of a human translation termination complex. *Nucleic Acids Res.* **2015**, *43*, 8615–8626,

doi:10.1093/NAR/GKV909.

250. Cridge, A.G.; Crowe-McAuliffe, C.; Mathew, S.F.; Tate, W.P. Eukaryotic translational termination efficiency is influenced by the 3' nucleotides within the ribosomal mRNA channel. *Nucleic Acids Res.* **2018**, *46*, 1927, doi:10.1093/NAR/GKX1315.
251. Floquet, C.; Hatin, I.; Rousset, J.-P.; Bidou, L. Statistical Analysis of Readthrough Levels for Nonsense Mutations in Mammalian Cells Reveals a Major Determinant of Response to Gentamicin. *PLoS Genet* **2012**, *8*, 1002608, doi:10.1371/journal.pgen.1002608.
252. Harrell, L.; Melcher, U.; Atkins, J.F. Predominance of six different hexanucleotide recoding signals 3' of read-through stop codons. *Nucleic Acids Res.* **2002**, *30*, 2011, doi:10.1093/NAR/30.9.2011.
253. Tork, S.; Hatin, I.; Rousset, J.-P.; Fabret, C. The major 5' determinant in stop codon read-through involves two adjacent adenines. *Nucleic Acids Res.* **2004**, *32*, 415, doi:10.1093/NAR/GKH201.
254. Bertram, G.; Innes, S.; Minella, O.; Richardson, J.; Stansfield, I. Endless possibilities: translation termination and stop codon recognition. *Microbiol.* **2001**, *147*, 255–269, doi:10.1099/00221287-147-2-255.
255. Cohn, W.E.; Volkin, E. Nucleoside-5'-phosphates from ribonucleic acid. *Nature* **1951**, *167*, 483–484, doi:10.1038/167483a0.
256. Cantara, W.A.; Crain, P.F.; Rozenski, J.; McCloskey, J.A.; Harris, K.A.; Zhang, X.; Vendeix, F.A.P.; Fabris, D.; Agris, P.F. The RNA modification database, RNAMDB: 2011 update. *Nucleic Acids Res.* **2011**, *39*, D195, doi:10.1093/nar/gkq1028.
257. Machnicka, M.A.; Milanowska, K.; Oglou, O.O.; Purta, E.; Kurkowska, M.; Olchowik, A.; Januszewski, W.; Kalinowski, S.; Dunin-Horkawicz, S.; Rother, K.M.; et al. MODOMICS: A database of RNA modification pathways - 2013 update. *Nucleic Acids Res.* **2013**, *41*, D262, doi:10.1093/nar/gks1007.
258. Valadon, C.; Namy, O. The Importance of the Epi-Transcriptome in Translation Fidelity. *Non-Coding RNA* **2021**, *Vol. 7*, *Page 51* **2021**, *7*, 51, doi:10.3390/NCRNA7030051.
259. Prokhorova, I.; Altman, R.B.; Djumagulov, M.; Shrestha, J.P.; Urzhumtsev, A.; Ferguson, A.; Chang, C.-W.T.; Yusupov, M.; Blanchard, S.C.; Yusupova, G. Aminoglycoside interactions and impacts on the eukaryotic ribosome. *Proc. Natl. Acad. Sci.* **2017**, *114*, E10899–E10908, doi:10.1073/PNAS.1715501114.
260. Mort, M.; Ivanov, D.; Cooper, D.N.; Chuzhanova, N.A. A meta-analysis of nonsense mutations causing human genetic disease. *Hum. Mutat.* **2008**, *29*, 1037–1047, doi:10.1002/humu.20763.

261. Burke, J.F.; Mogg, A.E. Suppression of a nonsense mutation in mammalian cells in vivo by the aminoglycoside antibiotics G-418 and paromomycin. *Nucleic Acids Res.* **1985**, *13*, 6265–6272, doi:10.1093/nar/13.17.6265.
262. Kerem, E. Pharmacologic therapy for stop mutations: how much CFTR activity is enough? *Curr. Opin. Pulm. Med.* **2004**, *10*, 547–552, doi:10.1097/01.mcp.0000141247.22078.46.
263. Ramalho, A.S.; Beck, S.; Meyer, M.; Penque, D.; Cutting, G.R.; Amaral, M.D. Five percent of normal cystic fibrosis transmembrane conductance regulator mRNA ameliorates the severity of pulmonary disease in cystic fibrosis. *Am. J. Respir. Cell Mol. Biol.* **2002**, *27*, 619–627, doi:10.1165/rcmb.2001-0004OC.
264. Howard, M.; Frizzell, R.A.; Bedwell, D.M. Aminoglycoside antibiotics restore CFTR function by overcoming premature stop mutations. *Nat. Med.* **1996**, *2*, 467–469, doi:10.1038/nm0496-467.
265. Xue, X.; Mutyam, V.; Tang, L.; Biswas, S.; Du, M.; Jackson, L.A.; Dai, Y.; Belakhov, V.; Shalev, M.; Chen, F.; et al. Synthetic aminoglycosides efficiently suppress cystic fibrosis transmembrane conductance regulator nonsense mutations and are enhanced by ivacaftor. *Am. J. Respir. Cell Mol. Biol.* **2014**, *50*, 805–816, doi:10.1165/rcmb.2013-0282OC.
266. Arakawa, M.; Shiozuka, M.; Nakayama, Y.; Hara, T.; Hamada, M.; Kondo, S.; Ikeda, D.; Takahashi, Y.; Sawa, R.; Nonomura, Y.; et al. Negamycin Restores Dystrophin Expression in Skeletal and Cardiac Muscles of mdx Mice. *J. Biochem.* **2003**, *134*, 751–758, doi:10.1093/jb/mvg203.
267. Du, L.; Damoiseaux, R.; Nahas, S.; Gao, K.; Hu, H.; Pollard, J.M.; Goldstine, J.; Jung, M.E.; Henning, S.M.; Bertoni, C.; et al. Nonaminoglycoside compounds induce readthrough of nonsense mutations. *J. Exp. Med.* **2009**, *206*, 2285–2297, doi:10.1084/jem.20081940.
268. Du, L.; Jung, M.E.; Damoiseaux, R.; Completo, G.; Fike, F.; Ku, J.M.; Nahas, S.; Piao, C.; Hu, H.; Gatti, R.A. A new series of small molecular weight compounds induce read through of all three types of nonsense mutations in the ATM gene. *Mol. Ther.* **2013**, *21*, 1653–1660, doi:10.1038/mt.2013.150.
269. Gonzalez-Hilarion, S.; Beghyn, T.; Jia, J.; Debreuck, N.; Berte, G.; Mamchaoui, K.; Mouly, V.; Gruenert, D.C.; Déprez, B.; Lejeune, F. Rescue of nonsense mutations by amlexanox in human cells. *Orphanet J. Rare Dis.* **2012**, *7*, doi:10.1186/1750-1172-7-58.
270. Moosajee, M.; Tracey-White, D.; Smart, M.; Weetall, M.; Torriano, S.; Kalatzis, V.; da Cruz, L.; Coffey, P.; Webster, A.R.; Welch, E. Functional rescue of REP1 following treatment with PTC124 and novel derivative PTC-414 in human choroideremia fibroblasts and the nonsensemediated zebrafish model. *Hum. Mol. Genet.* **2016**,

- 25, 3416–3431, doi:10.1093/hmg/ddw184.
271. Mutyam, V.; Du, M.; Xue, X.; Keeling, K.M.; Lucile White, E.; Robert Bostwick, J.; Rasmussen, L.; Liu, B.; Mazur, M.; Hong, J.S.; et al. Discovery of clinically approved agents that promote suppression of cystic fibrosis transmembrane conductance regulator nonsense mutations. *Am. J. Respir. Crit. Care Med.* **2016**, *194*, 1092–1103, doi:10.1164/rccm.201601-0154OC.
272. Pibiri, I.; Lentini, L.; Melfi, R.; Tutone, M.; Baldassano, S.; Ricco Galluzzo, P.; Di Leonardo, A.; Pace, A. Rescuing the CFTR protein function: Introducing 1,3,4-oxadiazoles as translational readthrough inducing drugs. *Eur. J. Med. Chem.* **2018**, *159*, 126–142, doi:10.1016/j.ejmech.2018.09.057.
273. Welch, E.M.; Barton, E.R.; Zhuo, J.; Tomizawa, Y.; Friesen, W.J.; Trifillis, P.; Paushkin, S.; Patel, M.; Trotta, C.R.; Hwang, S.; et al. PTC124 targets genetic disorders caused by nonsense mutations. *Nature* **2007**, *447*, 87–91, doi:10.1038/nature05756.
274. Friesen, W.J.; Trotta, C.R.; Tomizawa, Y.; Zhuo, J.; Johnson, B.; Sierra, J.; Roy, B.; Weetall, M.; Hedrick, J.; Sheedy, J.; et al. The nucleoside analog cliticine is a potent and efficacious readthrough agent. *RNA* **2017**, *23*, 567–577, doi:10.1261/rna.060236.116.
275. Konstan, M.W.; VanDevanter, D.R.; Rowe, S.M.; Wilschanski, M.; Kerem, E.; Sermet-Gaudelus, I.; DiMango, E.; Melotti, P.; McIntosh, J.; Boeck, K. De Efficacy and safety of ataluren in patients with nonsense-mutation cystic fibrosis not receiving chronic inhaled aminoglycosides: The international, randomized, double-blind, placebo-controlled Ataluren Confirmatory Trial in Cystic Fibrosis (ACT CF). *J. Cyst. Fibros.* **2020**, *19*, 595–601, doi:10.1016/J.JCF.2020.01.007.
276. Wargo, K.A.; Edwards, J.D. Aminoglycoside-induced nephrotoxicity. *J. Pharm. Pract.* **2014**, *27*, 573–577, doi:10.1177/0897190014546836.
277. Selimoglu, E. Aminoglycoside-Induced Ototoxicity. *Curr. Pharm. Des.* **2006**, *13*, 119–126, doi:10.2174/138161207779313731.
278. Roy, B.; Friesen, W.J.; Tomizawa, Y.; Leszyk, J.D.; Zhuo, J.; Johnson, B.; Dakka, J.; Trotta, C.R.; Xue, X.; Mutyam, V.; et al. Ataluren stimulates ribosomal selection of near-cognate tRNAs to promote nonsense suppression. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 12508–12513, doi:10.1073/pnas.1605336113.
279. Trzaska, C.; Amand, S.; Bailly, C.; Leroy, C.; Marchand, V.; Duvernois-Berthet, E.; Saliou, J.M.; Benhabiles, H.; Werkmeister, E.; Chassat, T.; et al. 2,6-Diaminopurine as a highly potent corrector of UGA nonsense mutations. *Nat. Commun.* **2020**, *11*, doi:10.1038/s41467-020-15140-z.
280. Okada, N.; Harada, F.; Nishimura, S. Specific replacement of Q base in the anticodon of tRNA by guanine catalyzed by a cell-free extract of rabbit

- reticulocytes. *Nucleic Acids Res.* **1976**, *3*, 2593–2604, doi:10.1093/NAR/3.10.2593.
281. Shindo-Okada, N.; Okada, N.; Ohgi, T.; Goto, T.; Nishimura, S. Transfer ribonucleic acid guanine transglycosylase isolated from rat liver. *Biochemistry* **2002**, *19*, 395–400, doi:10.1021/BI00543A023.
282. Boland, C.; Hayes, P.; Santa-Maria, I.; Nishimura, S.; Kelly, V.P. Queuosine Formation in Eukaryotic tRNA Occurs via a Mitochondria-localized Heteromeric Transglycosylase \*. *J. Biol. Chem.* **2009**, *284*, 18218–18227, doi:10.1074/JBC.M109.002477.
283. Harada, F.; Nishimura, S. Possible Anticodon Sequences of tRNA<sup>His</sup>, tRNA<sup>Asn</sup>, and tRNA<sup>Asp</sup> from Escherichia coli B. Universal Presence of Nucleoside Q in the First Position of the Anticodons of These Transfer Ribonucleic Acids. *Biochemistry* **1972**, *11*, 301–308, doi:10.1021/bi00752a024.
284. Walden, T.; Reyniers, J.P.; Hiatt, V.; Farkas, W.R. Yeast Cells Cannot Incorporate Queuine into Their tRNA. *Proc. Soc. Exp. Biol. Med.* **1982**, *170*, 328–332, doi:10.3181/00379727-170-41438.
285. Okada, N.; Shindo-okada, N.; Nishimura, S. Isolation of mammalian tRNA<sup>Asp</sup> and tRNA<sup>Tyr</sup> by lectin-sepharose affinity column chromatography. *Nucleic Acids Res.* **1977**, *4*, 415–423, doi:10.1093/nar/4.2.415.
286. Costa, A.; De Barros, J.P.P.; Keith, G.; Baranowski, W.; Desgrès, J. Determination of queuosine derivatives by reverse-phase liquid chromatography for the hypomodification study of Q-bearing tRNAs from various mammal liver cells. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2004**, *801*, 237–247, doi:10.1016/j.jchromb.2003.11.022.
287. Farkas, W. Effect of diet on the queuosine family of tRNAs of germ-free mice - PubMed. *J Biol Chem.* **1980**.
288. Okada, N.; Nishimura, S. Enzymatic synthesis of Q nucleoside containing mannose in the anticodon of tRNA: Isolation of a novel mannosyltransferase from a cell-free extract of rat liver. *Nucleic Acids Res.* **1977**, *4*, 2931–2937, doi:10.1093/nar/4.8.2931.
289. Thumbs, P.; Ensfelder, T.T.; Hillmeier, M.; Wagner, M.; Heiss, M.; Scheel, C.; Schön, A.; Müller, M.; Michalakis, S.; Kellner, S.; et al. Synthesis of Galactosyl-Queuosine and Distribution of Hypermodified Q-Nucleosides in Mouse Tissues. *Angew. Chemie - Int. Ed.* **2020**, *59*, 12352–12356, doi:10.1002/anie.202002295.
290. Reyniers, J.; Pleasants, J.; Wostmann, B.; Katze, J.; Farkas, W. Administration of exogenous queuine is essential for the biosynthesis of the queuosine-containing transfer RNAs in the mouse - PubMed. *J Biol Chem.* **1981**.

291. Katze, J.R.; Gündüz, U.; Smith, D.L.; Cheng, C.S.; McCloskey, J.A. Evidence that the Nucleic Acid Base Queuine is Incorporated Intact into tRNA by Animal Cells. *Biochemistry* **1984**, *23*, 1171–1176, doi:10.1021/bi00301a022.
292. Grosjean, H.J.; De Henau, S.; Crotherst, D.M. *On the physical basis for ambiguity in genetic coding interactions (wobble rules/tRNA/codon-anticodon interactions/base pairing)*; 1978; Vol. 75;.
293. Meier, F.; Suter, B.; Grosjean, H.; Keith, G.; Kubli, E. Queuosine modification of the wobble base in tRNAHis influences “in vivo” decoding properties. *EMBO J.* **1985**, *4*, 823–827, doi:10.1002/j.1460-2075.1985.tb03704.x.
294. Zaborske, J.M.; Bauer DuMont, V.L.; Wallace, E.W.J.; Pan, T.; Aquadro, C.F.; Drummond, D.A. A Nutrient-Driven tRNA Modification Alters Translational Fidelity and Genome-wide Protein Coding across an Animal Genus. *PLoS Biol.* **2014**, *12*, doi:10.1371/journal.pbio.1002015.
295. BN, W.; GM, T. Activity of a transfer RNA modifying enzyme during the development of *Drosophila* and its relationship to the *su(s)* locus. *J. Mol. Biol.* **1973**, *74*, 635–651, doi:10.1016/0022-2836(73)90054-5.
296. Guzzi, N.; Cieśła, M.; Ngoc, P.C.T.; Lang, S.; Arora, S.; Dimitriou, M.; Pimková, K.; Sommarin, M.N.E.; Munita, R.; Lubas, M.; et al. Pseudouridylation of tRNA-Derived Fragments Steers Translational Control in Stem Cells. *Cell* **2018**, *173*, 1204–1216.e26, doi:10.1016/J.CELL.2018.03.008.
297. T, M.; WR, F. Effects of a diet deficient in tyrosine and queuine on germfree mice. *Biochem. Biophys. Res. Commun.* **1997**, *230*, 233–237, doi:10.1006/BBRC.1996.5768.
298. Rakovich, T.; Boland, C.; Bernstein, I.; Chikwana, V.M.; Iwata-Reuyl, D.; Kelly, V.P. Queuosine deficiency in eukaryotes compromises tyrosine production through increased tetrahydrobiopterin oxidation. *J. Biol. Chem.* **2011**, *286*, 19354–19363, doi:10.1074/jbc.M111.219576.
299. Zaringhalam, M.; Papavasiliou, F.N. Pseudouridylation meets next-generation sequencing. *Methods* **2016**, *107*, 63–72, doi:10.1016/J.YMETH.2016.03.001.
300. Schwartz, S.; Bernstein, D.A.; Mumbach, M.R.; Jovanovic, M.; Herbst, R.H.; León-Ricardo, B.X.; Engreitz, J.M.; Guttman, M.; Satija, R.; Lander, E.S.; et al. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **2014**, *159*, 148–162, doi:10.1016/j.cell.2014.08.028.
301. Kierzek, E.; Malgowska, M.; Lisowiec, J.; Turner, D.H.; Gdaniec, Z.; Kierzek, R. The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Res.* **2014**, *42*, 3492, doi:10.1093/NAR/GKT1330.

302. Lorenz, C.; Lünse, C.E.; Mörl, M. tRNA Modifications: Impact on Structure and Thermal Adaptation. *Biomolecules* **2017**, *7*, doi:10.3390/BIOM7020035.
303. Rintala-Dempsey, A.C.; Kothe, U. Eukaryotic stand-alone pseudouridine synthases – RNA modifying enzymes and emerging regulators of gene expression? *RNA Biol.* **2017**, *14*, 1185, doi:10.1080/15476286.2016.1276150.
304. Behm-Ansmant, I.; Urban, A.; Ma, X.; Yu, Y.T.; Motorin, Y.; Branlant, C. The *Saccharomyces cerevisiae* U2 snRNA:pseudouridine-synthase Pus7p is a novel multisite-multisubstrate RNA:ψ-synthase also acting on tRNAs. *RNA* **2003**, *9*, 1371–1382, doi:10.1261/rna.5520403.
305. Assi, H.A.; Shi, H.; Liu, B.; Clay, M.; Erharter, K.; Kreutz, C.; Holley, C.; Al-Hashimi, H. 2'- O -methylation alters the RNA secondary structural ensemble. *bioRxiv* **2020**, 2020.05.28.121996, doi:10.1101/2020.05.28.121996.
306. Blanchet, S.; Cornu, D.; Hatin, I.; Grosjean, H.; Bertin, P.; Namy, O. Deciphering the reading of the genetic code by near-cognate tRNA. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 3018–3023, doi:10.1073/pnas.1715578115.
307. Addepalli, B.; Limbach, P.A. Pseudouridine in the Anticodon of *Escherichia coli* tRNA<sup>Tyr</sup>(QΨA) Is Catalyzed by the Dual Specificity Enzyme RluF. *J. Biol. Chem.* **2016**, *291*, 22327, doi:10.1074/JBC.M116.747865.
308. Cui, Q.; Yin, K.; Zhang, X.; Ye, P.; Chen, X.; Chao, J.; Meng, H.; Wei, J.; Roeth, D.; Li, L.; et al. Targeting PUS7 suppresses tRNA pseudouridylation and glioblastoma tumorigenesis. *Nat. Cancer* **2021**, *2*, 932–949, doi:10.1038/s43018-021-00238-0.
309. Ayadi, L.; Galvanin, A.; Pichot, F.; Marchand, V.; Motorin, Y. RNA ribose methylation (2'-O-methylation): Occurrence, biosynthesis and biological functions. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2019**, *1862*, 253–269, doi:10.1016/J.BBAGRM.2018.11.009.
310. Decatur, W.A.; Fournier, M.J. rRNA modifications and ribosome function. *Trends Biochem. Sci.* **2002**, *27*, 344–351.
311. Baudin-Baillieu, A.; Namy, O. *Saccharomyces cerevisiae*, a powerful model for studying rRNA modifications and their effects on translation fidelity. *Int. J. Mol. Sci.* **2021**, *22*.
312. Dimitrova, D.G.; Teyssset, L.; Carré, C. RNA 2'-O-Methylation (Nm) Modification in Human Diseases. *Genes (Basel)*. **2019**, *10*, doi:10.3390/GENES10020117.
313. Kurth, H.M.; Mochizuki, K. 2'-O-methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in *Tetrahymena*. *RNA* **2009**, *15*, 675, doi:10.1261/RNA.1455509.
314. Sproat, B.S.; Lamond, A.I.; Beijer, B.; Neuner, P.; Ryder, U. Highly efficient chemical

- synthesis of 2'-O-methyloligoribonucleotides and tetrabiotinylated derivatives; novel probes that are resistant to degradation by RNA or DNA specific nucleases. *Nucleic Acids Res.* **1989**, *17*, 3373, doi:10.1093/NAR/17.9.3373.
315. Byszewska, M.; Śmietański, M.; Purta, E.; Bujnicki, J.M. RNA methyltransferases involved in 5' cap biosynthesis. *RNA Biol.* **2014**, *11*, 1597, doi:10.1080/15476286.2015.1004955.
316. Kumar, S.; Mapa, K.; Maiti, S. Understanding the Effect of Locked Nucleic Acid and 2'-O-Methyl Modification on the Hybridization Thermodynamics of a miRNA-mRNA Pair in the Presence and Absence of AfPwi Protein. *Biochemistry* **2014**, *53*, 1607–1615, doi:10.1021/BI401677D.
317. Yildirim, I.; Kierzek, E.; Kierzek, R.; Schatz, G.C. Interplay of LNA and 2'-O-Methyl RNA in the Structure and Thermodynamics of RNA Hybrid Systems: A Molecular Dynamics Study Using the Revised AMBER Force Field and Comparison with Experimental Results. *J. Phys. Chem. B* **2014**, *118*, 14177–14187, doi:10.1021/JP506703G.
318. Inoue, H.; Hayase, Y.; Imura, A.; Iwai, S.; Miura, K.; Ohtsuka, E. Synthesis and hybridization studies on two complementary nona(2'-O-methyl)ribonucleotides. *Nucleic Acids Res.* **1987**, *15*, 6131, doi:10.1093/NAR/15.15.6131.
319. Kawai, G.; Yamamoto, Y.; Kamimura, T.; Masegi, T.; Sekine, M.; Hata, T.; Iimori, T.; Watanabe, T.; Miyazawa, T.; Yokoyama, S. Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2'-hydroxyl group. *Biochemistry* **2002**, *31*, 1040–1046, doi:10.1021/BI00119A012.
320. Majlessi, M.; Nelson, N.C.; Becker, M.M. Advantages of 2'-O-methyl oligoribonucleotide probes for detecting RNA targets. *Nucleic Acids Res.* **1998**, *26*, 2224, doi:10.1093/NAR/26.9.2224.
321. Tsourkas, A.; Behlke, M.A.; Bao, G. Hybridization of 2'-O-methyl and 2-deoxy molecular beacons to RNA and DNA targets. *Nucleic Acids Res.* **2002**, *30*, 5168–5174.
322. Lebars, I.; Legrand, P.; Aimé, A.; Pinaud, N.; Fribourg, S.; Primo, C. Di Exploring TAR-RNA aptamer loop-loop interaction by X-ray crystallography, UV spectroscopy and surface plasmon resonance. *Nucleic Acids Res.* **2008**, *36*, 7146, doi:10.1093/NAR/GKN831.
323. Hou, Y.-M.; Zhang, X.; Holland, J.A.; Davis, D.R. An important 2'-OH group for an RNA-protein interaction. *Nucleic Acids Res.* **2001**, *29*, 976, doi:10.1093/NAR/29.4.976.
324. Lacoux, C.; Marino, D. Di; Boyl, P.P.; Zalfa, F.; Yan, B.; Ciotti, M.T.; Falconi, M.;

- Urlaub, H.; Achsel, T.; Mouglin, A.; et al. BC1-FMRP interaction is modulated by 2'-O-methylation: RNA-binding activity of the tudor domain and translational regulation at synapses. *Nucleic Acids Res.* **2012**, *40*, 4086, doi:10.1093/NAR/GKR1254.
325. Jones, S.; Daley, D.T.A.; Luscombe, N.M.; Berman, H.M.; Thornton, J.M. Protein–RNA interactions: a structural analysis. *Nucleic Acids Res.* **2001**, *29*, 943, doi:10.1093/NAR/29.4.943.
326. Treger, M.; Westhof, E. Statistical analysis of atomic contacts at RNA–protein interfaces. *J. Mol. Recognit.* **2001**, *14*, 199–214, doi:10.1002/JMR.534.
327. Lo Monaco, P.; Marcel, V.; Diaz, J.J.; Catez, F. 2'-O-methylation of ribosomal RNA: Towards an epitranscriptomic control of translation? *Biomolecules* **2018**, *8*.
328. Gautier, T.; Bergès, T.; Tollervey, D.; Hurt, E. Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. *Mol. Cell. Biol.* **1997**, *17*, 7088, doi:10.1128/MCB.17.12.7088.
329. Lafontaine, D.L.J. Noncoding RNAs in eukaryotic ribosome biogenesis and function. *Nat. Struct. Mol. Biol.* **2014**, *221* **2015**, *22*, 11–19, doi:10.1038/NSMB.2939.
330. Sergeeva, O. V.; Bogdanov, A.A.; Sergiev, P. V. What do we know about ribosomal RNA methylation in *Escherichia coli*? *Biochimie* **2015**, *117*, 110–118, doi:10.1016/J.BIOCHI.2014.11.019.
331. Marcel, V.; Kielbassa, J.; Marchand, V.; Natchiar, K.S.; Paraqindes, H.; Nguyen Van Long, F.; Ayadi, L.; Bourguignon-Igel, V.; Lo Monaco, P.; Monchiet, D.; et al. Ribosomal RNA 2'-O-methylation as a novel layer of inter-tumour heterogeneity in breast cancer. *NAR Cancer* **2020**, *2*, doi:10.1093/NARCAN/ZCAA036.
332. Schimmang, T.; Tollervey, D.; Kern, H.; Frank, R.; Hurt, E.C. A yeast nucleolar protein related to mammalian fibrillarin is associated with small nucleolar RNA and is essential for viability. *EMBO J.* **1989**, *8*, 4015.
333. Newton, K.; Petfalski, E.; Tollervey, D.; Cáceres, J.F. Fibrillarin Is Essential for Early Development and Required for Accumulation of an Intron-Encoded Small Nucleolar RNA in the Mouse. *Mol. Cell. Biol.* **2003**, *23*, 8519, doi:10.1128/MCB.23.23.8519-8527.2003.
334. Bouffard, S.; Dambroise, E.; Brombin, A.; Lempereur, S.; Hatin, I.; Simion, M.; Corre, R.; Bourrat, F.; Joly, J.S.; Jamen, F. Fibrillarin is essential for S-phase progression and neuronal differentiation in zebrafish dorsal midbrain and retina. *Dev. Biol.* **2018**, *437*, 1–16, doi:10.1016/J.YDBIO.2018.02.006.
335. Watanabe-Susaki, K.; Takada, H.; Enomoto, K.; Miwata, K.; Ishimine, H.; Intoh, A.;

- Ohtaka, M.; Nakanishi, M.; Sugino, H.; Asashima, M.; et al. Biosynthesis of Ribosomal RNA in Nucleoli Regulates Pluripotency and Differentiation Ability of Pluripotent Stem Cells. *Stem Cells* **2014**, *32*, 3099–3111, doi:10.1002/stem.1825.
336. Watkins, N.J.; Bohnsack, M.T. The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 397–414, doi:10.1002/wrna.117.
337. Tessarz, P.; Santos-Rosa, H.; Robson, S.C.; Sylvestersen, K.B.; Nelson, C.J.; Nielsen, M.L.; Kouzarides, T. Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nat. 2013 5057484* **2013**, *505*, 564–568, doi:10.1038/NATURE12819.
338. Marcel, V.; Ghayad, S.E.; Belin, S.; Therizols, G.; Morel, A.P.; Solano-González, E.; Vendrell, J.A.; Hacot, S.; Mertani, H.C.; Albaret, M.A.; et al. P53 Acts as a Safeguard of Translational Control by Regulating Fibrillarin and rRNA Methylation in Cancer. *Cancer Cell* **2013**, *24*, 318–330, doi:10.1016/j.ccr.2013.08.013.
339. Su, H.; Xu, T.; Ganapathy, S.; Shadfan, M.; Long, M.; Huang, T.H.M.; Thompson, I.; Yuan, Z.M. Elevated snoRNA biogenesis is essential in breast cancer. *Oncogene* **2014**, *33*, 1348–1358, doi:10.1038/onc.2013.89.
340. Zhou, F.; Liu, Y.; Rohde, C.; Pauli, C.; Gerloff, D.; Köhn, M.; Misiak, D.; Bäumer, N.; Cui, C.; Göllner, S.; et al. AML1-ETO requires enhanced C/D box snoRNA/RNP formation to induce self-renewal and leukaemia. *Nat. Cell Biol. 2017 197* **2017**, *19*, 844–855, doi:10.1038/NCB3563.
341. Krogh, N.; Asmar, F.; Côme, C.; Munch-Petersen, H.F.; Grønbaek, K.; Nielsen, H. Profiling of ribose methylations in ribosomal RNA from diffuse large B-cell lymphoma patients for evaluation of ribosomes as drug targets. *NAR Cancer* **2020**, *2*, doi:10.1093/NARCAN/ZCAA035.
342. Zhu, Y.; Pirnie, S.P.; Carmichael, G.G. High-throughput and site-specific identification of 2'-O-methylation sites using ribose oxidation sequencing (RibOxi-seq). *RNA* **2017**, *23*, 1303, doi:10.1261/RNA.061549.117.
343. Nachmani, D.; Bothmer, A.H.; Grisendi, S.; Mele, A.; Bothmer, D.; Lee, J.D.; Monteleone, E.; Cheng, K.; Zhang, Y.; Bester, A.C.; et al. Germline NPM1 mutations lead to altered rRNA 2'-O-methylation and cause dyskeratosis congenita. *Nat. Genet. 2019 5110* **2019**, *51*, 1518–1529, doi:10.1038/S41588-019-0502-Z.
344. Ruggero, D.; Pandolfi, P.P. Does the ribosome translate cancer? *Nat. Rev. Cancer 2003 33* **2003**, *3*, 179–192, doi:10.1038/NRC1015.
345. White, R.J. RNA polymerases I and III, growth control and cancer. *Nat. Rev. Mol. Cell Biol. 2005 61* **2005**, *6*, 69–78, doi:10.1038/NRM1551.
346. Truitt, M.L.; Ruggero, D. New frontiers in translational control of the cancer

- genome. *Nat. Rev. Cancer* **2016**, *16*, 288, doi:10.1038/NRC.2016.27.
347. Koh, C.M.; Gurel, B.; Sutcliffe, S.; Aryee, M.J.; Schultz, D.; Iwata, T.; Uemura, M.; Zeller, K.I.; Anele, U.; Zheng, Q.; et al. Alterations in Nucleolar Structure and Gene Expression Programs in Prostatic Neoplasia Are Driven by the MYC Oncogene. *Am. J. Pathol.* **2011**, *178*, 1824–1834, doi:10.1016/J.AJP.2010.12.040.
348. Schlosser, I.; Hölzel, M.; Mürnseer, M.; Burtscher, H.; Weidle, U.H.; Eick, D. A role for c-Myc in the regulation of ribosomal RNA processing. *Nucleic Acids Res.* **2003**, *31*, 6148–6156, doi:10.1093/NAR/GKG794.
349. Coller, H.A.; Grandori, C.; Tamayo, P.; Colbert, T.; Lander, E.S.; Eisenman, R.N.; Golub, T.R. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3260, doi:10.1073/PNAS.97.7.3260.
350. Erales, J.; Marchand, V.; Panthu, B.; Gillot, S.; Belin, S.; Ghayad, S.E.; Garcia, M.; Laforêts, F.; Marcel, V.; Baudin-Baillieu, A.; et al. Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 12934–12939, doi:10.1073/pnas.1707674114.
351. Natchiar, S.K.; Myasnikov, A.G.; Kratzat, H.; Hazemann, I.; Klaholz, B.P. Visualization of chemical modifications in the human 80S ribosome structure. *Nature* **2017**, *551*, 472–477, doi:10.1038/nature24482.
352. Bugaud, O.; Barbier, N.; Chommy, H.; Fiszman, N.; Gall, A. Le; Dulin, D.; Saguy, M.; Westbrook, N.; Perronet, K.; Namy, O. Kinetics of CrPV and HCV IRES-mediated eukaryotic translation using single-molecule fluorescence microscopy. *RNA* **2017**, *23*, 1626, doi:10.1261/RNA.061523.117.
353. Motorin, Y.; Helm, M. tRNA Stabilization by Modified Nucleotides. *Biochemistry* **2010**, *49*, 4934–4944, doi:10.1021/BI100408Z.
354. Chen, X.; Zaro, J.L.; Shen, W.C. Fusion Protein Linkers: Property, Design and Functionality. *Adv. Drug Deliv. Rev.* **2013**, *65*, 1357, doi:10.1016/J.ADDR.2012.09.039.
355. Bidou, L.; Stahl, G.; Hatin, I.; Namy, O.; Rousset, J.P.; Farabaugh, P.J. Nonsense-mediated decay mutants do not affect programmed -1 frameshifting. *RNA* **2000**, *6*, 952, doi:10.1017/S1355838200000443.
356. Baudin-Baillieu, A.; Hatin, I.; Legendre, R.; Namy, O. Translation Analysis at the Genome Scale by Ribosome Profiling. *Methods Mol. Biol.* **2016**, *1361*, 105–124, doi:10.1007/978-1-4939-3079-1\_7.
357. François, P.; Arbes, H.; Demais, S.; Baudin-Baillieu, A.; Namy, O. RiboDoc: A Docker-based package for ribosome profiling analysis. *Comput. Struct.*

- Biotechnol. J.* **2021**, *19*, 2851–2860, doi:10.1016/J.CSBJ.2021.05.014.
358. Zhang, W.; Xu, R.; Matuszek, Z.; Cai, Z.; Pan, T. Detection and quantification of glycosylated queuosine modified tRNAs by acid denaturing and APB gels. *RNA* **2020**, *26*, 1291–1298, doi:10.1261/RNA.075556.120.
359. Han, L.; Phizicky, E.M. A rationale for tRNA modification circuits in the anticodon loop. *RNA* **2018**, *24*, 1277, doi:10.1261/RNA.067736.118.
360. Ingolia, N.T. Ribosome Footprint Profiling of Translation throughout the Genome. *Cell* **2016**, *165*, 22, doi:10.1016/J.CELL.2016.02.066.
361. Geslain, R.; Pan, T. Functional analysis of human tRNA isodecoders. *J. Mol. Biol.* **2010**, *396*, 821, doi:10.1016/J.JMB.2009.12.018.
362. Torres, A.G.; Rodríguez-Escribà, M.; Marcet-Houben, M.; Santos Vieira, H.G.; Camacho, N.; Catena, H.; Murillo Recio, M.; Rafels-Ybern, À.; Reina, O.; Torres, F.M.; et al. Human tRNAs with inosine 34 are essential to efficiently translate eukarya-specific low-complexity proteins. *Nucleic Acids Res.* **2021**, *49*.
363. Beznosková, P.; Bidou, L.; Namy, O.; Valášek, L.S. Increased expression of tryptophan and tyrosine tRNAs elevates stop codon readthrough of reporter systems in human cell lines. *Nucleic Acids Res.* **2021**, *49*, 5202, doi:10.1093/NAR/GKAB315.
364. Floquet, C.; Rousset, J.P.; Bidou, L. La réactivation par translecture du gène p53 possédant une mutation non-sens induit l'apoptose de cellules cancéreuses. *médecine/sciences* **2011**, *27*, 585–586, doi:10.1051/MEDSCI/2011276008.
365. Fergus, C.; Barnes, D.; Alqasem, M.A.; Kelly, V.P. The Queuine Micronutrient: Charting a Course from Microbe to Man. *Nutrients* **2015**, *7*, 2897, doi:10.3390/NU7042897.
366. Decatur, W.A.; Schnare, M.N. Different Mechanisms for Pseudouridine Formation in Yeast 5S and 5.8S rRNAs. *Mol. Cell. Biol.* **2008**, *28*, 3089, doi:10.1128/MCB.01574-07.
367. Penzo, M.; Carnicelli, D.; Montanaro, L.; Brigotti, M. A reconstituted cell-free assay for the evaluation of the intrinsic activity of purified human ribosomes. *Nat. Protoc.* **2016**, *11*, 1309–1325, doi:10.1038/nprot.2016.072.
368. Montero, H.; Pérez-Gil, G.; Sampieri, C.L. Eukaryotic initiation factor 4A (eIF4A) during viral infections. *Virus Genes* **2019**, *55*, 267–273, doi:10.1007/S11262-019-01641-7/FIGURES/3.
369. Cencica, R.; Pelletier, J. Hippuristanol - A potent steroid inhibitor of eukaryotic initiation factor 4A. *Translation* **2016**, *4*, doi:10.1080/21690731.2015.1137381.
370. Struwe, W.B.; Robinson, C. V. Relating glycoprotein structural heterogeneity to

- function – insights from native mass spectrometry. *Curr. Opin. Struct. Biol.* **2019**, *58*, 241, doi:10.1016/J.SBI.2019.05.019.
371. Weng, T.Y.; Chiu, W.T.; Liu, H.S.; Cheng, H.C.; Shen, M.R.; Mount, D.B.; Chou, C.Y. Glycosylation regulates the function and membrane localization of KCC4. *Biochim. Biophys. Acta - Mol. Cell Res.* **2013**, *1833*, 1133–1146, doi:10.1016/J.BBAMCR.2013.01.018.
372. Flynn, R.A.; Pedram, K.; Malaker, S.A.; Batista, P.J.; Smith, B.A.H.; Johnson, A.G.; George, B.M.; Majzoub, K.; Villalta, P.W.; Carette, J.E.; et al. Small RNAs are modified with N-glycans and displayed on the surface of living cells. *Cell* **2021**, *184*, 3109–3124.e22, doi:10.1016/J.CELL.2021.04.023/ATTACHMENT/3B503342-FA9B-4309-87D5-732F8708C5A5/MMC5.XLSX.
373. Honda, S.; Shigematsu, M.; Morichika, K.; Telonis, A.G.; Kirino, Y. Four-leaf clover qRT-PCR: A convenient method for selective quantification of mature tRNA. *RNA Biol.* **2015**, *12*, 501, doi:10.1080/15476286.2015.1031951.
374. Johansson, M.; Chen, J.; Tsai, A.; Kornberg, G.; Puglisi, J.D. Sequence-dependent elongation dynamics on macrolide-bound ribosomes. *Cell Rep.* **2014**, *7*, 1534–1546, doi:10.1016/J.CELREP.2014.04.034/ATTACHMENT/0B284E86-3D7B-4089-BBE7-90453E35C15C/MMC1.PDF.
375. Morisaki, T.; Stasevich, T.J. Quantifying Single mRNA Translation Kinetics in Living Cells. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a032078, doi:10.1101/CSHPERSPECT.A032078.
376. Schott, J.; Reitter, S.; Lindner, D.; Grosser, J.; Bruer, M.; Shenoy, A.; Geiger, T.; Mathes, A.; Dobрева, G.; Stoecklin, G. Nascent Ribo-Seq measures ribosomal loading time and reveals kinetic impact on ribosome density. *Nat. Methods* **2021**, *18*, 1068–1074, doi:10.1038/s41592-021-01250-z.
377. Sharma, A.K.; Sormanni, P.; Ahmed, N.; Ciryam, P.; Friedrich, U.A.; Kramer, G.; O'Brien, E.P. A chemical kinetic basis for measuring translation initiation and elongation rates from ribosome profiling data. *PLoS Comput. Biol.* **2019**, *15*, doi:10.1371/JOURNAL.PCBI.1007070.
378. Szavits-Nossan, J.; Ciandrini, L. Inferring efficiency of translation initiation and elongation from ribosome profiling. *Nucleic Acids Res.* **2020**, *48*, 9478, doi:10.1093/NAR/GKAA678.

## IX. Annexes

### Annexe 1 : références de la revue

1. Cohn, W.E.; Volkin, E. Nucleoside-5<sup>0</sup>-phosphates from ribonucleic acid. *Nature* **1951**, *167*, 483–484. [[CrossRef](#)]
2. Cantara, W.A.; Crain, P.F.; Rozenski, J.; McCloskey, J.A.; Harris, K.A.; Zhang, X.; Vendeix, F.A.P.; Fabris, D.; Agris, P.F. The RNA modification database, RNAMDB: 2011 update. *Nucleic Acids Res.* **2011**, *39*, D195. [[CrossRef](#)] [[PubMed](#)]
3. Machnicka, M.A.; Milanowska, K.; Oglou, O.O.; Purta, E.; Kurkowska, M.; Olchowik, A.; Januszewski, W.; Kalinowski, S.; Dunin-Horkawicz, S.; Rother, K.M.; et al. MODOMICS: A database of RNA modification pathways - 2013 update. *Nucleic Acids Res.* **2013**, *41*, D262. [[CrossRef](#)]
4. Song, J.; Yi, C. Chemical Modifications to RNA: A New Layer of Gene Expression Regulation. *ACS Chem. Biol.* **2017**, *12*, 316–325. [[CrossRef](#)] [[PubMed](#)]
5. Song, X.; Nazar, R.N. Modification of rRNA as a quality control mechanism in ribosome biogenesis. *FEBS Lett.* **2002**, *523*, 182–186. [[CrossRef](#)]
6. Agris, P.F. Decoding the genome: A modified view. *Nucleic Acids Res.* **2004**, *32*, 223–238. [[CrossRef](#)] [[PubMed](#)]
7. Ranjan, N.; Leidel, S.A. The epitranscriptome in translation regulation: mRNA and tRNA modifications as the two sides of the same coin? *FEBS Lett.* **2019**, *593*, 1483–1493. [[CrossRef](#)]
8. Kadumuri, R.V.; Janga, S.C. Epitranscriptomic Code and Its Alterations in Human Disease. *Trends Mol. Med.* **2018**, *24*, 886–903. [[CrossRef](#)]
9. Jonkhout, N.; Tran, J.; Smith, M.A.; Schonrock, N.; Mattick, J.S.; Novoa, E.M. The RNA modification landscape in human disease. *RNA* **2017**, *23*, 1754–1769. [[CrossRef](#)]
10. Hsu, P.J.; Shi, H.; He, C. Epitranscriptomic influences on development and disease. *Genome Biol.* **2017**, *18*, 1–9. [[CrossRef](#)]
11. Haran, V.; Lenka, N. Deciphering the Epitranscriptomic Signatures in Cell Fate Determination and Development. *Stem Cell Rev. Reports* **2019**, *15*, 474–496. [[CrossRef](#)]
12. Barbieri, I.; Kouzarides, T. Role of RNA modifications in cancer. *Nat. Rev. Cancer* **2020**, *20*, 303–322. [[CrossRef](#)]
13. Kapp, L.D.; Lorsch, J.R. The molecular mechanics of eukaryotic translation. *Annu. Rev. Biochem.* **2004**, *73*, 657–704. [[CrossRef](#)]
14. Pelletier, J.; Sonenberg, N. The Organizing Principles of Eukaryotic Ribosome Recruitment. *Annu. Rev. Biochem.* **2019**, *88*, 307–335. [[CrossRef](#)] [[PubMed](#)]
15. Jackson, R.; Hellen, C.; Pestova, T. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 113–127. [[CrossRef](#)] [[PubMed](#)]
16. Hinnebusch, A. Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends Biochem. Sci.* **2017**, *42*, 589–611. [[CrossRef](#)]
17. Hashem, Y.; Frank, J. The Jigsaw Puzzle of mRNA Translation Initiation in Eukaryotes: A Decade of Structures Unraveling the Mechanics of the Process. *Annu. Rev. Biophys.* **2018**, *47*, 125–151. [[CrossRef](#)]
18. Pestova, T.; Lomakin, I.; Lee, J.; Choi, S.; Dever, T.; Hellen, C. The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* **2000**, *403*, 332–335. [[CrossRef](#)] [[PubMed](#)]
19. Wang, J.; Wang, J.; Shin, B.; Kim, J.; Dever, T.; Puglisi, J.; Fernández, I. Structural basis for the transition from translation initiation to elongation by an 80S-eIF5B complex. *Nat. Commun.* **2020**, *11*, 1–18. [[CrossRef](#)]
20. Blanchet, S.; Rowe, M.; Von der Haar, T.; Fabret, C.; Demais, S.; Howard, M.; Namy, O. New insights into stop codon recognition by eRF1. *Nucleic Acids Res.* **2015**, *43*, 3298–3308. [[CrossRef](#)]
21. Brown, A.; Shao, S.; Murray, J.; Hegde, R.; Ramakrishnan, V. Structural basis for stop codon recognition in eukaryotes. *Nature* **2015**, *524*, 493–496. [[CrossRef](#)]
22. Young, D.; Guydosh, N.; Zhang, F.; Hinnebusch, A.; Green, R. Rli1/ABCE1 Recycles Terminating Ribosomes and Controls Translation Reinitiation in 3<sup>0</sup>UTRs In Vivo. *Cell* **2015**, *162*, 872–884. [[CrossRef](#)] [[PubMed](#)]
23. Dever, T.; Green, R. The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, 1–16. [[CrossRef](#)]
24. Mancera-Martínez, E.; Brito Querido, J.; Valasek, L.; Simonetti, A.; Hashem, Y. ABCE1: A special factor that orchestrates translation at the crossroad between recycling and initiation. *RNA Biol.* **2017**, *14*, 1279–1285. [[CrossRef](#)]
25. Loftfield, R.B. The frequency of errors in protein biosynthesis. *Biochem. J.* **1963**, *89*, 82–92. [[CrossRef](#)] [[PubMed](#)]
26. Kurland, C.; Gallant, J. Errors of heterologous protein expression. *Curr. Opin. Biotechnol.* **1996**, *7*, 489–493. [[CrossRef](#)]
27. Crick, F.H.C. Codon—anticodon pairing: The wobble hypothesis. *J. Mol. Biol.* **1966**, *19*, 548–555. [[CrossRef](#)]
28. Namy, O.; Rousset, J.P.; Naphine, S.; Brierley, I. Reprogrammed Genetic Decoding in Cellular Gene Expression. *Mol. Cell* **2004**, *13*, 157–168. [[CrossRef](#)]
29. Muramatsu, T.; Nishikawa, K.; Nemoto, F.; Kuchino, Y.; Nishimura, S.; Miyazawa, T.; Yokoyama, S. Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature* **1988**, *336*, 179–181. [[CrossRef](#)]

30. Sloan, K.E.; Warda, A.S.; Sharma, S.; Entian, K.D.; Lafontaine, D.L.J.; Bohnsack, M.T. Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol.* **2017**, *14*, 1138–1152. [[CrossRef](#)]
31. Tuorto, F.; Lyko, F. Genome recoding by tRNA modifications. *Open Biol.* **2016**, *6*, 160287. [[CrossRef](#)] [[PubMed](#)]
32. De Crécy-Lagard, V.; Boccaletto, P.; Mangleburg, C.G.; Sharma, P.; Lowe, T.M.; Leidel, S.A.; Bujnicki, J.M. Survey and summary: Matching tRNA modifications in humans to their known and predicted enzymes. *Nucleic Acids Res.* **2019**, *47*, 2143–2159. [[CrossRef](#)] [[PubMed](#)]
33. Schaffrath, R.; Leidel, S.A. Wobble uridine modifications—a reason to live, a reason to die?! *RNA Biol.* **2017**, *14*, 1209–1222. [[CrossRef](#)] [[PubMed](#)]
34. Liu, F.; Clark, W.; Luo, G.; Wang, X.; Fu, Y.; Wei, J.; Wang, X.; Hao, Z.; Dai, Q.; Zheng, G.; et al. ALKBH1-Mediated tRNA Demethylation Regulates Translation. *Cell* **2016**, *167*, 816–828.e16. [[CrossRef](#)] [[PubMed](#)]
35. Karlsborn, T.; Tükenmez, H.; Mahmud, A.K.M.F.; Xu, F.; Xu, H.; Byström, A.S. Elongator, a conserved complex required for wobble uridine modifications in Eukaryotes. *RNA Biol.* **2014**, *11*, 1519–1528. [[CrossRef](#)]
36. Bourgeois, G.; Létoquart, J.; van Tran, N.; Graille, M. Trm112, a protein activator of methyltransferases modifying actors of the eukaryotic translational apparatus. *Biomolecules* **2017**, *7*, 7. [[CrossRef](#)]
37. Songe-Møller, L.; van den Born, E.; Leihne, V.; Vågbo, C.B.; Kristoffersen, T.; Krokan, H.E.; Kirpekar, F.; Falnes, P.Ø.; Klungland, A. Mammalian ALKBH8 Possesses tRNA Methyltransferase Activity Required for the Biogenesis of Multiple Wobble Uridine Modifications Implicated in Translational Decoding. *Mol. Cell. Biol.* **2010**, *30*, 1814–1827. [[CrossRef](#)]
38. Noma, A.; Sakaguchi, Y.; Suzuki, T. Mechanistic characterization of the sulfur-relay system for eukaryotic 2-thiouridine biogenesis at tRNA wobble positions. *Nucleic Acids Res.* **2009**, *37*, 1335–1352. [[CrossRef](#)]
39. Blanchet, S.; Cornu, D.; Hatin, I.; Grosjean, H.; Bertin, P.; Namy, O. Deciphering the reading of the genetic code by near-cognate tRNA. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3018–3023. [[CrossRef](#)]
40. Patil, A.; Chan, C.T.Y.; Dyavaiah, M.; Rooney, J.P.; Dedon, P.C.; Begley, T.J. Translational infidelity-induced protein stress results from a deficiency in Trm9-catalyzed tRNA modifications. *RNA Biol.* **2012**, *9*, 990–1001. [[CrossRef](#)]
41. Johansson, M.J.O.; Esberg, A.; Huang, B.; Björk, G.R.; Byström, A.S. Eukaryotic Wobble Uridine Modifications Promote a Functionally Redundant Decoding System. *Mol. Cell. Biol.* **2008**, *28*, 3301–3312. [[CrossRef](#)]
42. Nedialkova, D.D.; Leidel, S.A. Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. *Cell* **2015**, *161*, 1606–1618. [[CrossRef](#)]
43. Hou, Y.M.; Gamper, H.; Yang, W. Post-transcriptional modifications to tRNA - A response to the genetic code degeneracy. *RNA* **2015**, *21*, 642–644. [[CrossRef](#)]
44. Licht, K.; Hartl, M.; Amman, F.; Anrather, D.; Janisiw, M.P.; Jantsch, M.F. Inosine induces context-dependent recoding and translational stalling. *Nucleic Acids Res.* **2019**, *47*, 3–14. [[CrossRef](#)]
45. Rafels-Ybern, À.; Torres, A.G.; Grau-Bove, X.; Ruiz-Trillo, I.; Ribas de Pouplana, L. Codon adaptation to tRNAs with Inosine modification at position 34 is widespread among Eukaryotes and present in two Bacterial phyla. *RNA Biol.* **2018**, *15*, 500–507. [[CrossRef](#)]
46. Gerber, A.P.; Keller, W. An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* **1999**, *286*, 1146–1149. [[CrossRef](#)]
47. Grosjean, H.; Auxilien, S.; Constantinesco, F.; Simon, C.; Corda, Y.; Becker, H.F.; Foiret, D.; Morin, A.; Jin, Y.X.; Fournier, M.; et al. Enzymatic conversion of adenosine to inosine and to N1-methylinosine in transfer RNAs: A review. *Biochimie* **1996**, *78*, 488–501. [[CrossRef](#)]
48. Torres, A.G.; Piñeyro, D.; Filonava, L.; Stracker, T.H.; Batlle, E.; Ribas de Pouplana, L. A-to-I editing on tRNAs: Biochemical, biological and evolutionary implications. *FEBS Lett.* **2014**, *588*, 4279–4286. [[CrossRef](#)] [[PubMed](#)]
49. Ramos, J.; Proven, M.; Halvardson, J.; Hagelskamp, F.; Kuchinskaya, E.; Phelan, B.; Bell, R.; Kellner, S.M.; Feuk, L.; Thuresson, A.C.; et al. Identification and rescue of a tRNA wobble inosine deficiency causing intellectual disability disorder. *RNA* **2020**, *26*, 1654–1666. [[CrossRef](#)] [[PubMed](#)]
50. Torres, A.G.; Rodríguez-Escribà, M.; Marcet-Houben, M.; Santos Vieira, H.G.; Camacho, N.; Catena, H.; Murillo Recio, M.; Rafels-Ybern, À.; Reina, O.; Torres, F.M.; et al. Human tRNAs with inosine 34 are essential to efficiently translate eukarya-specific low-complexity proteins. *Nucleic Acids Res.* **2021**, *49*, 7011–7034. [[CrossRef](#)] [[PubMed](#)]
51. Rafels-Ybern, À.; Torres, A.G.; Camacho, N.; Herencia-Roperro, A.; Frigolé, H.R.; Wulff, T.F.; Raboteg, M.; Bordons, A.; Grau-Bove, X.; Ruiz-Trillo, I.; et al. The expansion of inosine at the wobble position of tRNAs, and its role in the evolution of proteomes. *Mol. Biol. Evol.* **2019**, *36*, 650–662. [[CrossRef](#)]
52. Harada, F.; Nishimura, S. Possible Anticodon Sequences of tRNA<sup>His</sup>, tRNA<sup>Asn</sup>, and tRNA<sup>Asp</sup> from *Escherichia coli* B. Universal Presence of Nucleoside Q in the First Position of the Anticodons of These Transfer Ribonucleic Acids. *Biochemistry* **1972**, *11*, 301–308. [[CrossRef](#)]
53. Okada, N.; Shindo-okada, N.; Nishimura, S. Isolation of mammalian tRNA<sup>Asp</sup> and tRNA<sup>Tyr</sup> by lectin-sepharose affinity column chromatography. *Nucleic Acids Res.* **1977**, *4*, 415–423. [[CrossRef](#)]

54. Costa, A.; De Barros, J.P.P.; Keith, G.; Baranowski, W.; Desgrès, J. Determination of queuosine derivatives by reverse-phase liquid chromatography for the hypomodification study of Q-bearing tRNAs from various mammal liver cells. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2004**, *801*, 237–247. [[CrossRef](#)]
55. Nishimura, S. *Structure, Biosynthesis, and Function of Queuosine in Transfer RNA*; Elsevier: Amsterdam, The Netherlands, 1983.
56. Farkas, W. Effect of diet on the queuosine family of tRNAs of germ-free mice. *J. Biol. Chem.* **1980**, *255*, 6832–6835. [[CrossRef](#)]
57. Reyniers, J.; Pleasants, J.; Wostmann, B.; Katze, J.; Farkas, W. Administration of exogenous queuine is essential for the biosynthesis of the queuosine-containing transfer RNAs in the mouse. *J. Biol. Chem.* **1981**, *256*, 11591–11594. [[CrossRef](#)]
58. Katze, J.R.; Gündüz, U.; Smith, D.L.; Cheng, C.S.; McCloskey, J.A. Evidence that the Nucleic Acid Base Queuine is Incorporated Intact into tRNA by Animal Cells. *Biochemistry* **1984**, *23*, 1171–1176. [[CrossRef](#)]
59. Zaborske, J.M.; Bauer DuMont, V.L.; Wallace, E.W.J.; Pan, T.; Aquadro, C.F.; Drummond, D.A. A Nutrient-Driven tRNA Modification Alters Translational Fidelity and Genome-wide Protein Coding across an Animal Genus. *PLoS Biol.* **2014**, *12*. [[CrossRef](#)] [[PubMed](#)]
60. Powell, J.R.; Sezzi, E.; Moriyama, E.N.; Gleason, J.M.; Caccone, A. Analysis of a Shift in Codon Usage in *Drosophila*. In Proceedings of the Journal of Molecular Evolution. *J. Mol. Evol.* **2003**, *57*, S214–S225. [[CrossRef](#)]
61. Okada, N.; Nishimura, S. Enzymatic synthesis of Q nucleoside containing mannose in the anticodon of tRNA: Isolation of a novel mannosyltransferase from a cell-free extract of rat liver. *Nucleic Acids Res.* **1977**, *4*, 2931–2937. [[CrossRef](#)]
62. Magee, R.; Rigoutsos, I. On the expanding roles of tRNA fragments in modulating cell behavior. *Nucleic Acids Res.* **2020**, *48*, 9433–9448. [[CrossRef](#)] [[PubMed](#)]
63. Grosjean, H.J.; De Henau, S.; Crotherst, D.M. On the Physical Basis for Ambiguity in Genetic Coding Interactions. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 610–614. [[CrossRef](#)]
64. Morris, R.C.; Brown, K.G.; Elliott, M.S. The effect of queuosine on trna structure and function. *J. Biomol. Struct. Dyn.* **1999**, *16*, 757–774. [[CrossRef](#)] [[PubMed](#)]
65. Manickam, N.; Joshi, K.; Bhatt, M.J.; Farabaugh, P.J. Effects of tRNA modification on translational accuracy depend on intrinsic codon-anticodon strength. *Nucleic Acids Res.* **2015**, *44*, 1871–1881. [[CrossRef](#)] [[PubMed](#)]
66. Meier, F.; Suter, B.; Grosjean, H.; Keith, G.; Kubli, E. Queuosine modification of the wobble base in tRNAHis influences “in vivo” decoding properties. *EMBO J.* **1985**, *4*, 823–827. [[CrossRef](#)] [[PubMed](#)]
67. Brzezicha, B.; Schmidt, M.; Makołowska, I.; Jarmołowski, A.; Pienkowska, J.; Szweykowska-Kulin'ska, Z. Identification of human tRNA: m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA(CAA)Leu. *Nucleic Acids Res.* **2006**, *34*, 6034–6043. [[CrossRef](#)] [[PubMed](#)]
68. Strobel, M.C.; Abelson, J. Effect of intron mutations on processing and function of *Saccharomyces cerevisiae* SUP53 tRNA in vitro and in vivo. *Mol. Cell. Biol.* **1986**, *6*, 2663–2673. [[CrossRef](#)]
69. Arragain, S.; Handelman, S.K.; Forouhar, F.; Wei, F.Y.; Tomizawa, K.; Hunt, J.F.; Douki, T.; Fontecave, M.; Mulliez, E.; Atta, M. Identification of eukaryotic and prokaryotic methylthiotransferase for biosynthesis of 2-methylthio-N6-threonylcarbamoyladenine in tRNA. *J. Biol. Chem.* **2010**, *285*, 28425–28433. [[CrossRef](#)]
70. Durant, P.C.; Bajj, A.C.; Sundaram, M.; Kumar, R.K.; Davis, D.R. Structural Effects of Hypermodified Nucleosides in the *Escherichia coli* and Human tRNA<sup>Lys</sup> Anticodon Loop: The Effect of Nucleosides s2U, mcm5U, mcm5s2U, mnm5s2U, t6A, and ms2t6A. *Biochemistry* **2005**, *44*, 8078–8089. [[CrossRef](#)]
71. Wei, F.Y.; Suzuki, T.; Watanabe, S.; Kimura, S.; Kaitesuka, T.; Fujimura, A.; Matsui, H.; Atta, M.; Michiue, H.; Fontecave, M.; et al. Deficit of tRNA<sup>Lys</sup> modification by Cdkal1 causes the development of type 2 diabetes in mice. *J. Clin. Investig.* **2011**, *121*, 3598–3608. [[CrossRef](#)]
72. Lamichhane, T.N.; Mattijssen, S.; Maraia, R.J. Human Cells Have a Limited Set of tRNA Anticodon Loop Substrates of the tRNA Isopentenyltransferase TRIT1 Tumor Suppressor. *Mol. Cell. Biol.* **2013**, *33*, 4900–4908. [[CrossRef](#)]
73. U, S.; S, B.; N, F.-V. The modified base isopentenyladenosine and its derivatives in tRNA. *RNA Biol.* **2017**, *14*, 1197–1208. [[CrossRef](#)]
74. Lamichhane, T.N.; Blewett, N.H.; Crawford, A.K.; Cherkasova, V.A.; Iben, J.R.; Begley, T.J.; Farabaugh, P.J.; Maraia, R.J. Lack of tRNA Modification Isopentenyl-A37 Alters mRNA Decoding and Causes Metabolic Deficiencies in Fission Yeast. *Mol. Cell. Biol.* **2013**, *33*, 2918–2929. [[CrossRef](#)] [[PubMed](#)]
75. Spinola, M.; Falvella, F.S.; Galvan, A.; Pignatiello, C.; Leoni, V.P.; Pastorino, U.; Paroni, R.; Chen, S.; Skaug, V.; Haugen, A.; et al. Ethnic differences in frequencies of gene polymorphisms in the MYCL1 region and modulation of lung cancer patients' survival. *Lung Cancer* **2007**, *55*, 271–277. [[CrossRef](#)]
76. Yarham, J.W.; Lamichhane, T.N.; Pyle, A.; Mattijssen, S.; Baruffini, E.; Bruni, F.; Donnini, C.; Vassilev, A.; He, L.; Blakely, E.L.; et al. Defective i6A37 Modification of Mitochondrial and Cytosolic tRNAs Results from Pathogenic Mutations in TRIT1 and Its Substrate tRNA. *PLoS Genet.* **2014**, *10*, e1004424. [[CrossRef](#)]

77. De Zoysa, M.D.; Yu, Y.T. Posttranscriptional RNA Pseudouridylation. In *Enzymes*; Academic Press: Cambridge, MA, USA, 2017; Volume 41, pp. 151–167.
78. Davis, D.R.; Veltri, C.A.; Nielsen, L. An rna model system for investigation of pseudouridine stabilization of the codon-anticodon interaction in trna<sup>Lys</sup>, trna<sup>His</sup> and trna<sup>Tyr</sup>. *J. Biomol. Struct. Dyn.* **1998**, *15*, 1121–1132. [[CrossRef](#)]
79. Deb, I.; Popena, Ł.; Sarzyn´ska, J.; Małgowska, M.; Lahiri, A.; Gdaniec, Z.; Kierzek, R. Computational and NMR studies of RNA duplexes with an internal pseudouridine-adenosine base pair. *Sci. Rep.* **2019**, *9*, 1–13. [[CrossRef](#)] [[PubMed](#)]
80. Lecointe, F.; Namy, O.; Hatin, I.; Simos, G.; Rousset, J.P.; Grosjean, H. Lack of Pseudouridine 38/39 in the anticodon arm of yeast cytoplasmic tRNA decreases in vivo recoding efficiency. *J. Biol. Chem.* **2002**, *277*, 30445–30453. [[CrossRef](#)] [[PubMed](#)]
81. De Brouwer, A.P.M.; Abou Jamra, R.; Körtel, N.; Soyris, C.; Polla, D.L.; Safra, M.; Zisso, A.; Powell, C.A.; Rebelo-Guimar, P.; Dinges, N.; et al. Variants in PUS7 Cause Intellectual Disability with Speech Delay, Microcephaly, Short Stature, and Aggressive Behavior. *Am. J. Hum. Genet.* **2018**, *103*, 1045–1052. [[CrossRef](#)]
82. Shaheen, R.; Han, L.; Faqeih, E.; Ewida, N.; Alobeid, E.; Phizicky, E.M.; Alkuraya, F.S. A homozygous truncating mutation in PUS3 expands the role of tRNA modification in normal cognition. *Hum. Genet.* **2016**, *135*, 707–713. [[CrossRef](#)]
83. Chou, H.J.; Donnard, E.; Gustafsson, H.T.; Garber, M.; Rando, O.J. Transcriptome-wide Analysis of Roles for tRNA Modifications in Translational Regulation. *Mol. Cell* **2017**, *68*, 978–992.e4. [[CrossRef](#)]
84. Guy, M.P.; Phizicky, E.M. Conservation of an intricate circuit for crucial modifications of the tRNA<sup>Phe</sup> anticodon loop in eukaryotes. *RNA* **2015**, *21*, 61–74. [[CrossRef](#)]
85. Guy, M.P.; Podyma, B.M.; Preston, M.A.; Shaheen, H.H.; Krivos, K.L.; Limbach, P.A.; Hopper, A.K.; Phizicky, E.M. Yeast Trm7 interacts with distinct proteins for critical modifications of the tRNA<sup>Phe</sup> anticodon loop. *RNA* **2012**, *18*, 1921–1933. [[CrossRef](#)]
86. Guy, M.P.; Shaw, M.; Weiner, C.L.; Hobson, L.; Stark, Z.; Rose, K.; Kalscheuer, V.M.; Gecz, J.; Phizicky, E.M. Defects in tRNA Anticodon Loop 2'-O-Methylation Are Implicated in Nonsyndromic X-Linked Intellectual Disability due to Mutations in FTSJ1. *Hum. Mutat.* **2015**, *36*, 1176–1187. [[CrossRef](#)]
87. Han, L.; Marcus, E.; D'Silva, S.; Phizicky, E.M.S. cerevisiae Trm140 has two recognition modes for 3-methylcytidine modification of the anticodon loop of tRNA substrates. *RNA* **2017**, *23*, 406–419. [[CrossRef](#)] [[PubMed](#)]
88. Müller, M.; Hartmann, M.; Schuster, I.; Bender, S.; Thüring, K.L.; Helm, M.; Katze, J.R.; Nellen, W.; Lyko, F.; Ehrenhofer-Murray, A.E. Dynamic modulation of Dnmt2-dependent tRNA methylation by the micronutrient queuine. *Nucleic Acids Res.* **2015**, *43*, 10952–10962. [[CrossRef](#)] [[PubMed](#)]
89. Rubio, M.A.T.; Ragone, F.L.; Gaston, K.W.; Ibba, M.; Alfonzo, J.D. C to U editing stimulates A to I editing in the anticodon loop of a cytoplasmic threonyl tRNA in Trypanosoma brucei. *J. Biol. Chem.* **2006**, *281*, 115–120. [[CrossRef](#)] [[PubMed](#)]
90. Pan, Q.; Han, T.; Li, G. Novel insights into the roles of tRNA-derived small RNAs. *RNA Biol.* **2021**, 1–11. [[CrossRef](#)]
91. Goll, M.G.; Kirpekar, F.; Maggert, K.A.; Yoder, J.A.; Hsieh, C.L.; Zhang, X.; Golic, K.G.; Jacobsen, S.E.; Bestor, T.H. Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2. *Science* **2006**, *311*, 395–398. [[CrossRef](#)] [[PubMed](#)]
92. Schaefer, M.; Pollex, T.; Hanna, K.; Tuorto, F.; Meusburger, M.; Helm, M.; Lyko, F. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev.* **2010**, *24*, 1590–1595. [[CrossRef](#)] [[PubMed](#)]
93. Chen, Z.; Qi, M.; Shen, B.; Luo, G.; Wu, Y.; Li, J.; Lu, Z.; Zheng, Z.; Dai, Q.; Wang, H. Transfer RNA demethylase ALKBH3 promotes cancer progression via induction of tRNA-derived small RNAs. *Nucleic Acids Res.* **2019**, *47*, 2533–2545. [[CrossRef](#)]
94. Shubina, M.Y.; Musinova, Y.R.; Sheval, E.V. Proliferation, cancer, and aging-novel functions of the nucleolar methyltransferase fibrillarin? *Cell Biol. Int.* **2018**, *42*, 1463–1466. [[CrossRef](#)]
95. Natchiar, S.K.; Myasnikov, A.G.; Kratzat, H.; Hazemann, I.; Klaholz, B.P. Visualization of chemical modifications in the human 80S ribosome structure. *Nature* **2017**, *551*, 472–477. [[CrossRef](#)]
96. Taoka, M.; Nobe, Y.; Yamaki, Y.; Sato, K.; Ishikawa, H.; Izumikawa, K.; Yamauchi, Y.; Hirota, K.; Nakayama, H.; Takahashi, N.; et al. Landscape of the complete RNA chemical modifications in the human 80S ribosome. *Nucleic Acids Res.* **2018**, *46*, 9289–9298. [[CrossRef](#)]
97. Baudin-Baillieu, A.; Namy, O. Saccharomyces cerevisiae, a powerful model for studying rRNA modifications and their effects on translation fidelity. *Int. J. Mol. Sci.* **2021**, *22*, 7419. [[CrossRef](#)]
98. Assi, H.A.; Shi, H.; Liu, B.; Clay, M.; Erharter, K.; Kreutz, C.; Holley, C.; Al-Hashimi, H. 2'-O-methylation alters the RNA secondary structural ensemble. *bioRxiv* **2020**, 2020, 121996. [[CrossRef](#)]
99. Watkins, N.J.; Bohnsack, M.T. The box C/D and H/ACA snoRNPs: Key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 397–414. [[CrossRef](#)] [[PubMed](#)]

100. Erales, J.; Marchand, V.; Panthu, B.; Gillot, S.; Belin, S.; Ghayad, S.E.; Garcia, M.; Laforêts, F.; Marcel, V.; Baudin-Baillieu, A.; et al. Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12934–12939. [[CrossRef](#)] [[PubMed](#)]
101. Sharma, S.; Marchand, V.; Motorin, Y.; Lafontaine, D.L.J. Identification of sites of 2'-O-methylation vulnerability in human ribosomal RNAs by systematic mapping. *Sci. Rep.* **2017**, *7*, 1–15. [[CrossRef](#)]
102. Krogh, N.; Jansson, M.D.; Häfner, S.J.; Tehler, D.; Birkedal, U.; Christensen-Dalsgaard, M.; Lund, A.H.; Nielsen, H. Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res.* **2016**, *44*, 7884–7895. [[CrossRef](#)]
103. Incarnato, D.; Anselmi, F.; Morandi, E.; Neri, F.; Maldotti, M.; Rapelli, S.; Parlato, C.; Basile, G.; Oliviero, S. High-throughput single-base resolution mapping of RNA 2-O-methylated residues. *Nucleic Acids Res.* **2017**, *45*, 1433–1441. [[CrossRef](#)]
104. Marcel, V.; Ghayad, S.E.; Belin, S.; Therizols, G.; Morel, A.P.; Solano-González, E.; Vendrell, J.A.; Hacot, S.; Mertani, H.C.; Albaret, M.A.; et al. P53 Acts as a Safeguard of Translational Control by Regulating Fibrillar and rRNA Methylation in Cancer. *Cancer Cell* **2013**, *24*, 318–330. [[CrossRef](#)] [[PubMed](#)]
105. Baudin-baillieu, A.; Fabret, C.; Liang, X.H.; Piekna-Przybylska, D.; Fournier, M.J.; Rousset, J.P. Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids Res.* **2009**, *37*, 7665–7677. [[CrossRef](#)] [[PubMed](#)]
106. Higa-Nakamine, S.; Suzuki, T.; Uechi, T.; Chakraborty, A.; Nakajima, Y.; Nakamura, M.; Hirano, N.; Suzuki, T.; Kenmochi, N. Loss of ribosomal RNA modification causes developmental defects in zebrafish. *Nucleic Acids Res.* **2012**, *40*, 391–398. [[CrossRef](#)] [[PubMed](#)]
107. Watanabe-Susaki, K.; Takada, H.; Enomoto, K.; Miwata, K.; Ishimine, H.; Intoh, A.; Ohtaka, M.; Nakanishi, M.; Sugino, H.; Asashima, M.; et al. Biosynthesis of Ribosomal RNA in Nucleoli Regulates Pluripotency and Differentiation Ability of Pluripotent Stem Cells. *Stem Cells* **2014**, *32*, 3099–3111. [[CrossRef](#)] [[PubMed](#)]
108. Su, H.; Xu, T.; Ganapathy, S.; Shadfan, M.; Long, M.; Huang, T.H.M.; Thompson, I.; Yuan, Z.M. Elevated snoRNA biogenesis is essential in breast cancer. *Oncogene* **2014**, *33*, 1348–1358. [[CrossRef](#)] [[PubMed](#)]
109. Schattner, P.; Barberan-Soler, S.; Lowe, T.M. A computational screen for mammalian pseudouridylation guide H/ACA RNAs. *RNA* **2006**, *12*, 15–25. [[CrossRef](#)]
110. Jack, K.; Bellodi, C.; Landry, D.M.; Niederer, R.O.; Meskauskas, A.; Musalgaonkar, S.; Kopmar, N.; Krasnykh, O.; Dean, A.M.; Thompson, S.R.; et al. rRNA Pseudouridylation Defects Affect Ribosomal Ligand Binding and Translational Fidelity from Yeast to Human Cells. *Mol. Cell* **2011**, *44*, 660–666. [[CrossRef](#)]
111. Hughes, D.G.; Maden, E.H. The pseudouridine contents of the ribosomal ribonucleic acids of three vertebrate species. Numerical correspondence between pseudouridine residues and 2'-O-methyl groups is not always conserved. *Biochem. J.* **1978**, *171*, 781–786. [[CrossRef](#)]
112. Maden, B.E.H. The Numerous Modified Nucleotides in Eukaryotic Ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **1990**, *39*, 241–303. [[CrossRef](#)]
113. Li, X.; Zhu, P.; Ma, S.; Song, J.; Bai, J.; Sun, F.; Yi, C. Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat. Chem. Biol.* **2015**, *11*, 592–597. [[CrossRef](#)]
114. Bakin, A.V.; Ofengand, J. Mapping of pseudouridine residues in RNA to nucleotide resolution. *Methods Mol. Biol.* **1998**, *77*, 297–309.
115. McMahon, M.; Contreras, A.; Holm, M.; Uechi, T.; Forester, C.M.; Pang, X.; Jackson, C.; Calvert, M.E.; Chen, B.; Quigley, D.A.; et al. A single H/ACA small nucleolar RNA mediates tumor suppression downstream of oncogenic RAS. *Elife* **2019**, *8*, e48847. [[CrossRef](#)] [[PubMed](#)]
116. Kiss, A.M.; Jády, B.E.; Bertrand, E.; Kiss, T. Human Box H/ACA Pseudouridylation Guide RNA Machinery. *Mol. Cell. Biol.* **2004**, *24*, 5797–5807. [[CrossRef](#)] [[PubMed](#)]
117. Liang, X.H.; Liu, Q.; Fournier, M.J. rRNA Modifications in an Intersubunit Bridge of the Ribosome Strongly Affect Both Ribosome Biogenesis and Activity. *Mol. Cell* **2007**, *28*, 965–977. [[CrossRef](#)] [[PubMed](#)]
118. Nishikura, K. Functions and regulation of RNA editing by ADAR deaminases. *Annu. Rev. Biochem.* **2010**, *79*, 321–349. [[CrossRef](#)] [[PubMed](#)]
119. Sakurai, M.; Yano, T.; Kawabata, H.; Ueda, H.; Suzuki, T. Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome. *Nat. Chem. Biol.* **2010**, *6*, 733–740. [[CrossRef](#)] [[PubMed](#)]
120. Sommer, B.; Köhler, M.; Sprengel, R.; Seeburg, P.H. RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* **1991**, *67*, 11–19. [[CrossRef](#)]
121. Kawahara, Y.; Ito, K.; Sun, H.; Aizawa, H.; Kanazawa, I.; Kwak, S. RNA editing and death of motor neurons: There is a glutamate-receptor defect in patients with amyotrophic lateral sclerosis. *Nature* **2004**, *427*, 801. [[CrossRef](#)]
122. Chan, T.H.M.; Lin, C.H.; Qi, L.; Fei, J.; Li, Y.; Yong, K.J.; Liu, M.; Song, Y.; Chow, R.K.K.; Ng, V.H.E.; et al. A disrupted RNA editing balance mediated by ADARs (Adenosine DeAminases that act on RNA) in human hepatocellular carcinoma. *Gut* **2014**, *63*, 832–843. [[CrossRef](#)]

123. Planchard, N.; Bertin, P.; Quadrado, M.; Dargel-Graffin, C.; Hatin, I.; Namy, O.; Mireau, H. The translational landscape of Arabidopsis mitochondria. *Nucleic Acids Res.* **2018**, *46*, 6218. [[CrossRef](#)] [[PubMed](#)]
124. Carlile, T.M.; Rojas-Duran, M.F.; Zinshteyn, B.; Shin, H.; Bartoli, K.M.; Gilbert, W.V. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* **2014**, *515*, 143–146. [[CrossRef](#)] [[PubMed](#)]
125. Carlile, T.M.; Martinez, N.M.; Schaening, C.; Su, A.; Bell, T.A.; Zinshteyn, B.; Gilbert, W.V. mRNA structure determines modification by pseudouridine synthase 1. *Nat. Chem. Biol.* **2019**, *15*, 966–974. [[CrossRef](#)]
126. Safra, M.; Nir, R.; Farouq, D.; Slutzkin, I.V.; Schwartz, S. TRUB1 is the predominant pseudouridine synthase acting on mammalian mRNA via a predictable and conserved code. *Genome Res.* **2017**, *27*, 393–406. [[CrossRef](#)]
127. Hüttenhofer, A.; Brosius, J.; Bachellerie, J.P. RNomics: Identification and function of small, non-messenger RNAs. *Curr. Opin. Chem. Biol.* **2002**, *6*, 835–843. [[CrossRef](#)]
128. Cavallé, J.; Buiting, K.; Kiefmann, M.; Lalande, M.; Brannan, C.I.; Horsthemke, B.; Bachellerie, J.P.; Brosius, J.; Hüttenhofer, A. Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 14311–14316. [[CrossRef](#)]
129. Vandivier, L.E.; Gregory, B.D. Reading the Epitranscriptome: New Techniques and Perspectives. In *Enzymes*; Academic Press: Cambridge, MA, USA, 2017; Volume 41, pp. 269–298.
130. Gilbert, W.V.; Bell, T.A.; Schaening, C. Messenger RNA modifications: Form, distribution, and function. *Science* **2016**, *352*, 1408–1412. [[CrossRef](#)]
131. Eyler, D.E.; Franco, M.K.; Batool, Z.; Wu, M.Z.; Dubuke, M.L.; Dobosz-Bartoszek, M.; Jones, J.D.; Polikanov, Y.S.; Roy, B.; Koutmou, K.S. Pseudouridylation of mRNA coding sequences alters translation. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 23068–23074. [[CrossRef](#)]
132. Karijolic, J.; Yu, Y.T. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* **2011**, *474*, 395–399. [[CrossRef](#)]
133. Fernández, I.S.; Ng, C.L.; Kelley, A.C.; Wu, G.; Yu, Y.T.; Ramakrishnan, V. Unusual base pairing during the decoding of a stop codon by the ribosome. *Nature* **2013**, *500*, 107–110. [[CrossRef](#)] [[PubMed](#)]
134. Svidritskiy, E.; Madireddy, R.; Korostelev, A.A. Structural Basis for Translation Termination on a Pseudouridylated Stop Codon. *J. Mol. Biol.* **2016**, *428*, 2228–2236. [[CrossRef](#)]
135. Jordan Ontiveros, R.; Stoute, J.; Liu, K.F. The chemical diversity of RNA modifications. *Biochem. J.* **2019**, *476*, 1227–1245. [[CrossRef](#)] [[PubMed](#)]
136. Meyer, K.D.; Saletore, Y.; Zumbo, P.; Elemento, O.; Mason, C.E.; Jaffrey, S.R. Comprehensive analysis of mRNA methylation reveals enrichment in 3<sup>0</sup> UTRs and near stop codons. *Cell* **2012**, *149*, 1635–1646. [[CrossRef](#)] [[PubMed](#)]
137. Dominissini, D.; Moshitch-Moshkovitz, S.; Schwartz, S.; Salmon-Divon, M.; Ungar, L.; Osenberg, S.; Cesarkas, K.; Jacob-Hirsch, J.; Amariglio, N.; Kupiec, M.; et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **2012**, *485*, 201–206. [[CrossRef](#)] [[PubMed](#)]
138. Mao, Y.; Dong, L.; Liu, X.M.; Guo, J.; Ma, H.; Shen, B.; Qian, S.B. m6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat. Commun.* **2019**, *10*, 1–11. [[CrossRef](#)]
139. Meyer, K.D.; Jaffrey, S.R. Rethinking m6A readers, writers, and erasers. *Annu. Rev. Cell Dev. Biol.* **2017**, *33*, 319–342. [[CrossRef](#)]
140. Balacco, D.L.; Soller, M. The m6A Writer: Rise of a Machine for Growing Tasks. *Biochemistry* **2019**, *58*, 363–378. [[CrossRef](#)]
141. Choi, J.; Jeong, K.W.; Demirci, H.; Chen, J.; Petrov, A.; Prabhakar, A.; O’Leary, S.E.; Dominissini, D.; Rechavi, G.; Soltis, S.M.; et al. N6-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. *Nat. Struct. Mol. Biol.* **2016**, *23*, 110–115. [[CrossRef](#)]
142. Jeong, K.-W.; Indrisiunaite, G.; Prabhakar, A.; Puglisi, J.D.; Ehrenberg, M. N 6-Methyladenosines in mRNAs reduce the accuracy of codon reading by transfer RNAs and peptide release factors. *Nucleic Acids Res.* **2021**, *49*, 2684–2699. [[CrossRef](#)]
143. You, C.; Dai, X.; Wang, Y. Position-dependent effects of regioisomeric methylated adenine and guanine ribonucleosides on translation. *Nucleic Acids Res.* **2017**, *45*, 9059–9067. [[CrossRef](#)]
144. Hoernes, T.P.; Heimdörfer, D.; Köstner, D.; Faserl, K.; Nußbaumer, F.; Plangger, R.; Kreutz, C.; Lindner, H.; Erlacher, M.D. Eukaryotic translation elongation is modulated by single natural nucleotide derivatives in the coding sequences of mRNAs. *Genes* **2019**, *10*, 84. [[CrossRef](#)] [[PubMed](#)]
145. Wang, X.; Lu, Z.; Gomez, A.; Hon, G.C.; Yue, Y.; Han, D.; Fu, Y.; Parisien, M.; Dai, Q.; Jia, G.; et al. N 6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **2014**, *505*, 117–120. [[CrossRef](#)] [[PubMed](#)]
146. Wang, X.; Zhao, B.S.; Roundtree, I.A.; Lu, Z.; Han, D.; Ma, H.; Weng, X.; Chen, K.; Shi, H.; He, C. N6-methyladenosine modulates messenger RNA translation efficiency. *Cell* **2015**, *161*, 1388–1399. [[CrossRef](#)]
147. Liu, N.; Dai, Q.; Zheng, G.; He, C.; Parisien, M.; Pan, T. N6 -methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **2015**, *518*, 560–564. [[CrossRef](#)] [[PubMed](#)]

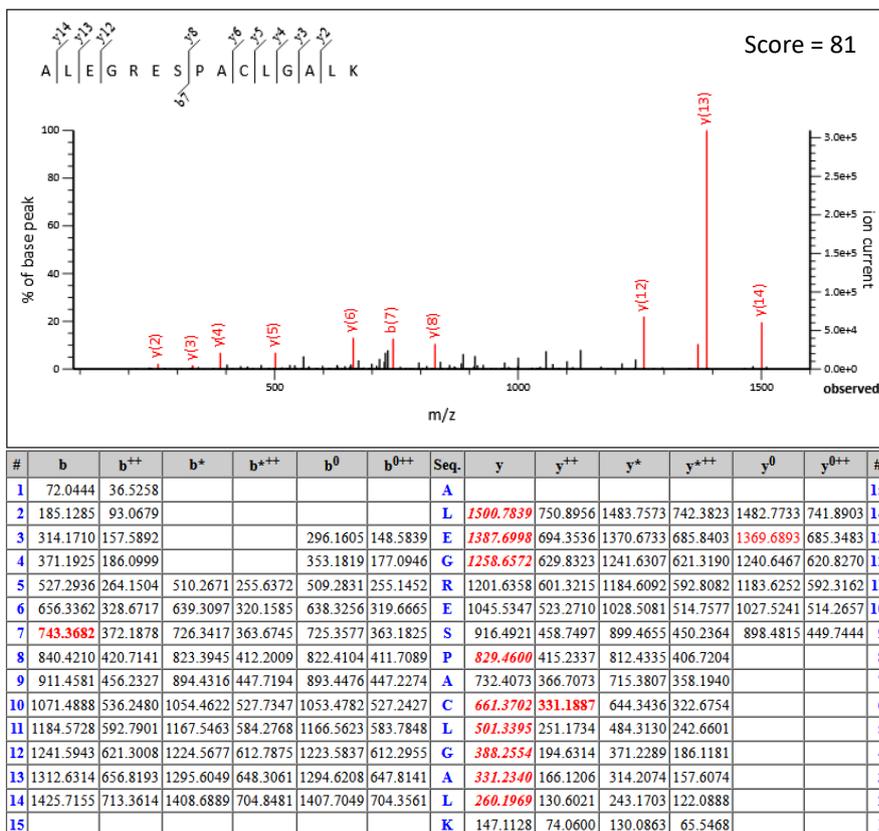
148. Tang, Y.; Chen, K.; Song, B.; Ma, J.; Wu, X.; Xu, Q.; Wei, Z.; Su, J.; Liu, G.; Rong, R.; et al. M6A-Atlas: A comprehensive knowledgebase for unraveling the N6-methyladenosine (m6A) epitranscriptome. *Nucleic Acids Res.* **2021**, *49*, D134–D143. [[CrossRef](#)]
149. Meyer, K.D.; Patil, D.P.; Zhou, J.; Zinoviev, A.; Skabkin, M.A.; Elemento, O.; Pestova, T.V.; Qian, S.B.; Jaffrey, S.R. 5<sup>0</sup> UTR m6A Promotes Cap-Independent Translation. *Cell* **2015**, *163*, 999–1010. [[CrossRef](#)] [[PubMed](#)]
150. He, L.; Li, H.; Wu, A.; Peng, Y.; Shu, G.; Yin, G. Functions of N6-methyladenosine and its role in cancer. *Mol. Cancer* **2019**, *18*, 1–15. [[CrossRef](#)]
151. García-Vílchez, R.; Sevilla, A.; Blanco, S. Post-transcriptional regulation by cytosine-5 methylation of RNA. *Biochim. Biophys. Acta Gene Regul. Mech.* **2019**, *1862*, 240–252. [[CrossRef](#)]
152. Squires, J.E.; Patel, H.R.; Nusch, M.; Sibbritt, T.; Humphreys, D.T.; Parker, B.J.; Suter, C.M.; Preiss, T. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res.* **2012**, *40*, 5023–5033. [[CrossRef](#)]
153. Blanco, S.; Dietmann, S.; Flores, J.V.; Hussain, S.; Kutter, C.; Humphreys, P.; Lukk, M.; Lombard, P.; Treps, L.; Popis, M.; et al. Aberrant methylation of t RNA s links cellular stress to neuro-developmental disorders. *EMBO J.* **2014**, *33*, 2020–2039. [[CrossRef](#)]
154. Hussain, S.; Sajini, A.A.; Blanco, S.; Dietmann, S.; Lombard, P.; Sugimoto, Y.; Paramor, M.; Gleeson, J.G.; Odom, D.T.; Ule, J.; et al. NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Rep.* **2013**, *4*, 255–261. [[CrossRef](#)]
155. Huang, T.; Chen, W.; Liu, J.; Gu, N.; Zhang, R. Genome-wide identification of mRNA 5-methylcytosine in mammals. *Nat. Struct. Mol. Biol.* **2019**, *26*, 380–388. [[CrossRef](#)] [[PubMed](#)]
156. Schumann, U.; Zhang, H.N.; Sibbritt, T.; Pan, A.; Horvath, A.; Gross, S.; Clark, S.J.; Yang, L.; Preiss, T. Multiple links between 5-methylcytosine content of mRNA and translation. *BMC Biol.* **2020**, *18*, 40. [[CrossRef](#)] [[PubMed](#)]
157. Selmi, T.; Hussain, S.; Dietmann, S.; Heiß, M.; Borland, K.; Flad, S.; Carter, J.M.; Dennison, R.; Huang, Y.L.; Kellner, S.; et al. Sequence- and structure-specific cytosine-5 mRNA methylation by NSUN6. *Nucleic Acids Res.* **2021**, *49*, 1006–1022. [[CrossRef](#)] [[PubMed](#)]
158. Khan, M.A.; Rafiq, M.A.; Noor, A.; Hussain, S.; Flores, J.V.; Rupp, V.; Vincent, A.K.; Malli, R.; Ali, G.; Khan, F.S.; et al. Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. *Am. J. Hum. Genet.* **2012**, *90*, 856–863. [[CrossRef](#)]
159. Abbasi-Moheb, L.; Mertel, S.; Gonsior, M.; Nouri-Vahid, L.; Kahrizi, K.; Cirak, S.; Wieczorek, D.; Motazacker, M.M.; Esmaeeli-Nieh, S.; Cremer, K.; et al. Mutations in NSUN2 cause autosomal- Recessive intellectual disability. *Am. J. Hum. Genet.* **2012**, *90*, 847–855. [[CrossRef](#)]
160. Martinez, F.J.; Lee, J.H.; Lee, J.E.; Blanco, S.; Nickerson, E.; Gabriele, S.; Frye, M.; Al-Gazali, L.; Gleeson, J.G. Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. *J. Med. Genet.* **2012**, *49*, 380–385. [[CrossRef](#)]
161. Blanco, S.; Bandiera, R.; Popis, M.; Hussain, S.; Lombard, P.; Aleksic, J.; Sajini, A.; Tanna, H.; Cortés-Garrido, R.; Gkatza, N.; et al. Stem cell function and stress response are controlled by protein synthesis. *Nature* **2016**, *534*, 335–340. [[CrossRef](#)] [[PubMed](#)]
162. Kaur, K.; Zangi, L. Modified mRNA as a Therapeutic Tool for the Heart. *Cardiovasc. Drugs Ther.* **2020**, *34*, 871–880. [[CrossRef](#)]
163. Mays, L.E.; Ammon-Treiber, S.; Mothes, B.; Alkhaled, M.; Rottenberger, J.; Müller-Hermelink, E.S.; Grimm, M.; Mezger, M.; Beer-Hammer, S.; Von Stebut, E.; et al. Modified Foxp3 mRNA protects against asthma through an IL-10-dependent mechanism. *J. Clin. Investig.* **2013**, *123*, 1216–1228. [[CrossRef](#)]
164. Haque, A.K.M.A.; Dewerth, A.; Antony, J.S.; Riethmüller, J.; Schweizer, G.R.; Weinmann, P.; Latifi, N.; Yasar, H.; Pedemonte, N.; Sondo, E.; et al. Chemically modified hCFTR mRNAs recuperate lung function in a mouse model of cystic fibrosis. *Sci. Rep.* **2018**, *8*, 1–14. [[CrossRef](#)]
165. Sahu, I.; Haque, A.K.M.A.; Weidensee, B.; Weinmann, P.; Kormann, M.S.D. Recent Developments in mRNA-Based Protein Supplementation Therapy to Target Lung Diseases. *Mol. Ther.* **2019**, *27*, 803–823. [[CrossRef](#)]
166. Nance, K.D.; Meier, J.L. Modifications in an Emergency: The Role of N1-Methylpseudouridine in COVID-19 Vaccines. *ACS Cent. Sci.* **2021**, *7*, 748–756. [[CrossRef](#)] [[PubMed](#)]

167. Svitkin, Y.V.; Cheng, Y.M.; Chakraborty, T.; Presnyak, V.; John, M.; Sonenberg, N. N1-methyl-pseudouridine in mRNA enhances translation through eIF2 $\alpha$ -dependent and independent mechanisms by increasing ribosome density. *Nucleic Acids Res.* **2017**, *45*, 6023–6036. [[CrossRef](#)] [[PubMed](#)]
168. Wek, R.C. Role of eIF2 $\alpha$  kinases in translational control and adaptation to cellular stress. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a032870. [[CrossRef](#)]
169. Rapino, F.; Delaunay, S.; Rambow, F.; Zhou, Z.; Tharun, L.; De Tullio, P.; Sin, O.; Shostak, K.; Schmitz, S.; Piepers, J.; et al. Codon-specific translation reprogramming promotes resistance to targeted therapy. *Nature* **2018**, *558*, 605–609. [[CrossRef](#)]
170. Dong, S.; Wu, Y.; Liu, Y.; Weng, H.; Huang, H. N<sup>6</sup>-methyladenosine Steers RNA Metabolism and Regulation in Cancer. *Cancer Commun.* **2021**, cac2.12161. [[CrossRef](#)]
171. Huang, J.; Chen, Z.; Chen, X.; Chen, J.; Cheng, Z.; Wang, Z. The role of RNA N6-methyladenosine methyltransferase in cancers. *Mol. Ther. Nucleic Acids* **2021**, *23*, 887–896. [[CrossRef](#)]
172. Chen, X.; Li, A.; Sun, B.F.; Yang, Y.; Han, Y.N.; Yuan, X.; Chen, R.X.; Wei, W.S.; Liu, Y.; Gao, C.C.; et al. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nat. Cell Biol.* **2019**, *21*, 978–990. [[CrossRef](#)]
173. Nombela, P.; Miguel-López, B.; Blanco, S. The role of m6A, m5C and  $\Psi$  RNA modifications in cancer: Novel therapeutic opportunities. *Mol. Cancer* **2021**, *20*, 1–30. [[CrossRef](#)]
174. Ruggero, D.; Grisendi, S.; Piazza, F.; Rego, E.; Mari, F.; Rao, P.H.; Cordon-Cardo, C.; Pandolfi, P.P. Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science* **2003**, *299*, 259–262. [[CrossRef](#)] [[PubMed](#)]
175. Bidou, L.; Bugaud, O.; Belakhov, V.; Baasov, T.; Namy, O. Characterization of new-generation aminoglycoside promoting premature termination codon readthrough in cancer cells. *RNA Biol.* **2017**, *14*, 378–388. [[CrossRef](#)]
176. Trzaska, C.; Amand, S.; Bailly, C.; Leroy, C.; Marchand, V.; Duvernois-Berthet, E.; Saliou, J.M.; Benhabiles, H.; Werkmeister, E.; Chassat, T.; et al. 2,6-Diaminopurine as a highly potent corrector of UGA nonsense mutations. *Nat. Commun.* **2020**, *11*, 1–12. [[CrossRef](#)] [[PubMed](#)]
177. McIntyre, A.B.R.; Gokhale, N.S.; Cerchietti, L.; Jaffrey, S.R.; Horner, S.M.; Mason, C.E. Limits in the detection of m6A changes using MeRIP/m6A-seq. *Sci. Rep.* **2020**, *10*. [[CrossRef](#)] [[PubMed](#)]
178. Thomas, N.; Poodari, V.; Jain, M.; Olsen, H.; Akeson, M.; Abu-Shumays, R. Direct Nanopore Sequencing of Individual Full Length tRNA Strands. *bioRxiv* **2021**. [[CrossRef](#)]

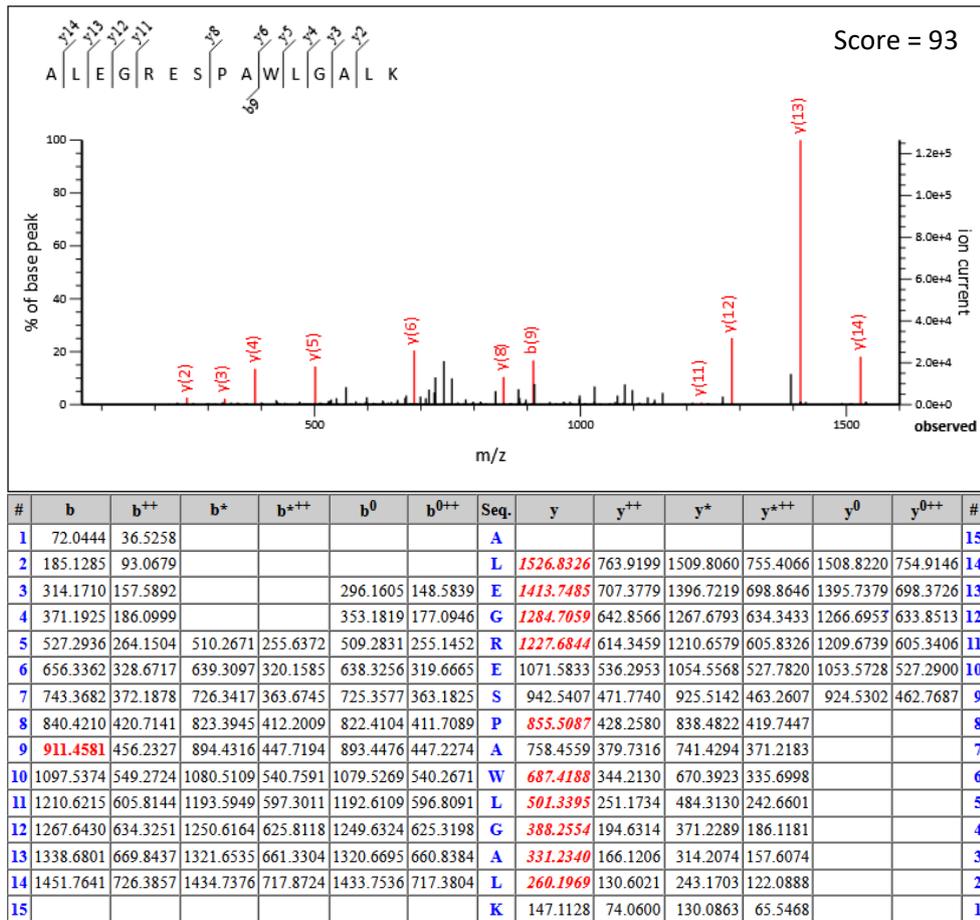
## Données supplémentaires de l'article de résultats

**Figure S1:** Identification and relative quantification by mass spectrometry of amino acids incorporated at stop codons. Amino acids incorporated at stop codons UAA, UAG and UGA were identified by nanoLC-MSMS analyses after LysC proteolysis. Fragmentation spectra of readthrough peptides 51-65 are shown (a)-(f). Labeled peaks correspond to N- and C-terminal fragments (b and y ions) of the triple charged precursor ion. Amino acids incorporated at stop codons (C, W, R, S at UGA and Q, Y at UAA and UAG) are surrounded on the sequence displayed on the top by y5-y6 or b9-b10 fragment ion pairs. Fragmentations tables indicate mass of different fragments. Mascot ion scores of the 6 spectra are displayed on top right corner. Two exemples of extracted ion chromatograms (XIC) of readthrough peptides from OPRL1-TGA *wt* and OPRL1-TAG *wt* with their synthetic heavy isotope labeled counterparts (AQUA peptides) are shown (g-h). The relative incorporations of amino acids are deduced from relative intensities of endogenous peptides corrected by AQUA peptides intensities spiked in mixtures. The strictly identical retention times between endogenous and synthetic peptides confirm amino acids inserted at stop codons positions. Endogenous and AQUA peptides with lysine at stop codon position (<sup>51</sup>ALEGRESPAK<sup>60</sup>) are not detected (h).

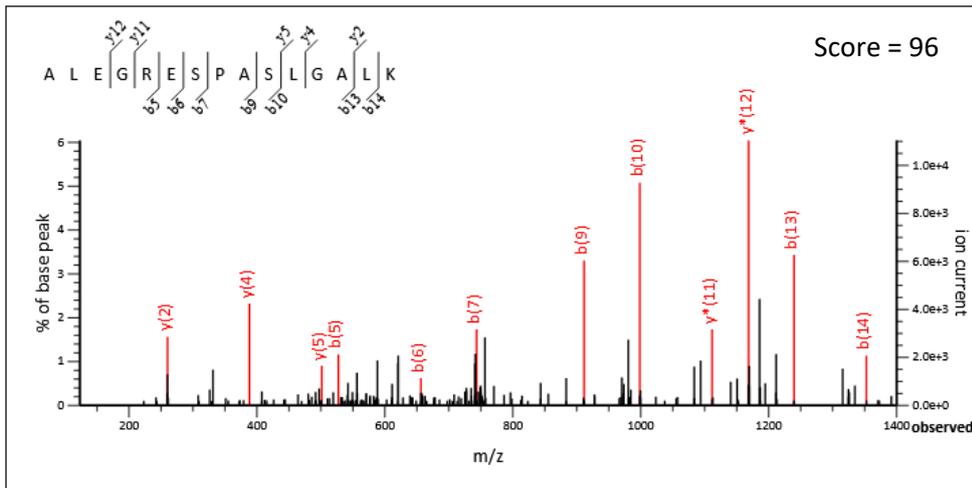
a)



b)

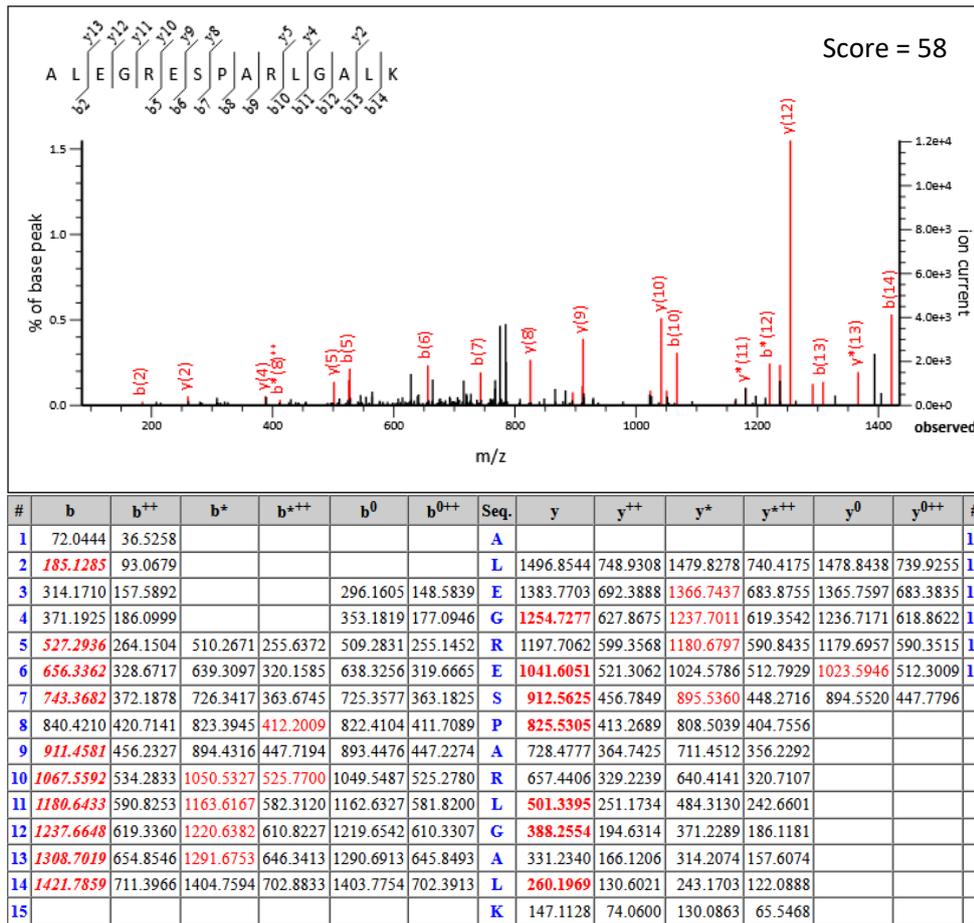


c)

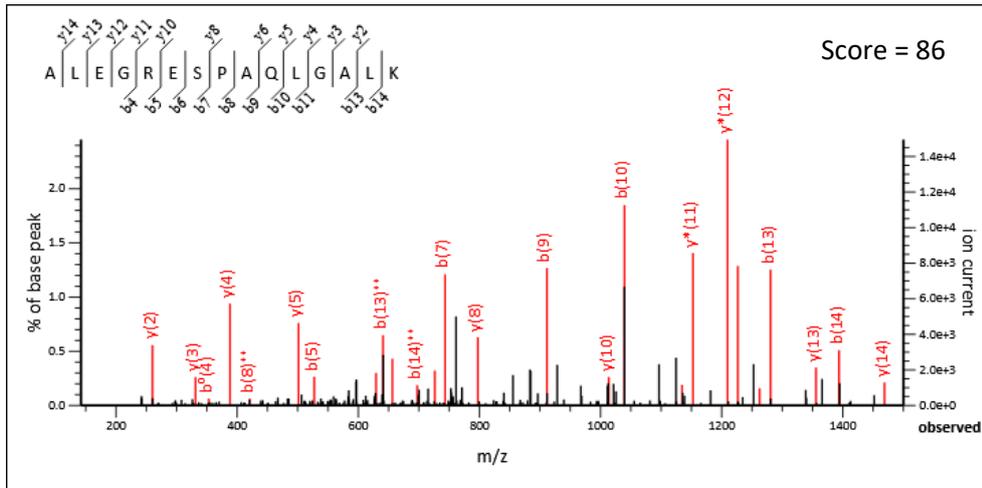


#	b	b <sup>++</sup>	b <sup>*</sup>	b <sup>+++</sup>	b <sup>0</sup>	b <sup>0++</sup>	Seq.	y	y <sup>++</sup>	y <sup>*</sup>	y <sup>+++</sup>	y <sup>0</sup>	y <sup>0++</sup>	#
1	72.0444	36.5258					A							15
2	185.1285	93.0679					L	1427.7853	714.3963	1410.7587	705.8830	1409.7747	705.3910	14
3	314.1710	157.5892			296.1605	148.5839	E	1314.7012	657.8542	1297.6747	649.3410	1296.6906	648.8490	13
4	371.1925	186.0999			353.1819	177.0946	G	1185.6586	593.3329	1168.6321	584.8197	1167.6480	584.3277	12
5	<b>527.2936</b>	264.1504	510.2671	255.6372	509.2831	255.1452	R	1128.6371	564.8222	1111.6106	556.3089	1110.6266	555.8169	11
6	<b>656.3362</b>	328.6717	639.3097	320.1585	638.3256	319.6665	E	972.5360	486.7717	955.5095	478.2584	954.5255	477.7664	10
7	<b>743.3682</b>	372.1878	726.3417	363.6745	725.3577	363.1825	S	843.4934	422.2504	826.4669	413.7371	825.4829	413.2451	9
8	840.4210	420.7141	823.3945	412.2009	822.4104	411.7089	P	756.4614	378.7343	739.4349	370.2211	738.4509	369.7291	8
9	<b>911.4581</b>	456.2327	894.4316	447.7194	893.4476	447.2274	A	659.4087	330.2080	642.3821	321.6947	641.3981	321.2027	7
10	<b>998.4901</b>	499.7487	981.4636	491.2354	980.4796	490.7434	S	588.3715	294.6894	571.3450	286.1761	570.3610	285.6841	6
11	<b>1111.5742</b>	556.2907	1094.5477	547.7775	1093.5636	547.2855	L	<b>501.3395</b>	251.1734	484.3130	242.6601			5
12	<b>1168.5957</b>	584.8015	1151.5691	576.2882	1150.5851	575.7962	G	<b>388.2554</b>	194.6314	371.2289	186.1181			4
13	<b>1239.6328</b>	620.3200	1222.6062	611.8068	1221.6222	611.3148	A	331.2340	166.1206	314.2074	157.6074			3
14	<b>1352.7169</b>	676.8621	1335.6903	668.3488	1334.7063	667.8568	L	<b>260.1969</b>	130.6021	243.1703	122.0888			2
15							K	147.1128	74.0600	130.0863	65.5468			1

d)

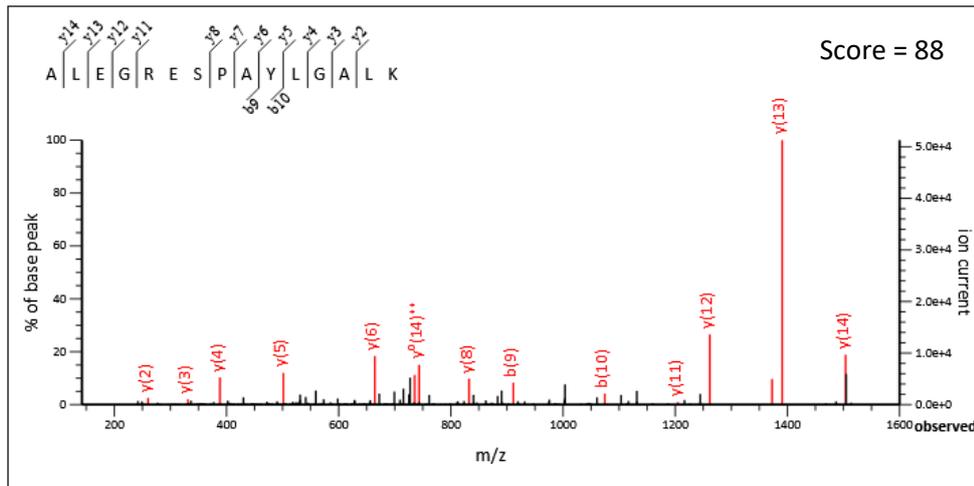


e)



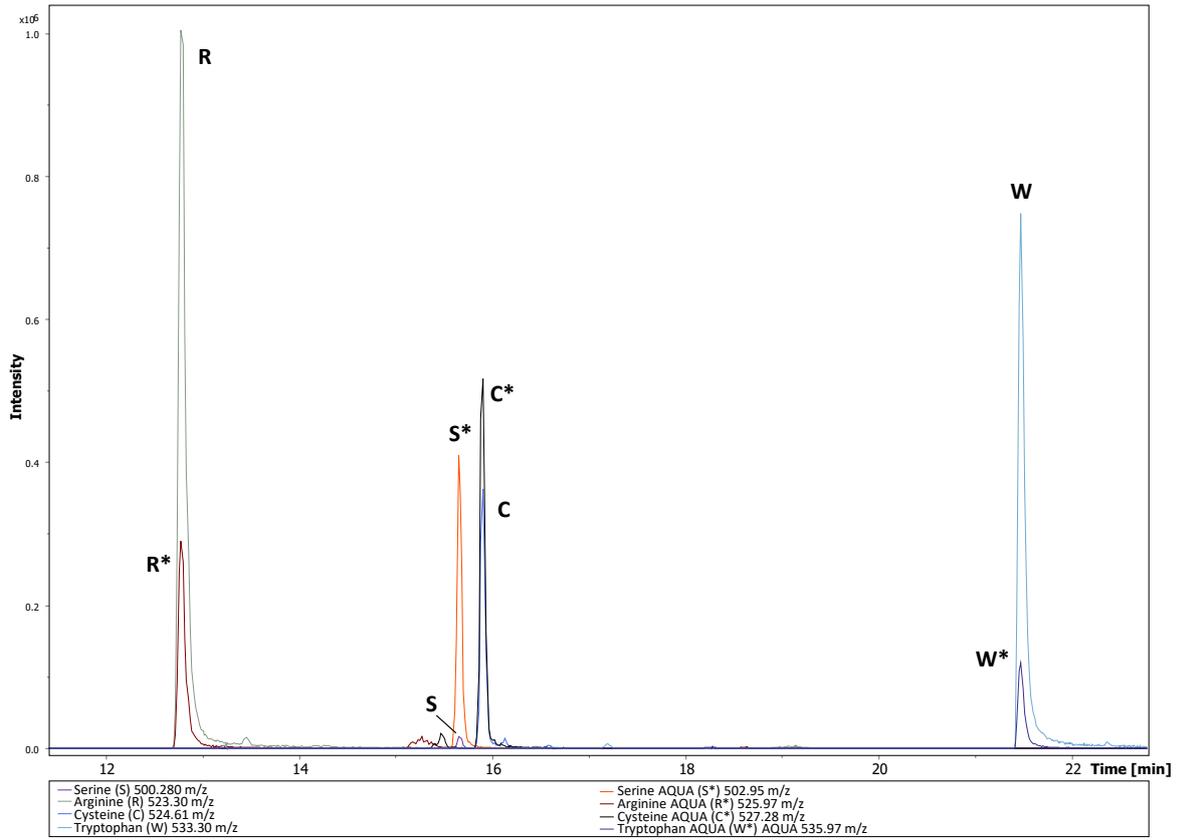
#	b	b <sup>++</sup>	b <sup>*</sup>	b <sup>***</sup>	b <sup>0</sup>	b <sup>0++</sup>	Seq.	y	y <sup>++</sup>	y <sup>*</sup>	y <sup>***</sup>	y <sup>0</sup>	y <sup>0++</sup>	#
1	72.0444	36.5258					A							15
2	185.1285	93.0679					L	1468.8118	734.9095	1451.7853	726.3963	1450.8013	725.9043	14
3	314.1710	157.5892			296.1605	148.5839	E	1355.7278	678.3675	1338.7012	669.8542	1337.7172	669.3622	13
4	371.1925	186.0999			353.1819	177.0946	G	1226.6852	613.8462	1209.6586	605.3329	1208.6746	604.8409	12
5	527.2936	264.1504	510.2671	255.6372	509.2831	255.1452	R	1169.6637	585.3355	1152.6371	576.8222	1151.6531	576.3302	11
6	656.3362	328.6717	639.3097	320.1585	638.3256	319.6665	E	1013.5626	507.2849	996.5360	498.7717	995.5520	498.2796	10
7	743.3682	372.1878	726.3417	363.6745	725.3577	363.1825	S	884.5200	442.7636	867.4934	434.2504	866.5094	433.7584	9
8	840.4210	420.7141	823.3945	412.2009	822.4104	411.7089	P	797.4880	399.2476	780.4614	390.7343			8
9	911.4581	456.2327	894.4316	447.7194	893.4476	447.2274	A	700.4352	350.7212	683.4087	342.2080			7
10	1039.5167	520.2620	1022.4901	511.7487	1021.5061	511.2567	Q	629.3981	315.2027	612.3715	306.6894			6
11	1152.6008	576.8040	1135.5742	568.2907	1134.5902	567.7987	L	501.3395	251.1734	484.3130	242.6601			5
12	1209.6222	605.3148	1192.5957	596.8015	1191.6117	596.3095	G	388.2554	194.6314	371.2289	186.1181			4
13	1280.6593	640.8333	1263.6328	632.3200	1262.6488	631.8280	A	331.2340	166.1206	314.2074	157.6074			3
14	1393.7434	697.3753	1376.7169	688.8621	1375.7328	688.3701	L	260.1969	130.6021	243.1703	122.0888			2
15							K	147.1128	74.0600	130.0863	65.5468			1

f)

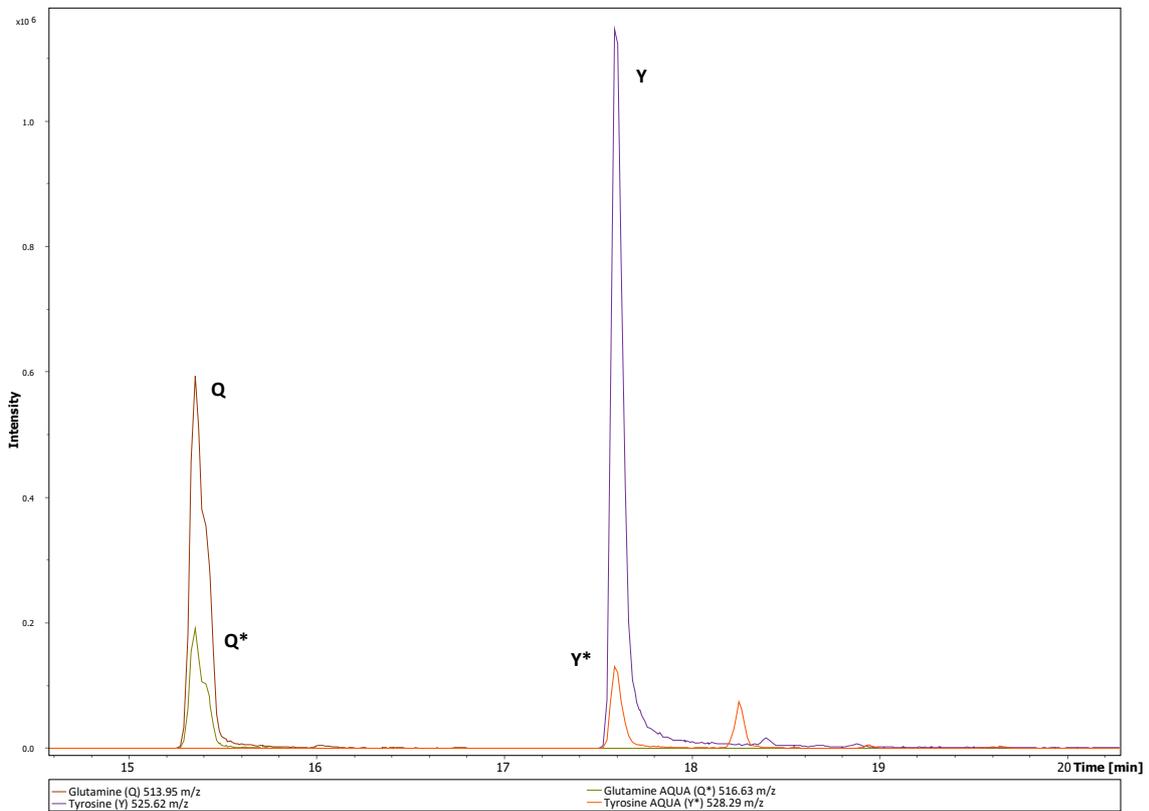


#	b	b <sup>++</sup>	b <sup>*</sup>	b <sup>+++</sup>	b <sup>0</sup>	b <sup>0++</sup>	Seq.	y	y <sup>++</sup>	y <sup>*</sup>	y <sup>+++</sup>	y <sup>0</sup>	y <sup>0++</sup>	#
1	72.0444	36.5258					A							15
2	185.1285	93.0679					L	1503.8166	752.4119	1486.7900	743.8986	1485.8060	743.4066	14
3	314.1710	157.5892			296.1605	148.5839	E	1390.7325	695.8699	1373.7060	687.3566	1372.7219	686.8646	13
4	371.1925	186.0999			353.1819	177.0946	G	1261.6899	631.3486	1244.6634	622.8353	1243.6793	622.3433	12
5	527.2936	264.1504	510.2671	255.6372	509.2831	255.1452	R	1204.6684	602.8379	1187.6419	594.3246	1186.6579	593.8326	11
6	656.3362	328.6717	639.3097	320.1585	638.3256	319.6665	E	1048.5673	524.7873	1031.5408	516.2740	1030.5568	515.7820	10
7	743.3682	372.1878	726.3417	363.6745	725.3577	363.1825	S	919.5247	460.2660	902.4982	451.7527	901.5142	451.2607	9
8	840.4210	420.7141	823.3945	412.2009	822.4104	411.7089	P	832.4927	416.7500	815.4662	408.2367			8
9	911.4581	456.2327	894.4316	447.7194	893.4476	447.2274	A	735.4400	368.2236	718.4134	359.7103			7
10	1074.5214	537.7644	1057.4949	529.2511	1056.5109	528.7591	Y	664.4028	332.7051	647.3763	324.1918			6
11	1187.6055	594.3064	1170.5790	585.7931	1169.5949	585.3011	L	501.3395	251.1734	484.3130	242.6601			5
12	1244.6270	622.8171	1227.6004	614.3039	1226.6164	613.8118	G	388.2554	194.6314	371.2289	186.1181			4
13	1315.6641	658.3357	1298.6375	649.8224	1297.6535	649.3304	A	331.2340	166.1206	314.2074	157.6074			3
14	1428.7482	714.8777	1411.7216	706.3644	1410.7376	705.8724	L	260.1969	130.6021	243.1703	122.0888			2
15							K	147.1128	74.0600	130.0863	65.5468			1

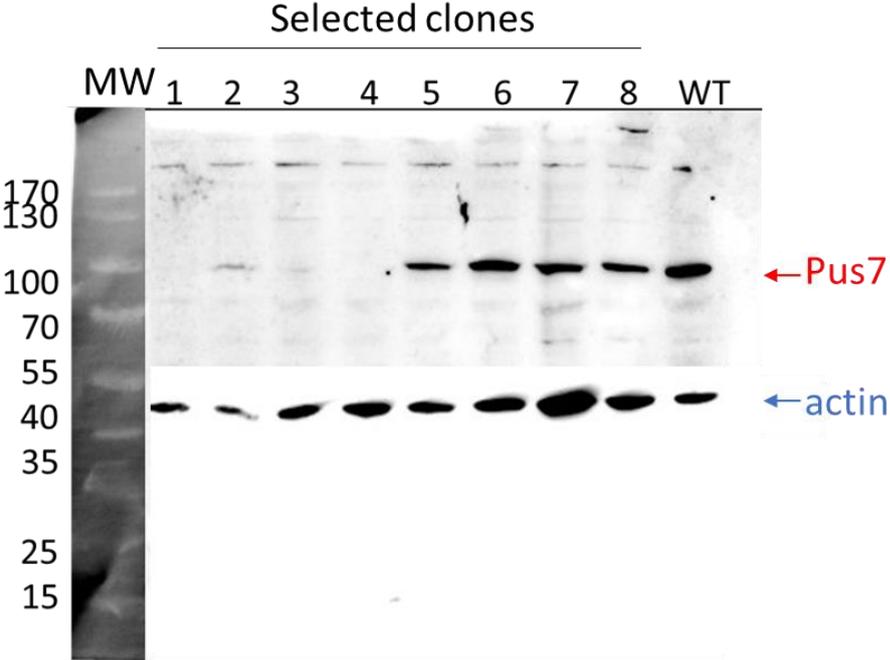
g)



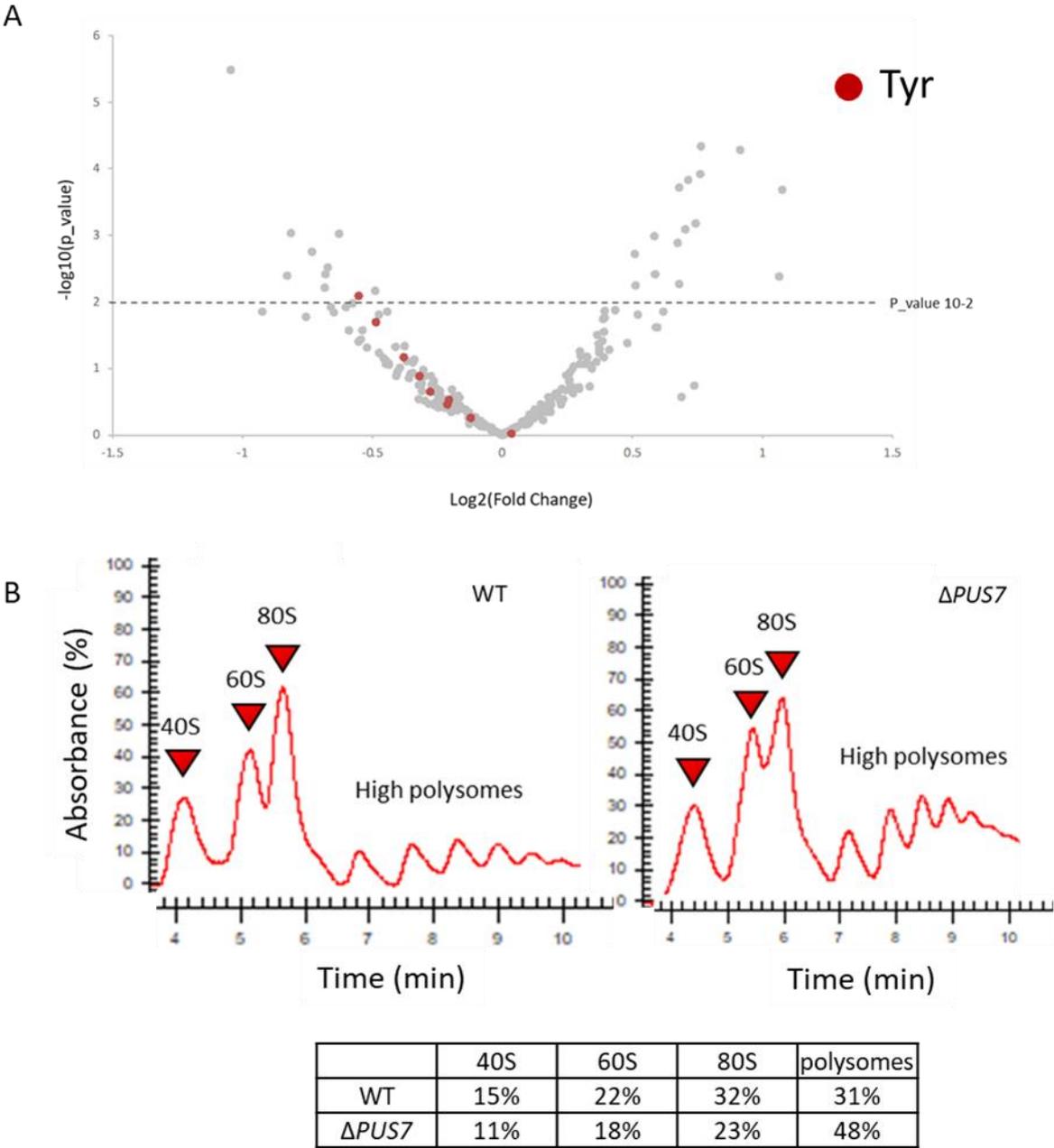
h)



**Figure S2:** Pus7 protein expression in Expi293F cells WT or selected for Pus7 deletion. The western blot represents the revealing of the proteins Pus7 and actin. The clones were selected by flow cytometry for GFP fluorescence, then by resistance to Puromycine antibiotic. A genomic sequencing of clone 4 showed a monoclonal cell line, with alleles with an insertion of CRISPR Cas9 cassette or a deletion of 8 bases in the beginning of *PUS7* gene.



**Figure S3** : Impact of Pus7 deletion on global translation. (A) Volcano plot of variations in tRNAs isodecoders pools between Expi293F cells  $\Delta$ PUS7 and WT. The Volcano plot was performed with the CPM (counts per million reads mapped) values calculating with multimapped reads. n=3 samples. (B) Polysomes profiles obtained for WT and  $\Delta$ PUS7 cells cultured with 200nM queuine. The table resumes the quantification of the various fractions of the polysome profiles obtained. n=2 samples.



**Titre :** Rôle des modifications des ARN dans la fidélité de la traduction chez les eucaryotes

**Mots clés :** Traduction, ARNt, Ribosomes spécialisés, Molécule Unique, Fidélité, modifications des ARN.

**Résumé :** La traduction possède un taux d'erreurs de  $10^{-3}$  à  $10^{-4}$  conservé chez l'ensemble du vivant. Chez l'Homme, plus de 100 modifications de l'ARN ont été décrites, certaines connues pour moduler la fidélité de la traduction. Lors de la terminaison de la traduction, la reconnaissance du codon stop par les facteurs de terminaison est en compétition avec la translecture, processus au cours duquel un ARNt proche cognat est incorporé au niveau du codon stop, le ribosome continuant alors la traduction jusqu'au prochain codon stop. La compréhension des mécanismes de translecture est d'autant importante qu'il s'agit d'un véritable enjeu thérapeutique pour les patients atteints de maladies génétiques à codon stop prématuré. Le premier axe de ma thèse consiste en l'étude de l'impact des modifications de la boucle anticodon des ARNt sur l'efficacité de translecture chez l'Homme. Au cours de ce projet j'ai réalisé la première analyse systématique des acides aminés incorporés au niveau des trois codons stop dans des cellules humaines en culture. Les résultats de spectrométrie de masse ont permis d'identifier les acides aminés

Gln et Tyr au niveau des codons UAA et UAG, et les acides aminés Trp, Arg, Cys et Ser au niveau du codon UGA. J'ai ensuite étudié le rôle des modifications GalQ<sub>34</sub> et  $\Psi_{35}$  de l'anticodon de l'ARNt<sub>Tyr</sub>. J'ai pu mettre en évidence que ces deux modifications jouent deux rôles opposés dans la fidélité de la traduction. Alors que la présence de GalQ<sub>34</sub> diminuait l'efficacité d'incorporation de l'ARNt<sub>Tyr</sub> au niveau des codons UAA et UAG, la présence de  $\Psi_{35}$  était importante pour permettre l'incorporation de l'ARNt. Ces dernières années il est apparu que les modifications 2'-O-méthylation de l'ARNr jouaient un rôle important dans la régulation de la traduction en particulier dans les cellules cancéreuses. Le deuxième axe de ma thèse portait sur l'impact de la modification 2'-O-méthylation de l'ARNr dans la cinétique du ribosome en molécule unique par microscopie TIRF. Durant les étapes de mise au point du protocole, j'ai pu mettre en évidence que la cinétique d'initiation, à partir de l'IRES de l'EMCV, était bien plus rapide qu'à partir d'autres IRES (HCV ou CrPV).

**Title :** Role of RNAs' modifications in eukaryotic Translation fidelity

**Keywords :** Translation, tRNAs, specialized ribosomes, Single Molecule, fidelity, RNAs modifications.

**Abstract :** Translation has an error rate of  $10^{-3}$  to  $10^{-4}$  that is conserved in all living organisms. In humans, more than 100 RNA modifications have been described, some of which are known to modulate the fidelity of translation. During translation termination, recognition of the stop codon by termination factors competes with readthrough, a process in which a cognate tRNA is incorporated at the stop codon and the ribosome continues translation to the next stop codon. The understanding of readthrough mechanisms is important because it is a real therapeutic issue for patients with genetic diseases with premature stop codon. The first axis of my thesis consists in the study of the impact of modifications of the tRNA anticodon loop on the efficiency of readthrough in humans. During this project I performed the first systematic analysis of amino acids incorporated at the three stop codons in human cells in culture. The mass spectrometry results identified the amino

acids Gln and Tyr at the UAA and UAG codons, and the amino acids Trp, Arg, Cys and Ser at the UGA codon. I then investigated the role of the GalQ<sub>34</sub> and Ψ<sub>35</sub> modifications of the tRNA<sub>Tyr</sub> anticodon. I was able to highlight that these two modifications play two opposing roles in translation fidelity. While the presence of GalQ<sub>34</sub> decreased the efficiency of tRNA<sub>Tyr</sub> incorporation at the UAA and UAG codons, the presence of Ψ<sub>35</sub> was important to allow tRNA incorporation. In recent years, it appeared that 2'-O-methylation modifications of rRNA play an important role in the regulation of translation, especially in cancer cells. The second focus of my thesis was on the impact of 2'-O-methylation modification of rRNA in single molecule ribosome kinetics by TIRF microscopy. During the protocol development stages, I was able to demonstrate that the initiation kinetics from the EMCV IRES was much faster than from other IRES (HCV or CrPV).