

Introduction

Préambule	11
I – L’acteur principal de la traduction : le ribosome	11
1 – Les acteurs de la traduction	11
2 – La structure et la composition du ribosome	21
3 – La biogenèse du ribosome.....	27
A – La synthèse et la maturation des composants du ribosome.....	27
a – La synthèse et la maturation des protéines ribosomiques	27
b – La synthèse et la maturation des ARN ribosomiques	29
c – L’export du noyau et l’assemblage des composants du ribosome	31
d – Mutations dans les protéines ribosomiques et facteurs d’assemblage du ribosome et pathologies associées	33
B – Les pseudouridylations et 2’-O-méthylation des ARNr et les complexes snoRNPs qui les synthétisent	35
a – Les snoRNPs à boîte H/ACA, responsables des pseudouridylations	35
b – Les snoRNPs à boîte C/D, responsables des 2’-O-méthylation	37
c – Expression et synthèse des snoARNs	41
d – Effets d’une perte totale ou partielle des pseudouridylations et 2’-O-méthylation	41
e – Défauts des machineries snoRNPs et pathologies	51
II – La traduction chez les Eucaryotes.....	55
1 – L’initiation de la traduction	55
A – L’initiation cap-dépendante	55
B – L’initiation IRES-dépendante	59
C – Autres modes d’initiation	61
2 – L’elongation de la traduction	65
3 – La terminaison de la traduction	69
4 – Le recyclage de la machinerie de traduction et la réinitiation	71
III – La régulation et la fidélité de la traduction	79
1 – L’initiation est l’étape la plus régulée	79
2 – Régulation et fidélité de l’elongation	87
3 – Régulation et fidélité de la terminaison	93
4 – Régulation de la traduction et croissance cellulaire	99
5 – Voies de surveillance NGD, NSD et NMD	103
6 – Hétérogénéité des ribosomes et traduction	109
A – Hétérogénéité dans la composition des protéines	111
B – Hétérogénéité des séquences des ARN ribosomiques	117

C – Hétérogénéité des modifications chimiques des ARNr	119
Objectifs de thèse	123
Résultats et Discussion	127
Induction de la baisse du taux des 2'-O-méthylations des ARNr.....	127
I – Résultats	127
1 – Extinction de la fibrillarine par un shRNA	127
2 – Baisse du taux de fibrillarine et diminution globale de la méthylation des ARNr	129
3 – Effets de la diminution du taux de fibrillarine et de la baisse des méthylations des ARNr sur la traduction	133
II – Discussion	135
Analyses globale et ciblée des effets de la baisse de la méthylation des ARNr sur la traduction	143
I – Résultats	143
1 – Analyse globale à l'échelle du génome.....	143
2 – Analyse bioinformatique des données de séquençage.....	145
3 – Recherche d'éléments fonctionnels et/ou moléculaires communs aux gènes affectés .	149
4 – Impact sur la fidélité de la traduction	155
II – Discussion	157
Résumé de l'article « Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes »	169
Analyse des effets de la baisse de la méthylation des ARNr sur l'initiation de la traduction	189
I – Résultats	189
1 – Analyse globale à l'échelle du génome de l'initiation de la traduction	189
2 – L'initiation de la traduction est affectée par la baisse de la méthylation des ARNr	191
II – Discussion	195
Délétions ciblées de 2'-O-méthylations localisées dans le tunnel de sortie du peptide naissant chez <i>S. cerevisiae</i>	203
I – Résultats	203
1 – Localisation des 2'-O-méthylations affectées par la baisse du taux de fibrillarine dans les cellules humaines	203
2 – Construction de souches de levure déletées pour des snoARNs	203
3 – Impact sur la croissance cellulaire	205
Figure 35 : profil de polysomes de la souche FY1679	206
4 – Etude de l'impact sur la fidélité de la traduction	207
Figure 36 : profil de polysomes de la souche FYΔ(39B-58-24) pASC1	208
II – Discussion	209
Figure 38 : profil de polysomes de la souche FYΔ(39B-58-18) pEFB1	210
La protéine Asc1p est impliquée dans la translecture du codon stop	217

I - Résultats	217
1 – La protéine Asc1p est une protéine plateforme conservée chez les eucaryotes et impliquée dans de nombreux processus traductionnels et cellulaires.....	217
2 – L'absence de la protéine Asc1p entraîne une diminution de la translecture des codons stop	221
Figure 46 : Profil de polysomes de la souche FY1679	222
3 – La délétion du gène ASC1 augmente la sensibilité à des antibiotiques	223
Figure 48 : Profil de polysomes de la souche FYΔ(ASC1-U24) pASC1	224
II – Discussion	225
Figure 49 : profil de polysomes de la souche FYΔ(ASC1-U24) pACT-U24	226
Conclusions et Perspectives	235
Matériels et Méthodes	247
I - Les lignées cellulaires et les souches de levure	247
II – Les plasmides	248
III – Les milieux	249
IV – Les protocoles	249
Références bibliographiques	257

Introduction

Introduction

Préambule

L'expression des gènes aboutit à la synthèse de protéines remplissant des fonctions diverses dans les cellules et se compose de plusieurs étapes. Tout d'abord, les gènes, sous forme ADN double brin, sont transcrits en ARN pré-messagers simple brin qui sont maturés et exportés du noyau vers le cytoplasme. Ils sont alors pris en charge par les ribosomes qui synthétisent les protéines. Les acides aminés qui composent les protéines sont assemblés au sein des ribosomes par les ARN de transfert, qui suivent le code génétique. Ce code assigne un acide aminé pour chaque triplet de nucléotide – ou codon – de l'ARNm (Figure 1).

Les ribosomes sont composés de protéines et d'ARN. Leur biogenèse est divisée en plusieurs étapes de synthèse, de maturation, et d'assemblage des différents composants et nécessite d'être fortement régulée.

Au sein des ribosomes, ce sont les ARNs qui portent l'activité catalytique : les ribosomes sont donc considérés comme des ribozymes. Ces ARNs ribosomiques présentent des modifications chimiques qui jouent un rôle structural et fonctionnel. Durant ma thèse, je me suis intéressée au rôle d'une de ces modifications, la méthylation des riboses, sur la fidélité de la traduction.

Au cours de cette introduction, je vais vous présenter tout d'abord les différents acteurs de la traduction, et plus précisément le ribosome, qui est l'acteur principal. Je détaillerai sa biogenèse et les facteurs qui y participent. Puis je présenterai la traduction et chaque étape qui la compose. J'exposerai ensuite les mécanismes de régulation et de fidélité de la traduction mis en jeu à certaines étapes ainsi que les voies de surveillance de l'ARN messager. Enfin, j'aborderai les notions d'hétérogénéité des ribosomes et de ribosome spécialisé, qui ont émergées ces dernières années, et sur lesquelles une partie importante de mes objectifs de thèse reposent. Je terminerai l'introduction par exposer ces objectifs.

I – L'acteur principal de la traduction : le ribosome

1 – Les acteurs de la traduction

La traduction fait intervenir de très nombreux facteurs protéiques et nucléiques, dont les principaux sont les ARNs messager (ARNm), les ARNs de transfert (ARNt) et le complexe ribonucléoprotéique ribosome.

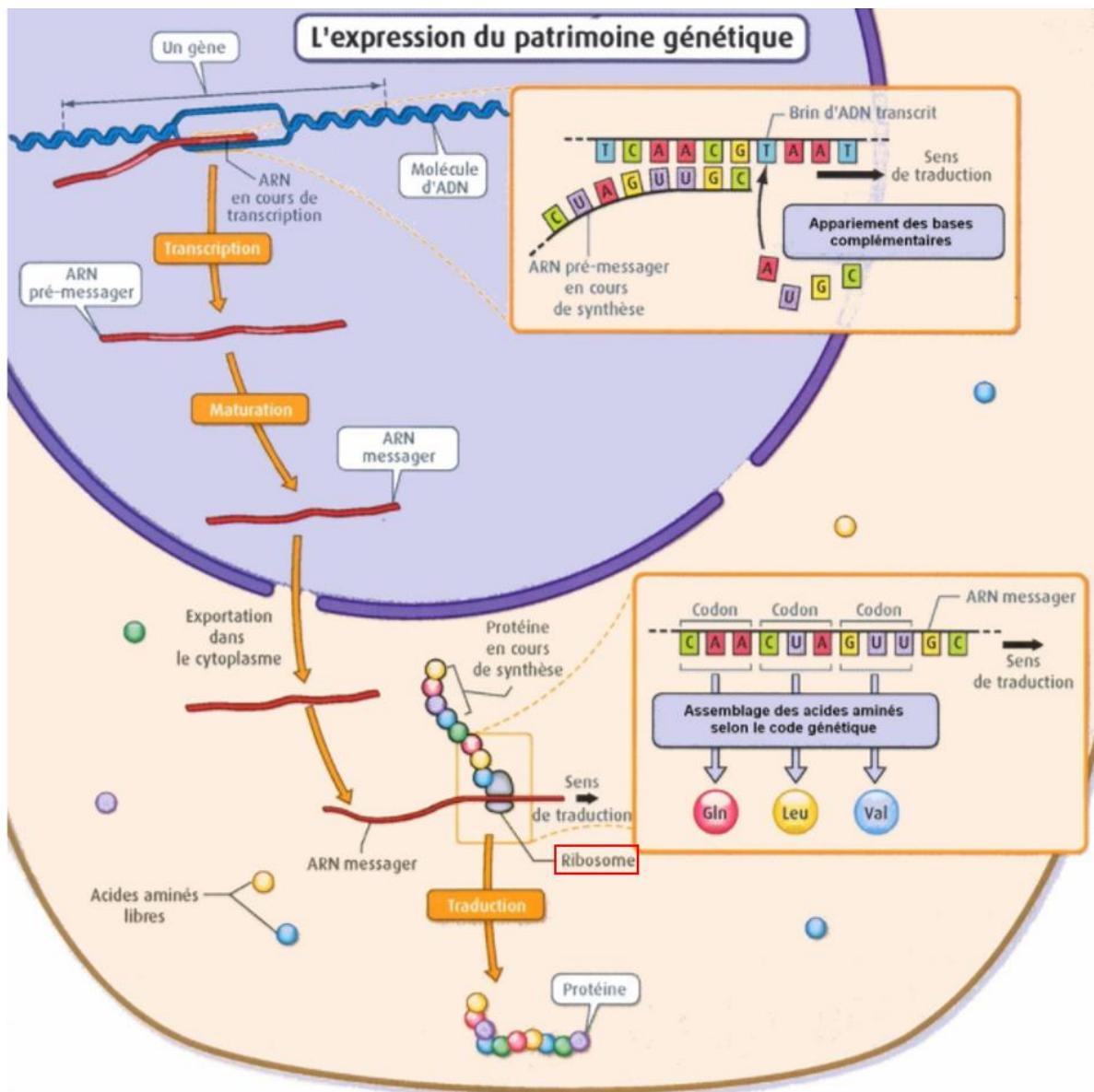


Figure 1 : Expression génique chez les eucaryotes

L'expression des gènes est composée de la transcription des gènes en pré-ARN messagers, de la maturation et de l'export de ces messagers puis de la traduction des ARNm en protéine par le ribosome. La synthèse des protéines se fait suivant le code génétique, qui assigne un acide aminé à chaque codon. *Adapté de <http://b.21-bal.com/pravo/1775/index.html>*

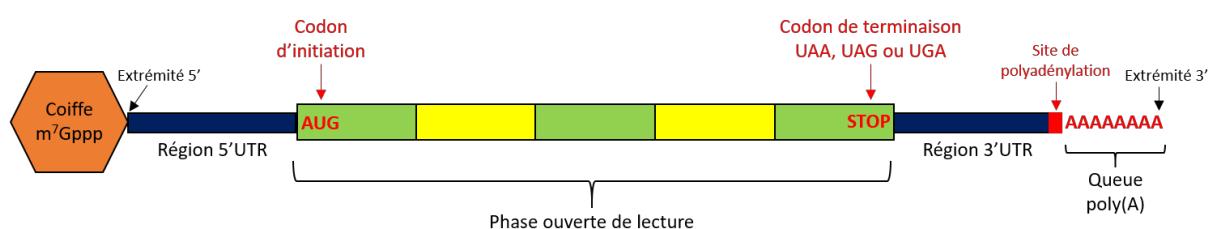


Figure 2 : Structure et éléments caractéristiques d'un ARN messager

L'ARN messager comprend une coiffe m⁷Gppp à l'extrémité 5', deux régions non codantes (UTR) en 5' et en 3', une queue poly(Adénine) en 3' et une phase ouverte de lecture délimitée par un codon d'initiation de la traduction AUG et un codon de terminaison de la traduction en 3' UAA, UAG ou UGA.

❖ Les ARNs messagers (ARNm) :

Les ARNm représentent 5% des ARNs totaux de la cellule¹ et comprennent, chez les eucaryotes, une coiffe à l'extrémité 5', une région non traduite en 5' (5'UTR, UnTranslated Region), une région codante délimitée en 5' par un codon d'initiation de la traduction (généralement AUG) et en 3' par un codon de terminaison UAA, UAG ou UGA, une région non traduite en 3' (3'UTR) et une queue poly(Adénine) (poly(A)) à l'extrémité 3' (Figure 2).

Les ARNm sont synthétisés sous forme de pré-ARNm par l'ARN polymérase Pol II à partir de la séquence d'ADN : c'est l'étape de transcription. Ces pré-ARNm subissent alors une maturation importante, puis sont exportés du noyau vers le cytoplasme pour y être traduit. La maturation se compose de plusieurs étapes : ajout d'une coiffe 7-méthylguanosine triphosphate, épissage des introns, clivage en 3', ajout d'une queue poly(A) et ajout de modifications chimiques. Toutes ces étapes ont lieu de manière co-transcriptionnelle, bien que la transcription et la maturation soient indépendantes².

L'ajout de la coiffe en 5' a lieu de façon précoce au cours de la transcription, le plus souvent après la synthèse d'environ vingt nucléotides. Elle se déroule en trois étapes : (1) l'extrémité 5' triphosphorylée est convertie en ARN diphosphorylé, (2) une molécule GMP est ajoutée pour former une coiffe GpppRNA, (3) et enfin le carbone 7 de la guanine est méthylé pour donner un m⁷GpppRNA³ (Figure 3a). La coiffe est essentielle : elle permet de stabiliser le messager³, elle est reconnue par les facteurs d'initiation de la traduction, qui recrutent le ribosome³, et elle joue également un rôle essentiel dans l'épissage en le stimulant⁴.

De manière intéressante, certains ARNm présentent une coiffe triméthylée, qui intervient alors dans la régulation de leur traduction. Jusqu'à présent cette coiffe particulière n'a été identifiée que sur des ARNm codant des sélénoprotéines chez les mammifères⁵, même si la triméthylation de la coiffe a aussi été mise en évidence pour des snRNAs et snoRNAs⁶.

Les pré-ARNm eucaryotes acquièrent une queue poly(A), après un clivage en 3'. Cette queue d'adénines protège de la dégradation, a un rôle dans l'export de l'ARNm mature du noyau vers le cytoplasme et favorise la traduction³. Le clivage en 3' fait intervenir des éléments de reconnaissance, sous forme de motifs de nucléotides. Chez les mammifères, le motif nucléotidique majeur PAS (Polyadenylation (poly(A) signals) est un signal hexanucléotidique poly(A), souvent AAUAAA), dix à trente nucléotides avant le site de clivage. On trouve également le site de clivage lui-même (souvent après un dinucléotide CA) et un élément de séquence riche en U ou en G/U en aval

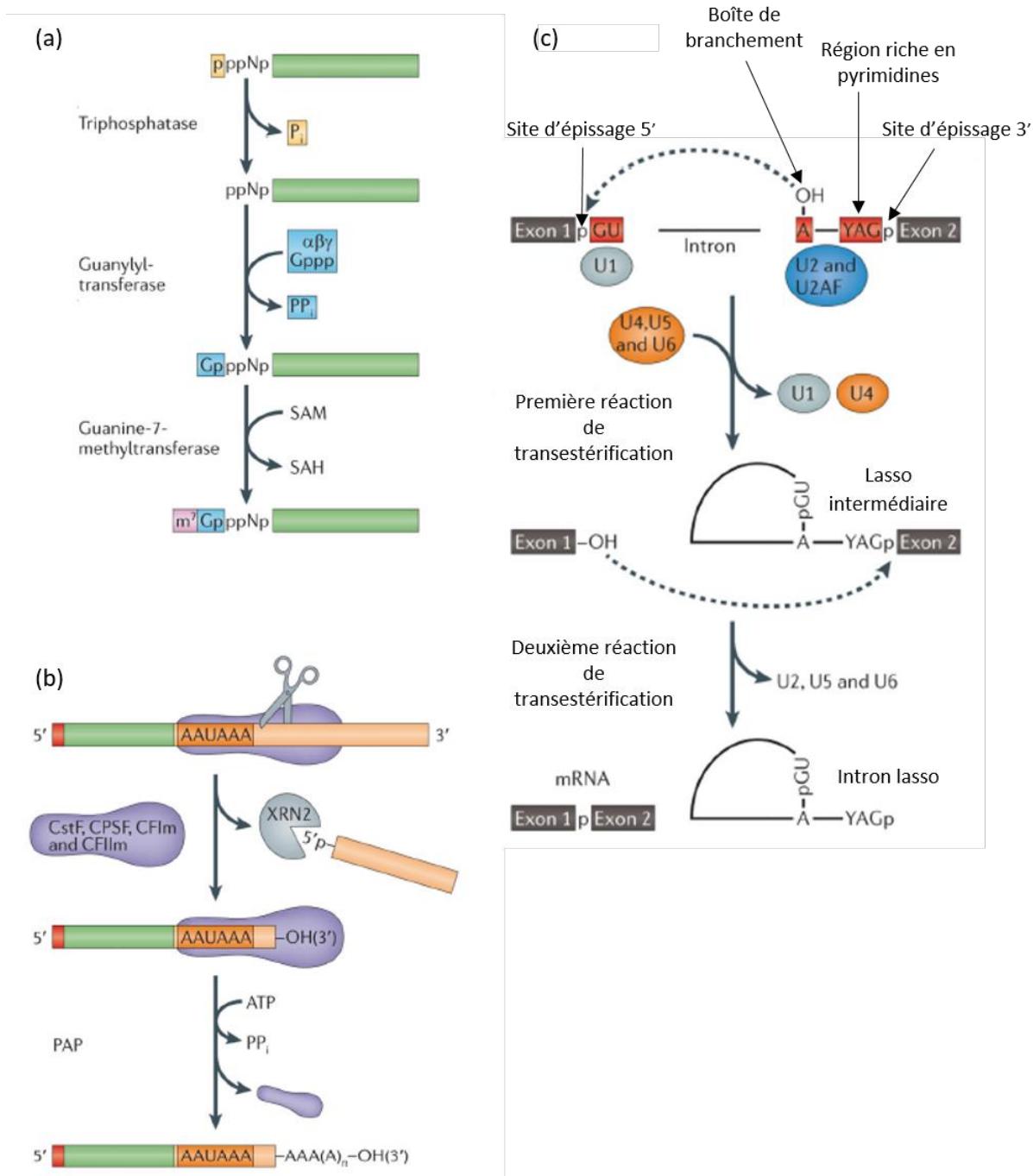


Figure 3 : Etapes de maturation du pré-ARN messager en ARN messager mature

(a) Ajout de la coiffe m^7Gppp en 5'. Une triphosphatase convertit l'extrémité 5' triphosphorylé en ARN diphosphorylé, puis une guanylyl-transférase ajoute une molécule GMP pour former une coiffe GpppRNA, et une guanyl-7-méthyltransférase méthyle le carbone 7 de la guanine à partir du donneur SAM pour donner un m^7Gppp RNA. (b) Ajout d'une queue poly(A). L'extrémité 3' est tout d'abord clivé 10 à 30 nucléotides après le signal de polyadénylation AAUAAA. Les complexes CstF, CPSF, CFIm et CFIIIm reconnaissent ce motif et clivent l'ARNm en 3'. Puis la queue poly(A), synthétisée par la poly(A) polymérase PAP est ligaturée au site de clivage. (c) Epissage. L'intron à épisser est composé de deux sites d'épissage en 5' et en 3', d'une boîte de branchement avec une Adénine qui joue un rôle central, et d'une région riche en pyrimidines chez les eucaryotes supérieurs. La première réaction de transestérification correspond à l'attaque nucléophile du 2'-OH du ribose de l'adénosine

(DSE)^{3,7}. La machinerie responsable du clivage de l'ARNm en 3' et de l'ajout de la queue poly(A) est complexe et comprend plusieurs facteurs dont : le complexe CPSF (cleavage and polyadenylation specificity factor), le complexe CstF (cleavage stimulation factor), le facteur CFI_m (cleavage factor I), le facteur CFI_m et d'autres protéines comme la poly(A) polymérase (PAP). CPSF-73, sous-unité du complexe CPSF, est l'endoribonucléase responsable de la réaction de clivage en 3', CPSF-160 reconnaît le PAS, CstF-64 reconnaît le DSE³. La queue poly(A) est synthétisée à l'extrémité 3' par PAP et sa taille est contrôlée par PABPN1. Elle atteint environ 80 nucléotides chez la levure *S. cerevisiae* et 250 nucléotides chez les mammifères⁸. Plusieurs études rapportent des observations de variations de site de poly(A), dans différents types cellulaires, différentes conditions pathologiques mais sans lien avec la stabilité des ARNm isoformes⁹⁻¹² (Figure 3b). En effet, plus de 70% des ARNm humains présentent plusieurs PAS, c'est ce qu'on appelle la polyadénylation alternative (APA).

Avant l'initiation de la traduction, des interactions reliant la coiffe et la queue poly(A) se mettent en place, ce qui aboutit à une circularisation de l'ARNm^{13,14}. Les mécanismes mis en place lors de cette circularisation seront détaillés dans le II-1-a qui traite de l'initiation de la traduction.

Les ARNm des eucaryotes possèdent des séquences codantes nommées exons interrompues par des séquences non codantes appelées introns, éliminées lors de l'étape d'épissage.

Les introns possèdent des éléments dans leur séquence qui leur permettent d'être reconnus comme tels par le spliceosome, la machinerie qui permet l'épissage. Ces éléments sont deux sites d'épissage en 5' et en 3' et une boîte de branchement, qui contient une adénine qui a un rôle central et qui est localisé 18 à 40 nucléotides en amont du site d'épissage 3'¹⁵.

Le spliceosome contient les snRNP (small nuclear ribonucleoprotein) U1, U2, U4/U6 et U5 et d'autres facteurs protéiques. D'abord les snRNPs U1 et U2 s'assemblent respectivement aux sites d'épissage 5' et 3'. Puis le complexe ternaire U4/6-U5 est recruté pour former le spliceosome mature et catalyser la coupure de l'ARNm, l'excision des introns et la religature des exons¹⁶.

L'épissage peut se découper en deux étapes : (1) Le 2'-OH du ribose de l'adénine de la boîte de branchement effectue une attaque nucléophile sur le phosphate de la jonction exon-intron en 5'. (2) Le 3' de l'exon alors libre attaque le phosphate de la jonction intron-exon en 3'. L'intron forme alors un lasso et est excisé de l'ARNm, les deux exons de part et d'autre de l'intron sont reliés¹⁵ (Figure 3c).

En moyenne, les ARNm humains contiennent huit exons et sept introns¹⁷. Cependant, la combinaison d'exons retenus dans l'ARNm mature peut varier, ce qui augmente considérablement

de la boîte de branchement sur le phosphate de la jonction exon1-intron. L'intron et l'exon2 forment alors un lasso intermédiaire. La deuxième réaction de transestérification correspond à l'attaque du 3' de l'exon1 libre sur le phosphate de la jonction intron-exon2. L'intron forme alors un lasso et est excisé de l'ARNm, les exons 1 et 2 sont reliés entre eux. *Adapté de Bentley, 2014*².

la diversité du protéome. C'est l'épissage alternatif. Chez l'homme, plus de 95% des ARNm subissent l'épissage alternatif¹⁸.

Les ARNm présentent des modifications chimiques. 110 modifications chimiques des ARNs (ARNm, ARNt, ARNr et autres) ont été décrites¹⁹. Bien que les ARNt soient les ARNs les plus modifiés²⁰, les ARNm présentent également une variété importante de modifications chimiques - dont notamment douze types d'ajout de groupement méthyle sur les quatre types de nucléotides¹⁹ - parmi lesquelles on retrouve des N⁶-méthyladénosines (m⁶A), des 5-méthylcytosines (m⁵C), des pseudouridines (Ψ), des N¹-méthyladénosines (m¹A) ou encore des déaminations des adénosines et des cytosines. Ces modifications sont un des mécanismes de régulation post-transcriptionnelle de l'expression des gènes et peuvent avoir un effet sur la vitesse d'elongation de la traduction, sur la fidélité de la traduction ou encore sur la terminaison prématuée de la traduction²¹. La découverte et l'étude des nombreuses modifications chimiques des ARNm ont élargi les mécanismes impliqués dans l'épigénétique, qui repose sur le fait que des modifications non codées par la séquence d'ADN seraient porteuses d'information et permettraient la régulation de l'expression des gènes.

❖ Les ARNs de transfert (ARNt) :

Les ARNt représentent 15% des ARNs totaux dans les cellules¹. Ces ARNs, synthétisés par l'ARN polymérase Pol III²², font une taille d'environ 75nt²³.

Francis Crick, dans sa lettre de 1955 (« On Degenerate Templates and the Adaptor Hypothesis »), émet une hypothèse sur le rôle des ARNs de transfert dans la traduction. Cette hypothèse est la suivante : les acides aminés sont attachés à des molécules d'ARN qui décodent l'ARNm par complémentarité de séquence avec cet ARNm. En 1958, l'interaction entre les acides aminés et les ARNt est démontrée par Zachau et ses collaborateurs²⁴ : la liaison de l'acide aminé à l'extrémité 3' adénine de l'ARNt a lieu grâce à une liaison ester avec le groupement OH en 2' ou en 3' du ribose de l'adénosine.

L'étude des séquences des ARNt a permis de dessiner une structure secondaire en forme de trèfle à quatre feuilles, basée sur les complémentarités Watson-Crick, similaire pour tous les ARNt²². Cette structure est composée de quatre hélices en épingle à cheveux et trois boucles majeures. L'extrémité 3' se termine par le triplet de nucléotide CCA non apparié et impliqué dans le chargement de l'acide aminé²⁵. Au sein de la structure en forme de trèfle, on peut distinguer différentes régions : la tige acceptrice sur laquelle l'acide aminé est attaché par estérification, la boucle de l'anticodon qui reconnaît le codon à décoder par complémentarité de

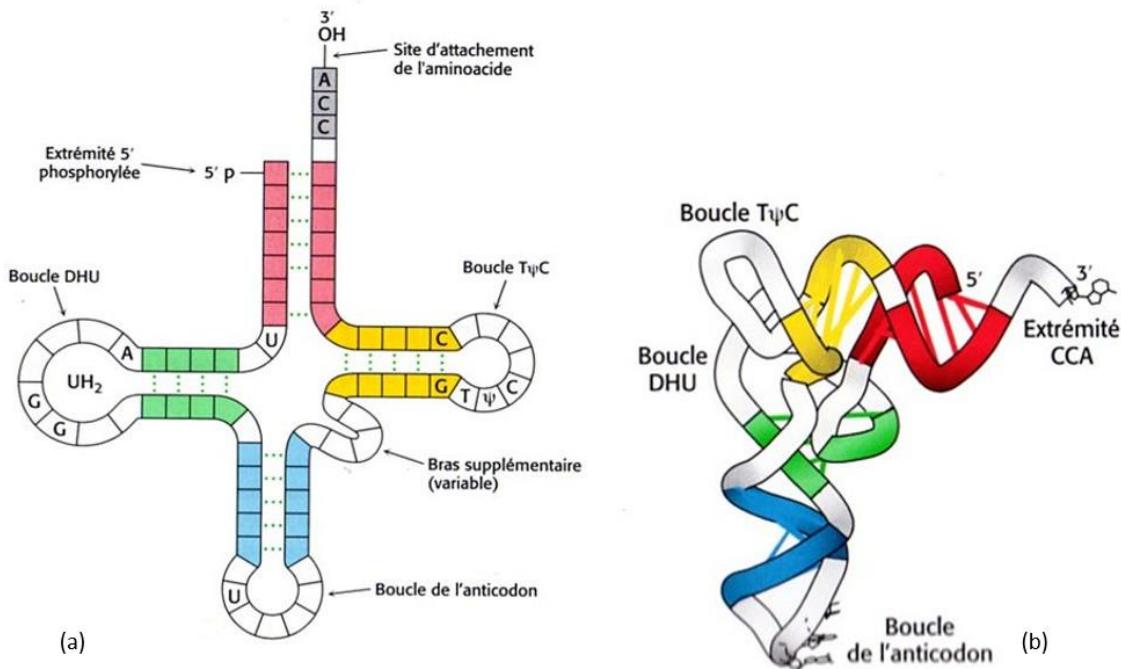


Figure 4 : Structures secondaire et tertiaire des ARNs de transfert

(a) Structure secondaire d'un ARN de transfert, en forme caractéristique de trèfle à quatre feuilles. On peut reconnaître les structures remarquables suivantes : la tige acceptrice, en rose, sur laquelle l'acide aminé est attaché par estérification, la boucle de l'anticodon qui reconnaît le codon à décoder par complémentarité de séquence, la boucle DHU, la boucle T Ψ C et le bras supplémentaire variable (Barciszewska, 2016). (b) La structure tertiaire en forme caractéristique de L est obtenue par repliement de la tige DHU (en vert) sur la tige T Ψ C (en jaune) (Barciszewska, 2016). *Adapté de Berg, Tymoczko & Stryer, Biochemistry, 5th ed., 2001: Fig. 29.4: General secondary structure of tRNA molecules (the "cloverleaf" secondary structure).*

séquence, la boucle DHU (qui doit son nom au ribonucléotide modifié dihydrouridine), la boucle T ψ C (qui doit son nom à la présence d'une ribothymidine et d'une pseudouridine) et la boucle variable²². La structure en trois dimensions a également été déterminée, pour la première fois en 1973²⁶. Tous les ARNt présentent une structure 3D en forme de L dans laquelle la tige DHU se replie sur la tige T ψ C. Cette structure est composée de deux hélices : la tige DHU et la tige anticodon forment le bras anticodon et la tige acceptrice et la tige T ψ C forment le bras accepteur²² (Figure 4).

Les ARNt sont fortement maturés après leur synthèse et présentent de nombreuses modifications chimiques de leurs nucléotides²². La maturation s'effectue en plusieurs étapes : coupure de l'extrémité 5' par la RNase P, coupure de l'extrémité 3' et addition du triplet CCA, épissage des introns et modifications post-transcriptionnelles des nucléotides²². Au sein des ARNt, on retrouve environ 80 types de modifications différentes²⁷ catalysées par plus de 100 enzymes²⁸. Ces modifications stabilisent les ARNt²⁹ et jouent un rôle très important dans l'efficacité et la fidélité de la traduction. Elles sont principalement situées dans la boucle anticodon et à proximité et sont impliquées dans la complémentarité codon-anticodon et le processus de décodage de l'ARNm³⁰. Ces modifications sont conservées au sein de chacun des trois règnes du vivant (bactéries, archées, eucaryotes) et quelques-unes sont conservées au sein des trois règnes, ce qui confirme leur importance fonctionnelle³¹.

Les ARNt sont chargés en acides aminés par les aminoacyl-ARNt synthétases (aaRS), qui catalysent la réaction d'estérification entre les ARNt et les acides aminés, pour former les aminoacyl-ARNt (aa-ARNt). Cette étape est irréversible, le taux d'erreur est de 10^{-4} à 10^{-5} chez les procaryotes³². Il existe une aaRS pour chaque acide aminé qui sera chargé sur ses différents ARNt³³.

❖ Le ribosome :

Le ribosome est un complexe ribonucléoprotéique d'environ 4,6 MDa, composé d'ARN (ARNr pour ARN ribosomique) et de protéines, dans lequel on retrouve des interactions ARN-ARN, protéine-protéine et ARN-protéine. Ce complexe dynamique est la plateforme de décodage de l'ARN messager et est au cœur de la synthèse des protéines. Ce sont les ARNr qui présentent l'activité catalytique : le ribosome est un ribozyme³⁴. Il se compose de deux sous-unités asymétriques, nommées par leur coefficient de sédimentation : 30S pour la petite sous-unité procaryote, 40S pour la petite sous-unité eucaryote et 50S pour la grande sous-unité procaryote, 60S pour la grande sous-unité eucaryote. L'ensemble forme le ribosome 70S procaryote et le ribosome 80S eucaryote. Ces deux sous-unités vont s'assembler autour d'un ARNm et se déplacer le long de celui-ci tout en synthétisant

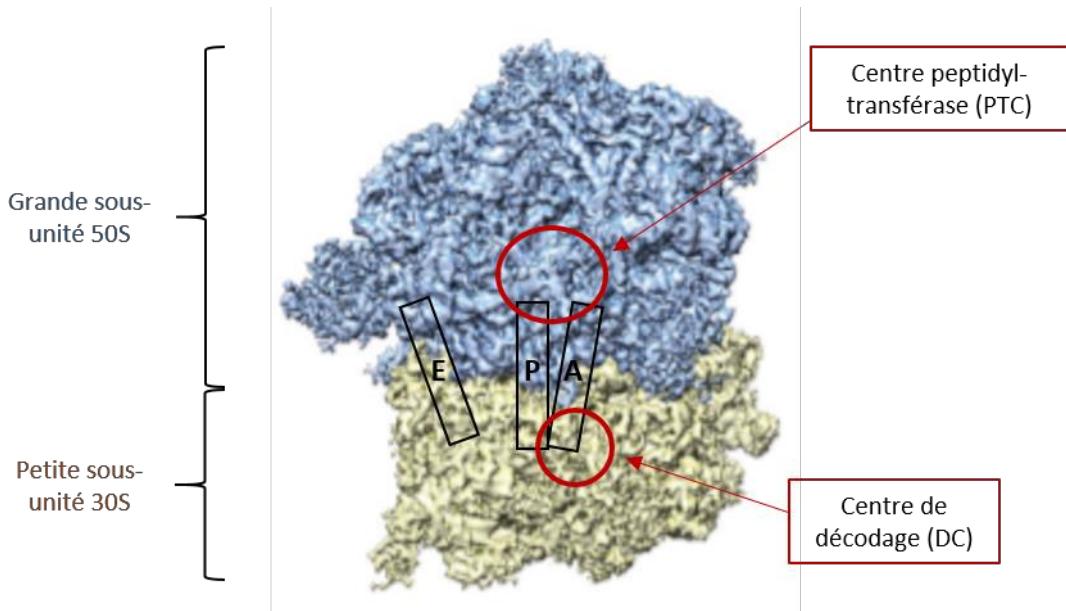


Figure 5 : Structure générale du ribosome eucaryote 80S

Le ribosome est un complexe composé de protéines et d'ARN. Chez les eucaryotes, le ribosome complet 80S est composé d'une petite sous-unité 40S (en beige) et d'une grande sous-unité 60S (en bleu). Il comprend, à l'interface de ces deux sous-unités, trois sites nommés A (aminoacyl site), P (peptidyl site) et E (exit site), dans lesquels se succèdent les ARNt (Selmer, 2006). Le ribosome présente également deux sites catalytiques très importants : le centre de décodage (DC), où a lieu le décodage de l'ARNm, et le centre peptidyl-transférase (PTC) où est synthétisée la liaison peptidique qui relie un nouvel acide aminé au peptide naissant en cours d'elongation (Selmer, 2006). *Adapté de Sun, 2015³⁵.*

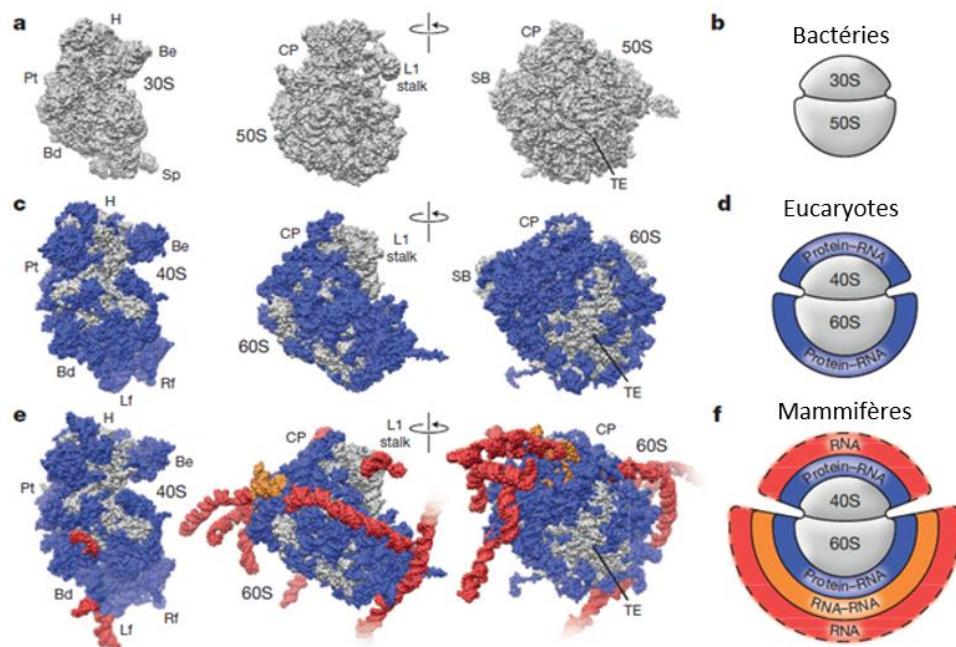


Figure 6 : Evolution et extensions du ribosome

Représentations en surface (a, c, e) et schématique (b, d, f) des ribosomes de *T. thermophilus* (a, b), *S. cerevisiae* (c, d) et *H. sapiens* (e, f). Les régions en grises sont celles universellement conservées. En bleu sont représentées les protéines et extensions d'ARN spécifiques des eucaryotes, en orange et en rouge les ARNs et leurs extensions spécifiquement retrouvés chez les mammifères. *Adapté de Anger, 2003³⁶.*

la protéine correspondante. Le ribosome comprend, à l'interface de ces deux sous-unités, trois sites nommés A (site aminoacyl), P (peptidyl site) et E (exit site), dans lesquels se succèdent les ARNt³⁷. Le ribosome présente également deux sites catalytiques très importants : le centre de décodage (DC), où a lieu le décodage de l'ARNm, et le centre peptidyl-transférase (PTC) où est synthétisée la liaison peptidique qui relie un nouvel acide aminé au peptide naissant en cours d'elongation³⁷ (Figure 5). Le cœur du ribosome est très conservé entre les espèces des trois règnes du vivant. Cependant, le ribosome eucaryote est plus complexe que le ribosome procaryote et présente des extensions des ARNr (environ 2650 nucléotides supplémentaires), des protéines ribosomiques supplémentaires ainsi que des extensions à certaines protéines (environ 2450 acides aminés supplémentaires)³⁶. Cette augmentation de la complexité concerne principalement la sous-unité 60S, et s'observe aussi entre les eucaryotes : ribosome humain est plus complexe que le ribosome de drosophile, lui-même plus complexe que le ribosome de levure³⁶ (Figure 6). Le ribosome sera décrit plus en détails dans le paragraphe suivant.

❖ Autres acteurs de la traduction :

La traduction fait également intervenir des centaines de facteurs protéiques et nucléotidiques au cours de la biosynthèse des différents acteurs principaux, au cours des étapes de la traduction ainsi que pour réguler la traduction.

2 – La structure et la composition du ribosome

Les ribosomes ont été observés pour la première fois en 1938 par Albert Claude. Alors qu'il étudiait le virus du sarcome de Rous, il détecte au microscope, en plus des virus, de petites particules présentes dans différents tissus, qu'il nomme « microsomes ». Ce n'est que 20 ans plus tard que George Palade obtient une image plus claire de ces particules par micrographie électronique, dont il estime la taille à environ 20 nanomètres. À la même époque, Paul Zamecnik et ses collègues utilisent des acides aminés radioactifs pour montrer que ces microsomes produisent des protéines. Ils seront renommés « ribosomes » en 1958, lorsque Richard Roberts établit que leur composition est riche en acide ribonucléique (ARN).

La petite sous-unité du ribosome est composée de l'ARNr 16S et de 20 protéines chez les procaryotes et de l'ARNr 18S et de 33 protéines chez les eucaryotes. Elle peut être divisée en différentes régions : la tête, le bec, la plateforme, l'épaule, le corps et les pieds gauche et droit. La grande sous-unité est composée des ARNr 23S et 5S et de 30 protéines chez les procaryotes et des

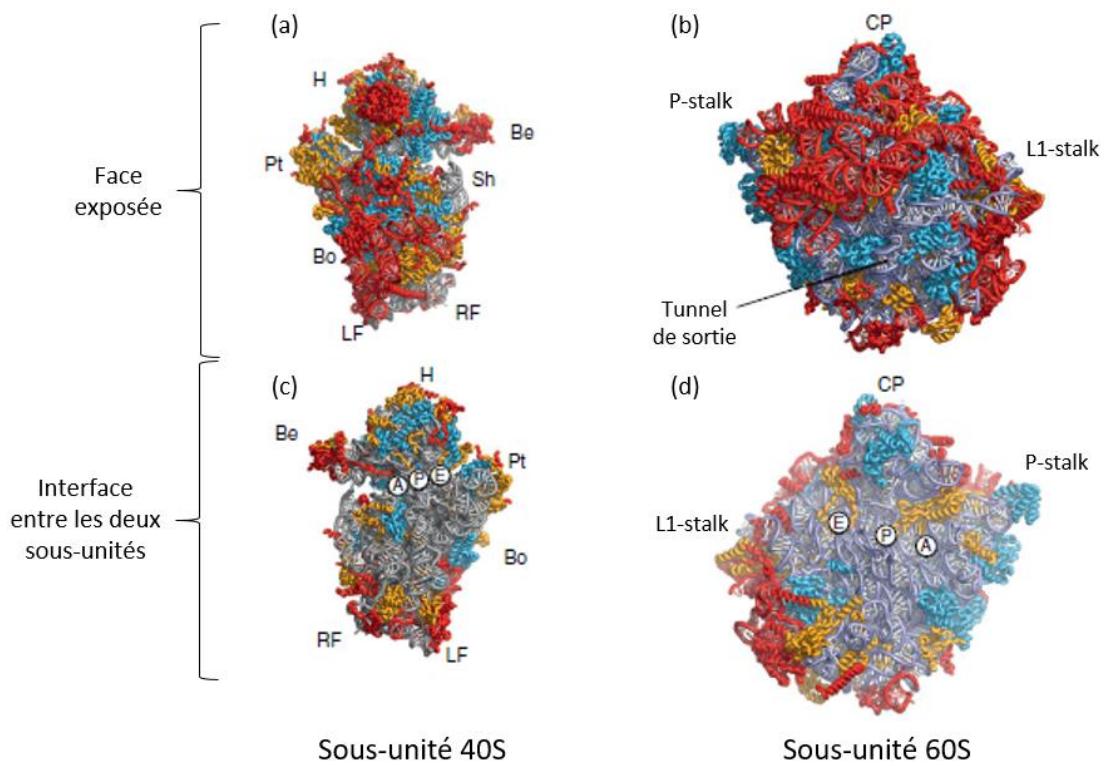


Figure 7 : Structure cristallographique du ribosome eucaryote 80S

Représentations de la face exposée (a, b) et de la face à l'interface entre les deux sous-unités (c, d), de la sous-unité 40S (a, c) et de la sous-unité 60S (b, d) du ribosome de *T. thermophila*. Les protéines universellement conservées sont représentées en bleu clair, les protéines présentes chez les archées et les eucaryotes sont en jaune et les protéines et les éléments ARN présents exclusivement chez les eucaryotes sont en rouge. On peut distinguer les structures remarquables suivantes : la tête (H), le bec (Be), la plateforme (Pt), l'épaule (Sh), le corps (Bo) et les pieds gauche (LF) et droit (RF) dans la sous-unité 40S; la protubérance centrale (CP), la L1-stalk et la P-stalk dans la sous-unité 60S. On retrouve également les sites de positionnement des ARNt A, P et E. *Adapté de Klinge, 2012*³⁸.

ARNr 28S (25S chez la levure *Saccharomyces cerevisiae*), 5S et 5.8S et de 47 protéines chez les eucaryotes. Elle présente trois régions remarquables : la protubérance centrale, la L1-stalk et la P-stalk³⁸ (Figure 7).

L'interface entre les deux sous-unités comporte plusieurs régions de contact afin de maintenir le ribosome complet 70S ou 80S au cours de la traduction. La région principale est entre l'hélice H69 (ARNr 28S) de la grande sous-unité et l'hélice h44 (ARNr 18S) de la petite sous-unité pour former les ponts B2a et B2b. Ces interactions entre les deux sous-unités sont de différentes natures : ribose-ribose, ribose-phosphate, ribose-base, protéine-base et protéine-protéine (Figure 8). Un type particulier d'interactions implique les 2'-O-méthylations des riboses des ARNr 18S et 28S : les groupements méthyles sont décalés, les liaisons hydrogènes entre les deux ARNr sont ainsi en forme de fourche³⁹.

D'un point de vue fonctionnel, le ribosome comprend, à l'interface entre les deux sous-unités, trois sites A, P et E brièvement décrits dans le paragraphe précédent. Le site A est le site où vient se fixer l'ARNt correspondant au codon de l'ARNm à ce site, grâce aux interactions codon-anticodon. Ce site fait partie du centre de décodage (DC) où a lieu le décodage de la séquence d'ARN messager et qui a un rôle très important dans la fidélité de la traduction. Le site P est le site où l'ARNt qui est présent porte le peptide en cours d'elongation. Le site E présente un ARNt dit déacétylé, c'est-à-dire dépourvu d'acide aminé. L'ARNt au site E est expulsé du ribosome. Le site catalytique du ribosome est le centre peptidyl-transférase (PTC) : c'est là qu'a lieu la formation de la liaison peptidique entre le nouvel acide aminé et le peptide en cours d'elongation³⁷.

Les protéines ribosomiques présentent des modifications post-traductionnelles et les ARNs ribosomiques des modifications chimiques. Une fraction significative de nucléotides des ARNr est modifiée : environ 2% soit plus de 100 positions chez la levure et plus de 200 chez l'homme⁴⁰. Il y a douze types de modifications distinctes, réalisées soit par une enzyme, soit par un complexe ribonucléoprotéique guidé par un ARN antisens, un snoRNP⁴⁰. Les 2'-O-méthylations et les pseudouridylations sont les modifications les plus fréquentes (55 2'-O-méthylations (2'-O-Me) et 45 pseudouridylations (ψ) chez la levure, 106 2'-O-Me et 97 ψ chez l'homme)⁴⁰. Les 2'-O-méthylations correspondent à un ajout d'un groupement méthyle sur le ribose du nucléoside et les pseudouridylations correspondent à une isomérisation d'une uridine par une rotation à 180° autour de l'axe N3-C6 (Figure 9). Ces deux types de modifications sont synthétisés par des snoRNPs constitués d'un ARN guide antisens complémentaire de la région à modifier, d'une protéine catalytique et de protéines accessoires. Les bases des ARNr peuvent aussi être modifiées. Des groupements méthyles peuvent être ajoutés sur les positions 1, 6 et 7 des purines (Adénine et Guanine) et sur les positions 1, 3 et 5 des pyrimidines (Uridine et Cytosine). Les

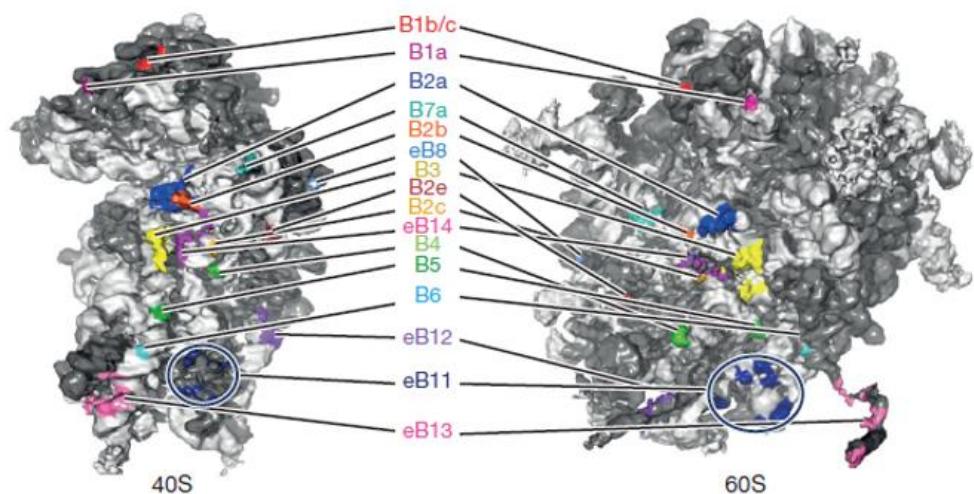


Figure 8 : Ponts entre les sous-unités du ribosome

Les deux sous-unités du ribosome sont en contact par des ponts, qui sont établis par des interactions de différentes natures : ribose-ribose, ribose-phosphate, ribose-base, protéine-base et protéine-protéine. *Adapté de Khatter, 2015*³⁹.

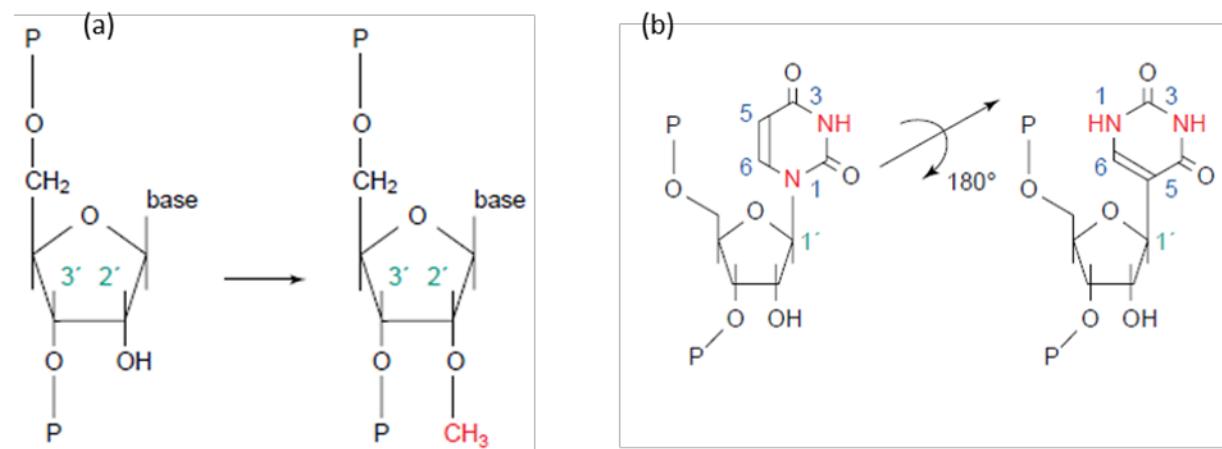


Figure 9 : Réactions de 2'-O-méthylation et de pseudouridylation

Les 2'-O-méthylations (a) correspondent à un ajout d'un groupement méthyle sur le ribose du nucléoside et les pseudouridylations (b) correspondent à une isomérisation d'une uridine par une rotation à 180° autour de l'axe N3-C6. *Adapté de Lafontaine, 1998*⁴¹.

pyrimidines peuvent également être aminocarboxypropylés en position 3 et acétylés en position 4⁴⁰. Certaines positions peuvent présenter plusieurs modifications. Toutes ces modifications sont présentes au cœur du ribosome, au niveau des régions fonctionnelles notamment (DC et PTC), et à l'interface entre les deux sous-unités, et absentes des régions périphériques (Figure 10). Leurs positions sont conservées entre les espèces, ce qui révèle un rôle fonctionnel important. En général elles modifient la structure et la réactivité des bases concernées. La fonction exacte de ces modifications n'est pas complètement comprise, et bien que l'absence de certaines modifications ou d'enzymes responsables de ces modifications soient impliquée dans des pathologies, on ne sait pas si c'est l'absence de modifications chimiques ou un mauvais assemblage du ribosome qui est à l'origine de la pathologie associée⁴⁰.

Certains facteurs protéiques peuvent également être associés au ribosome sans être strictement définis comme des protéines ribosomiques. Une étude protéomique réalisée chez *S. cerevisiae* a révélé 77 protéines associées au ribosome. La délétion de certaines de ces protéines entraîne des défauts de vitesse et de fidélité de la traduction⁴². Ces protéines pourraient donc participer à la régulation de la traduction.

Chez *D. melanogaster*, la protéine pro-apoptotique Reaper se lie à la sous-unité 40S et empêche la reconnaissance du codon d'initiation AUG par le complexe 48S lors de l'étape de scanning. La traduction cap-dépendante est ainsi inhibée. Reaper pourrait plutôt faciliter la traduction spécifique d'ARNm pro-apoptotique qui possèdent des IRES et dont la traduction ne nécessite pas de scanning, favorisant l'apoptose⁴³.

Dans des cellules HeLa, la protéine GYS1, qui est impliquée dans la biogenèse du glycogène et dont l'activité est inhibée par le glycogène lui-même, est également retrouvée associée aux ribosomes qui traduisent activement. Cette association nécessite la phosphorylation de GYS1 sur la Ser640. Une extinction de GYS1 entraîne une diminution des polysomes et un changement des ARNm associés aux polysomes. Aucun enrichissement fonctionnel ou motif structural commun aux gènes correspondants n'a été trouvé, cependant plusieurs de ces ARNm codent des protéines impliquées dans la synthèse protéique. GYS1 serait probablement incluse dans une boucle de rétroaction entre l'état énergétique de la cellule et la machinerie de traduction⁴⁴.

La protéine RACK1 (chez les mammifères, Asc1p chez la levure) est le récepteur de la protéine kinase C activée (PKC) et est une protéine plateforme qui s'associe à la sous-unité 40S avec un ratio 1:1. De nombreuses molécules de signalisation interagissent avec RACK1 au niveau du ribosome et affectent la synthèse protéique. Par exemple, la protéine PKC β II activée se lie à RACK1 associée au ribosome et phosphoryle le facteur d'initiation de la traduction eIF6, ce qui entraîne sa libération de la sous-unité 60S et permet donc l'assemblage des deux sous-unités⁴⁵. Parmi les autres

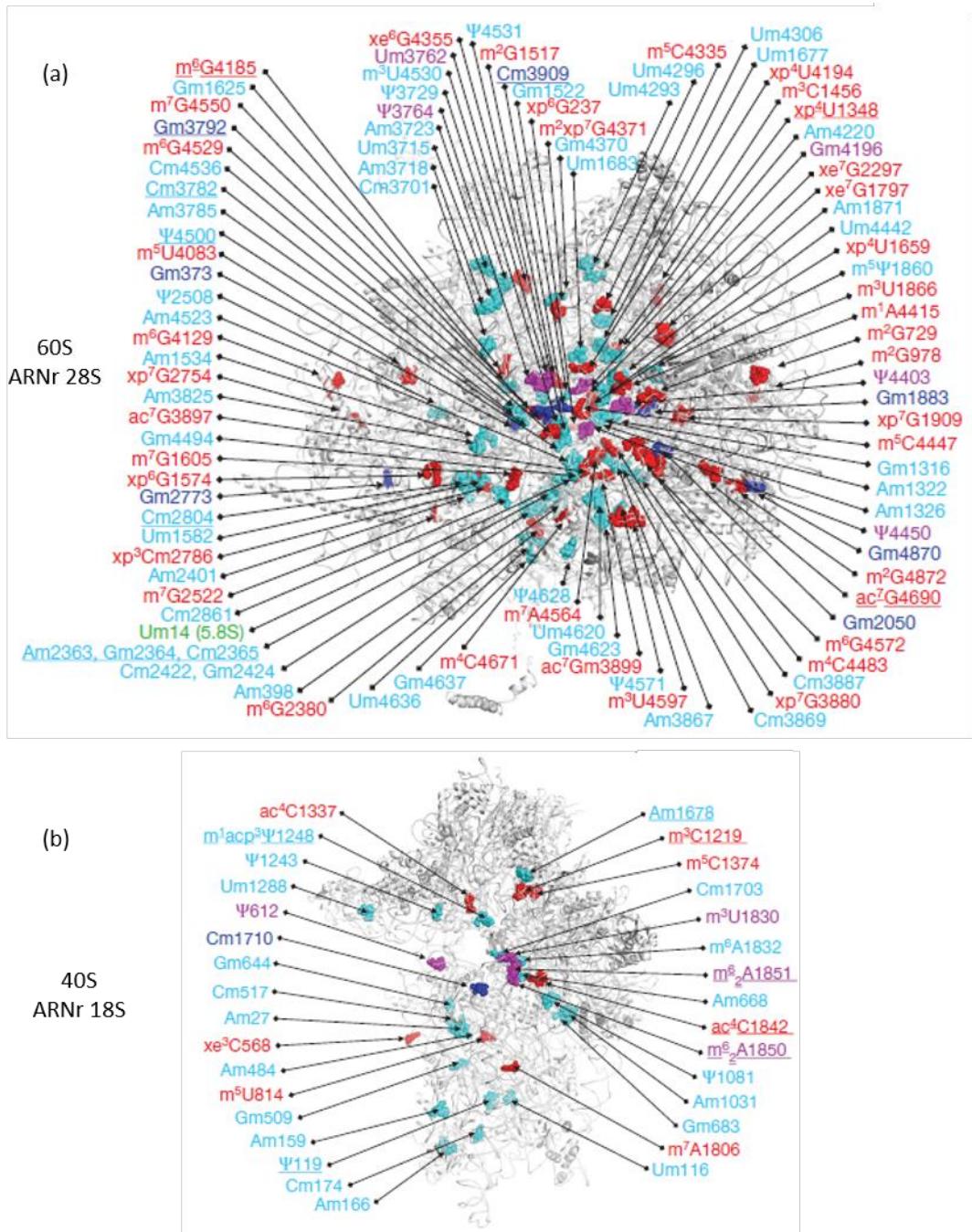


Figure 10 : Modifications chimiques des ARNr des sous-unités 40S et 60S du ribosome humain

Annotation des nucléotides modifiés dans l'ARNr 28S (sous-unité 60S) (a) et dans l'ARNr 18S (sous-unité 40S) (b). En violet : positions conservées entre *E. coli* et *H. sapiens* ; Bleu clair : sites prédis et révélés ; Bleu foncé : sites de 2'-O-méthylations non prédis ; Rouge : sites de modifications de base non prédis. *Adapté de Natchiar, 2017*⁴⁶.

partenaires d'interaction de RACK1 on trouve les récepteurs intégrines, qui peuvent potentiellement localiser les ribosomes à la membrane cellulaire. Chez la levure, Asc1p interagit avec la protéine de liaison à l'ARN Scp160p. Scp160p s'associe avec des ARNm spécifiques et sa délétion entraîne une réduction de l'association des ribosomes à ces ARNm au sein des polysomes⁴⁷.

Ces trois exemples illustrent la régulation de la traduction de sous-groupes d'ARNm par des facteurs non ribosomiques associés aux ribosomes⁴⁸.

3 – La biogenèse du ribosome

Dans une cellule de levure en croissance rapide, 60% de la transcription est dévouée aux ARNs ribosomiques par l'ARN Pol I (80% des ARNs d'une cellule sont des ARNr) et 50% de la transcription par l'ARN Pol II est dédiée aux ARNm codant des protéines ribosomiques¹. Dans les cellules de mammifères en prolifération, environ 35% de la transcription est dédiée à la production d'ARNr⁴⁹.

La biogenèse des ribosomes fait intervenir plusieurs centaines de facteurs d'assemblage et de complexes snoRNPs⁵⁰⁻⁵² et nécessite une importante régulation qui est étroitement liée à d'autres processus cellulaires comme la croissance et la division cellulaire. Des défauts dans la biogenèse du ribosome sont associés à des pathologies humaines, notamment des cancers⁵³.

La maturation des composants du ribosome a lieu à la fois dans le noyau et dans le cytoplasme après export des pré-sous-unités 40S et 60S⁵³.

A – La synthèse et la maturation des composants du ribosome

a – La synthèse et la maturation des protéines ribosomiques

Les protéines ribosomiques (PR) sont synthétisées à partir d'ARNm transcrits par l'ARN Pol III et contiennent très souvent des introns, qui seront éliminés lors de l'épissage.

Comme la plupart des protéines, les protéines ribosomiques, subissent des modifications post-traductionnelles communes, telles que la perte de la méthionine initiatrice, des acétylations, des méthylations, des phosphorylations, des ponts disulfures⁵⁴, et d'autres modifications plus spécifiques comme l'ajout de O-linked β -d-N-acétylglucosamine (O-GlcNAc). Chez l'Homme, au moins 11 PR de la petite sous-unité et la plupart des PR de la grande sous-unité sont modifiées post-traductionnellement^{54,55}.

Une fois synthétisées et maturées, les PR commencent à s'assembler aux ARNr de manière co-transcriptionnelle⁵⁰.

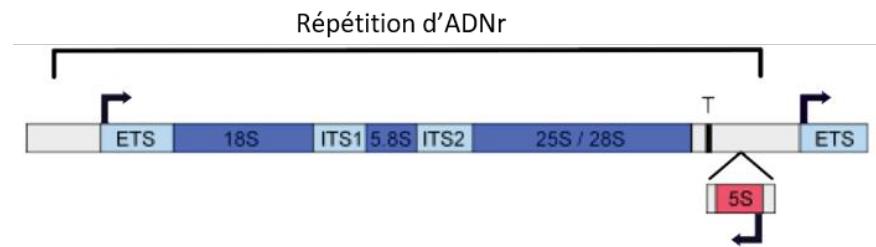


Figure 11 : Structure des répétitions d'ADNr

Les séquences des ARNr matures sont représentées en bleu foncé, et les séquences qui ne se retrouveront pas dans le ribosome (ETS et ITS) sont en bleu clair. Chez les levures, le gène ADNr 5S est dans le brin opposé à celui qui contient le pré-ARNr 35S, alors que chez les autres eucaryotes, le gène ADNr 5S est à un locus séparé. *Adapté de Schneider, 2012⁵⁶*.

b – La synthèse et la maturation des ARN ribosomiques

Les ARNr sont synthétisés au niveau d'un domaine nucléaire distinct, le nucléole, excepté l'ARNr 5S qui est synthétisé dans un domaine périnucléolaire⁵⁷. Le nucléole n'est pas seulement le siège de la transcription des ARNr mais aussi de leurs clivages, de leur maturation par des modifications chimiques et des premières étapes d'assemblage du pré-ribosome⁵⁸. Le nucléole comprend trois compartiments : le centre fibrillaire (FC), le composant fibrillaire dense (DF) et le composant granulaire (GC)⁵⁸ et est situé au niveau des séquences d'ADNr. La transcription étant intense, on peut voir les nombreuses ARN Pol I en cours de transcription sur les loci d'ADNr avec les ARNr en cours d'elongation, ce qui donne une structure particulière nommée « arbres de Noël »⁵⁹. Chez la levure, les gènes ADNr représentent 10% du génome et sont regroupés dans une seule région, avec environ 150 répétitions en tandem¹. Chez l'Homme, ils sont localisés en loci sur les chromosomes 13, 14, 15, 21 et 22 et sont répétés en tandem avec environ 400 copies au total. Les tandems d'ADNr sont appelés NORs (nucleolar organizer regions). Les gènes ADNr 5S sont également organisés en répétitions, en majorité sur le chromosome 1 chez l'homme (100 à 150 répétitions) mais d'autres copies sont aussi réparties dans tout le génome⁵⁷.

Les ARNr 18S, 28S et 5.8S sont cotranscrits par l'ARN Pol I sous la forme d'un transcrit unique polycistronique, tandis que l'ARNr 5S est synthétisé par l'ARN Pol III⁵⁷. La synthèse par la Pol I est ainsi une cible majeure de la régulation de la biosynthèse des ribosomes⁵⁶. La maturation des ARNr a lieu de façon concomitante à leur synthèse⁶⁰.

Au sein du long transcrit monocistronique ARNr 47S, les ARNr 18S, 5.8S et 28S sont séparés par des espaces internes ITS1 et ITS2 (internal transcribed spacers) et flanqués de part et d'autre par les espaces 5'-ETS et 3'-ETS (external transcribed spacers) (Figure 11). Ce transcrit primaire s'associe de façon co-transcriptionnelle avec des protéines ribosomiques, des facteurs pré-ribosomiques et des snoARNs (small nucleolar ARN) responsables du repliement et de la maturation du transcrit. Au cours de ce processus de maturation, les espaces sont séquentiellement éliminés suivant des clivages endo et exonucléolytique successifs. Avant la première coupure, les deux pré-sous-unités sont assemblées pour former un complexe 90S. Après la coupure au niveau d'ITS1, les deux pré-sous-unités sont séparées.

Le clivage du pré-ARNr commence par la coupure des extrémités du 47S pour donner un pré-ARNr 45S, puis l'enchaînement des clivages peut suivre deux voies différentes : dans la voie 1, le clivage a lieu au site A0 et 1 puis au site 2 tandis que dans la voie 2, le clivage a lieu d'abord au site 2

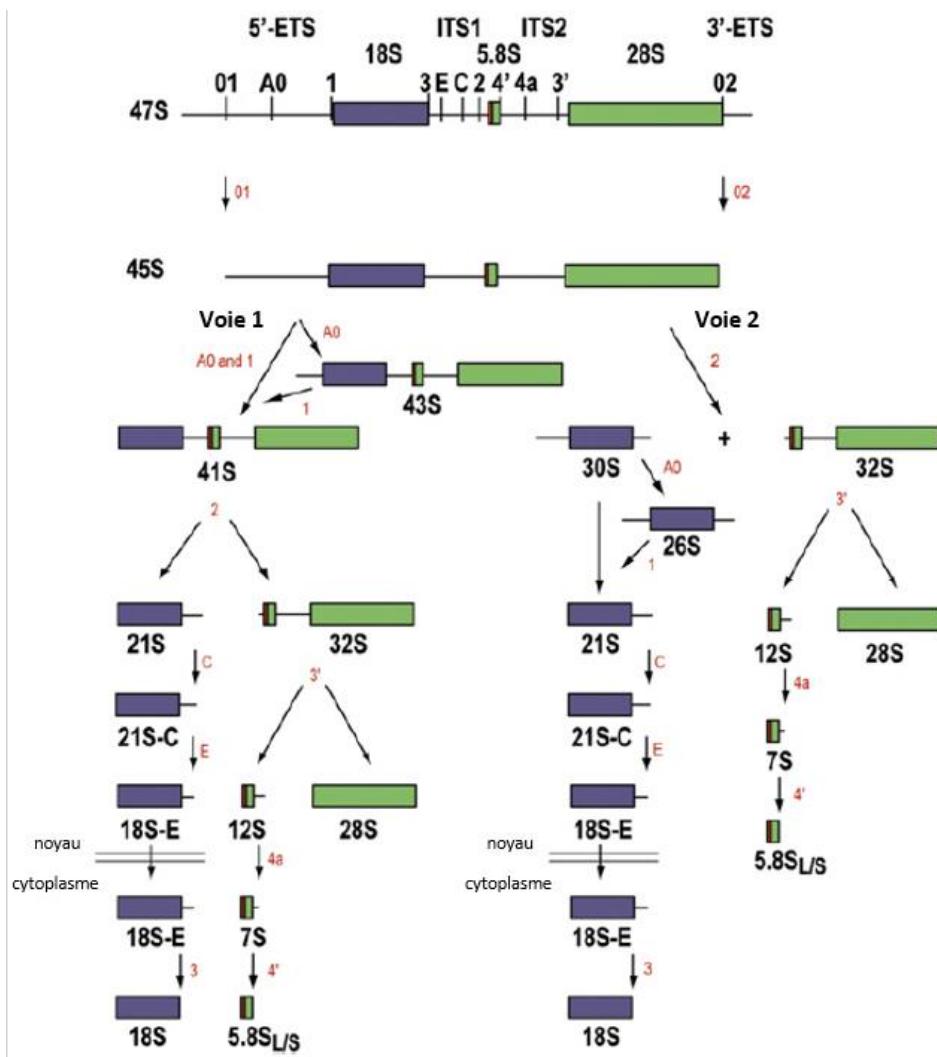


Figure 12 : Voies de clivage des pré-ARNr chez les mammifères

Le clivage du pré-ARNr commence par la coupure des extrémités du 47S pour donner un pré-ARNr 45S, puis l'enchaînement des clivages peut suivre deux voies différentes : dans la voie 1, le clivage a lieu au site A0 et 1 puis au site 2 tandis que dans la voie 2, le clivage a lieu d'abord au site 2 puis au site 1. Ces coupures génèrent un nombre important d'intermédiaires ARNr. *Adapté de Mullineux, 2012*⁶¹.

puis au site 1. Ces différentes coupures génèrent un nombre important d'intermédiaires ARNr^{61,62} (Figure 12). Leur enchaînement diffère d'un organisme à l'autre ou même d'un type cellulaire à un autre au sein d'un même organisme, révélant un niveau de modulation de la synthèse des ARNr⁶³.

Les ARNs ribosomiques subissent ensuite l'ajout de modifications chimiques. Comme vu au paragraphe I-2, les ARNr présentent douze types de modifications, notamment des ajouts de groupements méthyles, aminocarboxypropyles ou acétyles sur les bases, effectués par une enzyme, et des pseudouridylations et 2'-O-méthylations des riboses réalisées par des complexes riboucléoprotéiques snoRNP. Ces dernières sont les plus abondantes, on en retrouve plus de 100 chez la levure et plus de 200 chez l'Homme⁴⁰.

c – L'export du noyau et l'assemblage des composants du ribosome

L'export des pré-sous-unités ribosomiques a lieu alors que les sous-unités sont assemblées (ARNr + protéines) mais pas encore totalement maturées. Ainsi, cet export doit être fortement régulé, afin que des sous-unités non matures ne soient exportées dans le cytoplasme⁵³. Le mécanisme complet d'export n'est pas encore compris, mais il fait intervenir des interactions entre les sous-unités du ribosome et les complexes des pores nucléaires (NPC). Tout d'abord, les sous-unités 40S et 60S recrutent des protéines adaptatrices qui contiennent un signal d'export du noyau riche en leucine (NES). C'est la protéine Nmd3p qui est recrutée à la sous-unité 60S, la protéine adaptatrice recrutée à la sous-unité 40S n'a pas encore été déterminée⁵³. Ensuite, ces complexes sont reconnus grâce au signal NES par la protéine exportine Crm1p, qui médie leur export de façon dépendante de Ran-GTP⁶⁴.

Pour rentrer dans le noyau, les protéines ribosomiques ont besoin de transporteurs qui reconnaissent leurs signaux de localisation nucléaire (NLS)⁶⁵. Elles se lient ensuite au pré-ARN de manière co-transcriptionnelle pour former le pré-ribosome 90S. Cet assemblage est aidé par des protéines chaperonnes associées au ribosome, les complexes NAC (Nascent Polypeptide-associated Complex) et SSB-RAC (stress 70 B-ribosome-associated complex) chez la levure⁶⁶. L'assemblage des sous-unités suit une hiérarchie précise et est marqué par une stabilisation séquentielle des régions du ribosome. L'ordre est toujours le même et est conservé entre les espèces des différents règnes du vivant⁶⁷⁻⁶⁹. Ainsi le bec et la plateforme de la 40S sont structurés en dernier et l'assemblage de la sous-unité 60S suit toujours cet ordre tout d'abord la surface exposée, puis le tunnel de sortie du peptide naissant, l'interface entre

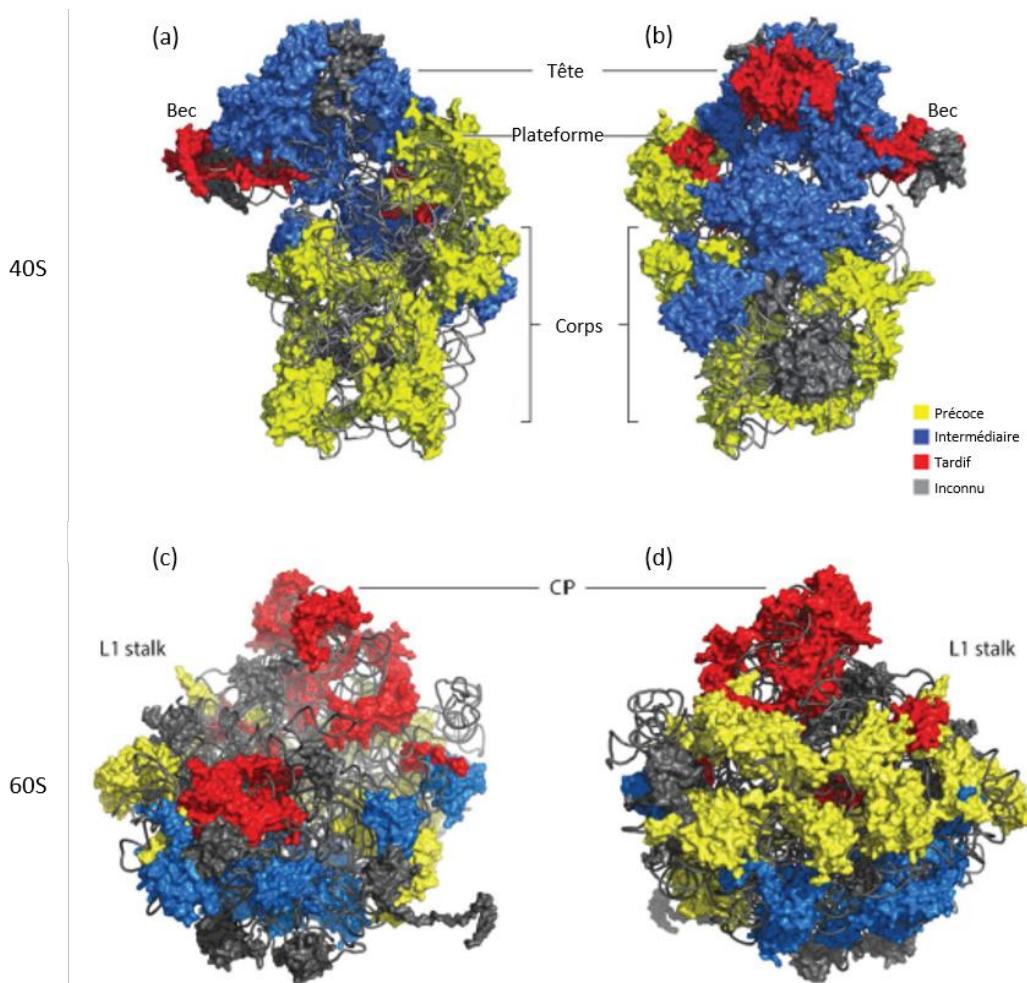


Figure 13 : Corrélation entre fonction, localisation et assemblage des protéines des sous-unités du ribosome de *S. cerevisiae*

Les protéines qui s'assemblent au ribosome précocement (en jaune), de manière intermédiaire (en bleu) et tardivement (en rouge) sont représentées sur la structure crystallographique, de la petite sous-unité 40S, vue sur l'interface (a) et sur la face exposée (b), et de la grande sous-unité 60S, vue sur l'interface (c) et sur la face exposée (d). *Adapté de de la Cruz, 2015*⁵⁰.

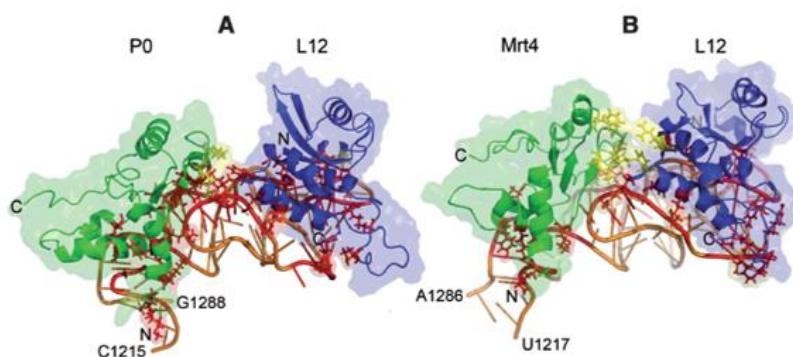


Figure 14 : Structure prédictive du complexe formée par l'ARNr 25S, la protéine uL11 et uL10/Mrt4 chez *S. cerevisiae*

Prédiction de la structure du complexe formé par l'ARNr 25S, la protéine uL11 et le domaine N-terminal de, soit la protéine ribosomique uL10 dans le ribosome mature, soit la protéine d'assemblage Mrt4p dans le ribosome en cours d'assemblage. Les résidus impliqués dans les interactions protéine-ARN et protéine-protéine sont marqués en rouge et jaune, respectivement. *Adapté de Rodriguez-Mateos, 2009*⁷⁰.

les deux sous-unités et enfin la protubérance centrale⁵⁰. Les sites actifs des deux sous-unités ne sont construits que tardivement, afin de prévenir toute initiation prématuée de la traduction (Figure 13).

Certains facteurs d'assemblage du ribosome se lient aux sites de fixation de certaines protéines ribosomiques afin d'éviter un ajout prématué de ces protéines. Par exemple, la liaison prématuée de la protéine eS10 est bloquée par Ltv1p et la liaison d'eS26 est bloquée par Pno1p/Dim2p⁷¹. Ces deux protéines sont localisées à l'entrée et à la sortie du tunnel de liaison à l'ARNm, ce qui suggère que le blocage de leur assemblage empêche un recrutement prématué de l'ARNm à la pré-sous-unité 40S. Idem pour la 60S, le recrutement des protéines uL16, eL24 et uL10 sont bloqués dans un premier temps par des protéines d'assemblage. Certaines de ces protéines d'assemblage qui bloquent la liaison de PR sont en fait des paralogues des PR qu'ils bloquent. Ces observations font émerger l'hypothèse que ces facteurs d'assemblages pourraient fonctionner comme substituants de leurs paralogues PR par interaction avec les mêmes structures ARNr dans le pré-ribosome que les PR dans le ribosome mature⁷². L'exemple le plus caractérisé est la protéine d'assemblage Mrt4p parologue de la PR uL10, dont le domaine N-terminal peut fonctionnellement remplacer le domaine de liaison à l'ARN d'uL10. Ainsi le recrutement de la protéine Mrt4p à la sous-unité 60S par interaction avec l'ARNr 25S empêche le recrutement de la PR uL10⁷⁰ (Figure 14).

d – Mutations dans les protéines ribosomiques et facteurs d'assemblage du ribosome et pathologies associées

La biogenèse des ribosomes est fortement régulée par de nombreuses voies cellulaires et un défaut dans le processus peut entraîner des pathologies. Ces pathologies, dont l'origine est la mutation dans une PR ou un facteur d'assemblage du ribosome, sont appelées des ribosomopathies⁷³. Elles partagent ainsi des défauts de biogenèse des ribosomes et de traduction, mais les phénotypes cliniques associés sont très variables et les mécanismes qui y conduisent sont fortement sujets à discussion. Les ribosomopathies présentent toutefois des caractéristiques communes, comme des effets sur les lignées cellulaires de la moelle osseuse et des tissus du squelette ainsi qu'une prédisposition au cancer.

Il a été montré que les mutations à l'origine des ribosomopathies entraînaient une activation de la protéine p53^{74,75}. Cette activation conduit alors à l'arrêt du cycle cellulaire et à l'apoptose. Mais d'autres facteurs sont aussi impliqués car la délétion de p53 ne restaure pas complètement les phénotypes⁷⁶.

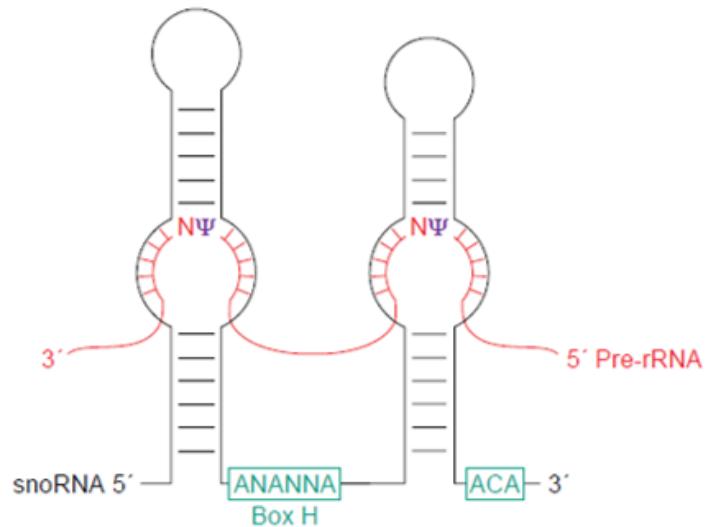


Figure 15 : SnoARN à boîte H/ACA chez les archées

Structure du complexe pré-ARNr – snoARN à boîte H/ACA. On distingue les deux structures en tige-boucle, séparées par la boîte H et suivies par la boîte ACA en 3'. Les uridines sont isomérisées dans les boucles internes des structures tiges-boucles. *Adapté de Lafontaine, 1998*⁴¹.

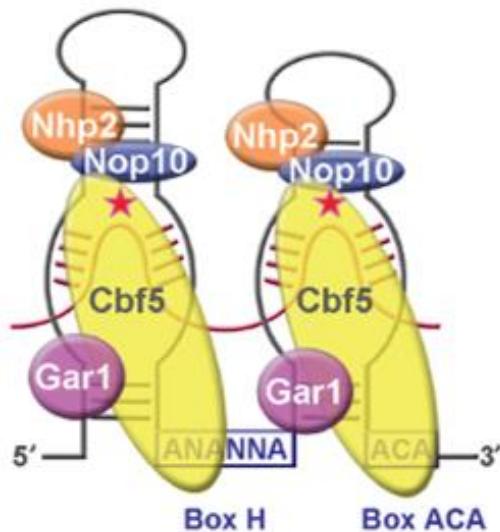


Figure 16 : Architecture du complexe snoRNP à boîte H/ACA chez les archées

Les snoRNPs à boîte H/ACA sont bipartites, composés de deux groupes des quatre protéines Cbf5p (ou dyskérine, la pseudouridyl transférase), Gar1p, Nhp2p et Nop10p, chacun associé à une tige-boucle. *Adapté de Reichow, 2007*⁷⁷.

Les mécanismes à l'origine de la variabilité des phénotypes associés aux mutations des PR suivant le tissu ou le type cellulaire ne sont pas déterminés. Plusieurs hypothèses ont été avancées⁷⁸. La première est que le seuil d'activation de p53 en réponse aux mutations des PR peut varier suivant le type cellulaire. La seconde hypothèse s'appuie sur la notion de « ribosomes spécialisés », qui repose sur le fait que des ribosomes de compositions différentes traduiraient des groupes d'ARNm différents^{48,79}. Ainsi, l'absence ou la déficience d'une protéine ribosomique affecterait spécifiquement la traduction d'un sous-groupe d'ARNm, plus ou moins spécifiques à un tissu ou un type cellulaire. Une troisième hypothèse est que certains tissus ou types cellulaires sont plus sensibles que d'autres au dysfonctionnement des ribosomes dans la traduction globale et/ou spécifique de sous-groupes d'ARNm⁸⁰.

L'étude des mécanismes qui lient les génotypes de mutation dans les PR aux variabilités phénotypiques observées dans les ribosomopathies permettrait d'améliorer la compréhension du rôle de chacune des PR dans la traduction, et inversement.

B – Les pseudouridylations et 2'-O-méthylations des ARNr et les complexes snoRNPs qui les synthétisent

97 pseudouridylations et 106 méthylations de ribose sont répertoriées dans les cellules humaines HeLa^{81,82}. Ces modifications chimiques des ARNr sont retrouvées chez tous les organismes, même si les machineries qui les catalysent sont différentes. En effet chez les bactéries, ces modifications sont assurées de façon similaire aux modifications des ARNt, c'est-à-dire par des enzymes qui reconnaissent directement la région (par le biais d'une séquence ou d'une structure) où se trouve la position à modifier. Chez les archées et les eucaryotes, le système fait intervenir des complexes ribonucléoprotéiques appelés snoRNPs, qui comportent une protéine catalytique, un snoARN guide et des protéines accessoires⁴¹.

Dans cette partie je vais m'intéresser à ces machineries ainsi qu'aux effets induits par une perte totale ou partielle de pseudouridylations et 2'-O-méthylations et par des défauts dans les machineries snoRNPs.

a – Les snoRNPs à boîte H/ACA, responsables des pseudouridylations

Les isomérisations d'uridines en pseudouridines sont catalysées par les snoRNPs H/ACA⁸³. Les snoARNs qui jouent le rôle de guides font entre 60 et 150 nucléotides. La version canonique de ces snoARNs possède un motif conservé H (5'-ANANNA-3') et un motif ACA en 3'. Ces snoARNs se replient suivant une conformation conservée qui présente deux structures tige-

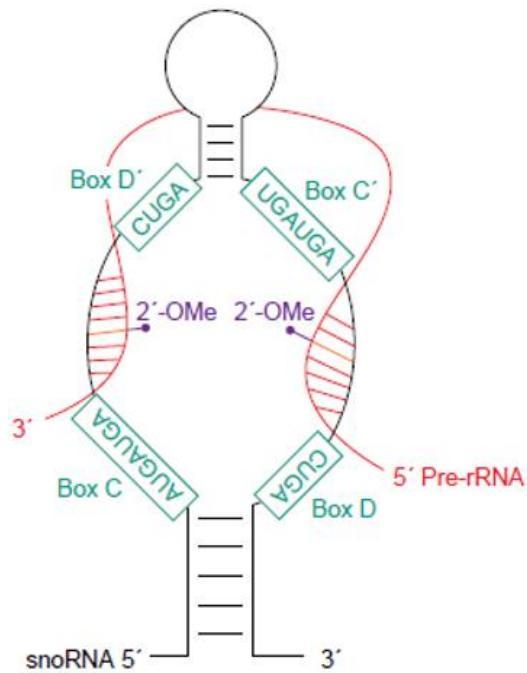


Figure 17 : SnoARN à boîte C/D chez les archées

Structure du complexe pré-ARNr – snoARN à boîte C/D. On distingue les motifs conservés C et D qui s'apparentent pour former un motif tige-boucle-tige interne appelé K-turn ainsi que les motifs C' et D' qui s'apparentent pour former un motif similaire de tige-boucle appelé K-loop. *Adapté de Lafontaine, 1998*⁴¹.

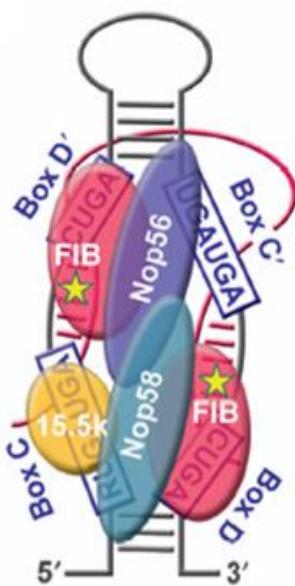


Figure 18 : Architecture du complexe snoRNP à boîte C/D chez les archées

Les snoRNPs à boîte C/D sont composés de deux parties asymétriques, chacune présentant une protéine méthyltransférase fibrillarine associée à une boîte D ou D'. *Adapté de Reichow, 2007*⁷⁷.

boucle contenant une boucle interne, séparées par une région charnière simple brin contenant le motif H. Ces deux tiges-boucles sont suivies par une queue contenant la boîte ACA positionnée trois nucléotides en amont de l'extrémité 3' de l'ARN⁸³ (Figure 15). Il existe des exceptions à cette structure canonique, par exemple, des snoARNs de type H/ACA avec une seule tige-boucle ont été découverts chez des trypanosomes et le protiste *Euglena gracilis*. Le motif ACA peut également présenter une variation de sa deuxième base, il devient alors le motif terminal « ANA »⁸⁴.

Les complexes snoRNPs H/ACA possèdent tous un groupe de quatre protéines présentes en deux exemplaires (un exemplaire pour chaque tige-boucle) : les protéines accessoires Nhp2p, Gar1p, Nop10p et la dyskérine (Cbf5p chez la levure) qui est la protéine catalytique. La dyskérine interagit avec le motif conservé ACA et la tige P1, tandis que Nhp2p se lie aux boucles des tiges-boucles dans la moitié supérieure de l'ARN. Gar1p n'interagit pas avec l'ARN et s'associe avec la dyskérine⁸⁵ (Figure 16).

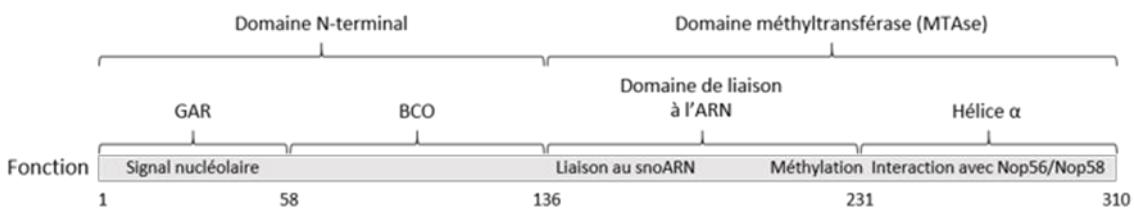
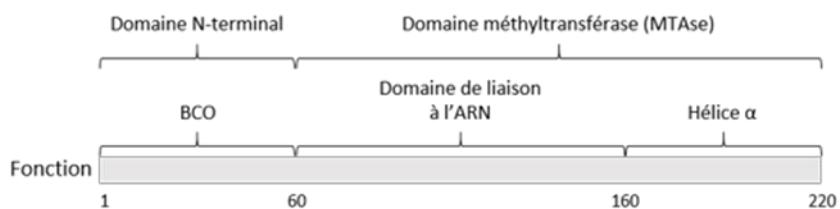
Les snoARNs H/ACA sélectionnent l'uridine à isomériser en pseudouridine en établissant, dans la plupart des cas, deux courtes interactions par complémentarité avec le snoARN, de part et d'autre de l'uridine à convertir. Ces deux régions qui interagissent avec l'ARNr sont les boucles internes des structures tiges-boucles, appelées « poches de pseudouridylation »⁸³ (Figure 16). Les interactions entre l'ARN guide et l'ARNr substrat forment une jonction à trois branches, au centre de laquelle sont exposés l'uridine cible et le nucléoside adjacent en 3'. Pour être isomérisée en pseudouridine, l'uridine cible ne doit pas être engagée dans une interaction Watson-Crick⁸⁴.

b – Les snoRNPs à boîte C/D, responsables des 2'-O-méthylation

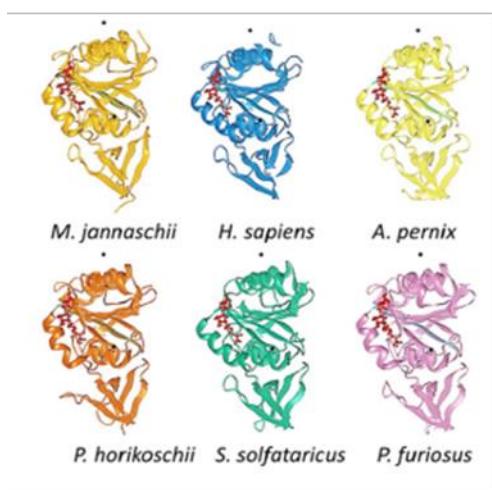
Les 2'-O-méthylation, site-spécifiques, sont catalysées par les snoRNPs à boîte C/D. Les snoARNs guides de ces complexes sont nommés « à boîte C/D » car ils possèdent deux motifs conservés C (5'-RUGAUGA-3' avec R est une purine) et D (5'-CUGA-3') respectivement positionnés proche des extrémités 5' et 3' du snoARN. Ils possèdent également deux motifs internes C' et D' dérivés des motifs C et D, mais ces motifs ne sont pas toujours présents chez les eucaryotes. Les motifs C et D s'apparent, formant un motif tige-boucle-tige interne appelé K-turn ; tandis que les motifs C' et D' s'apparent pour former un motif similaire de tige-boucle appelé K-loop⁸⁵ (Figure 17).

Les snoRNPs à boîte C/D eucaryotes possèdent quatre protéines : la protéine catalytique méthyltransférase, la fibrillarine (Nop1p chez la levure), et des protéines accessoires NOP56, NOP58 (Nop5 pour les deux chez les archées) et 15.5K (Snu13p chez la levure, L7Ae chez les archées). La fibrillarine possède un motif de liaison au SAM (S-аденосил méthionine)⁸⁴. Les motifs C/D et

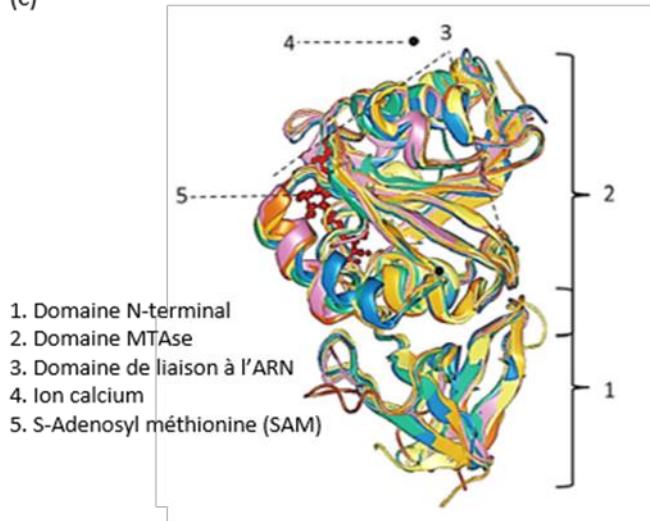
(a)

Fibrillarine chez les eucaryotesFibrillarine chez les archées

(b)



(c)

**Figure 19 : Alignement de structures de différentes fibrillarines archées et eucaryotes**

(a) Représentation de la structure primaire de la fibrillarine chez les eucaryotes et chez les archées. La séquence de la fibrillarine peut être divisée en trois ou quatre séquences : le domaine GAR, spécifique aux eucaryotes, est une séquence riche en glycine et en arginine et contient le signal nucléolaire ; le domaine BCO, dont l'activité n'est pas définie ; le domaine MTAse qui comprend à la fois l'activité enzymatique et la séquence de liaison à l'ARN et qui peut être séparé en deux domaines, une région qui se lie à l'ARN et une hélice α qui interagit avec Nop56p/Nop58p. (b) Six structures cristallographiques de cinq fibrillarines d'archées (*M. jannaschii*, *A. pernix*, *P. horikoschii*, *S. solfataricus* et *P. furiosus*) et une fibrillarine eucaryote (*H. sapiens*). (c) Alignement des six structures présentées en (b). Adapté de Rodriguez-Corona, 2015⁸⁶.

C'/D' comprennent deux centres distincts de liaison aux protéines, chacun est donc associé à un groupement des quatre protéines. Il y a donc deux fibrillarines par snoRNP, ce qui permet la méthylation de deux sites pour chaque complexe. La première protéine à se fixer est la protéine 15.5K, qui est ensuite reconnue par le complexe Nop5/fibrillarine chez les archées. Les deux protéines Nop5 interagissent entre elles par leurs domaines superhélices afin de lier les deux parties du snoRNP. Le domaine N-terminal de Nop5 interagit avec la fibrillarine, qui est positionnée avec son site actif tournée vers la cible ribose à méthyler. Chez les eucaryotes, la protéine Nop5 est remplacée par les protéines NOP56 et NOP58, qui se lient de façon préférentielle respectivement aux motifs C/D et C'/D'. L'assemblage des complexes snoRNPs se fait de façon séquentielle, avec en premier un recrutement de NOP58/fibrillarine puis NOP56/fibrillarine⁸⁵ (Figure 18).

Les snoARNs s'apparent à la région du nucléotide de l'ARNr à modifier sur environ 10 à 21 nucléotides au niveau de la séquence immédiatement en amont de la boîte D et/ou D'. Cet appariement guide l'ajout d'un groupement méthyle sur le ribose du 5^{ème} nucléotide de la région appariée en amont de la boîte D/D' (Figure 18). Il peut cependant y avoir quelques exceptions, notamment lorsque le snoARN guide la méthylation de deux nucléotides adjacents ou séparés par un nucléotide⁸⁷. De plus, un sous-ensemble (de 10 à 20 chez la levure, 17 chez l'homme) de snoARNs à boîte C/D a été identifié comme présentant des séquences conservées supplémentaires qui sont complémentaires aux régions de l'ARNr à proximité des sites de méthylation. Ces séquences stimulent la méthylation d'au moins un facteur cinq, probablement en stabilisant les interactions entre le snoRNA et l'ARNr⁸⁸.

La fibrillarine est une protéine essentielle conservée au cours de l'évolution⁸⁶. C'est la méthyltransférase responsable des 2'-O-méthylations des ARNr mais elle est également impliquée dans le clivage des ARNr et la méthylation de l'histone H2A aux loci d'ADNr⁸⁹. Chez l'homme, elle a un poids moléculaire de 35kDa. Elle fait partie de la superfamille des MTases, les méthyltransférases de S-adénosylméthionine (SAM) dont les caractéristiques sont un motif de liaison conservé au SAM, une triade ou tétrade catalytique et une structure α - β - α . La séquence de la protéine peut être divisée en deux domaines : le domaine N-terminal et le domaine MTase. Elle possède une région riche en arginine et glycine (GAR), absente chez les archées, qui interagit avec des protéines cellulaires et virales et qui possède un signal de rétention nucléolaire (Figure 19).

Le rôle de la fibrillarine dans la maturation des ARNr est connu depuis longtemps, tandis que son activité de méthylation de l'histone H2A au niveau de la Q104 (chez l'homme / Q105 chez la levure) n'a été révélée que ces dernières années⁸⁹. Cette modification entraîne un défaut de liaison de la chaperonne d'histones FACT (Facilitator of Transcription) et a lieu spécifiquement aux loci d'ADNr, ce qui entraîne une diminution de la transcription des pré-ARNr par l'ARN Pol I.

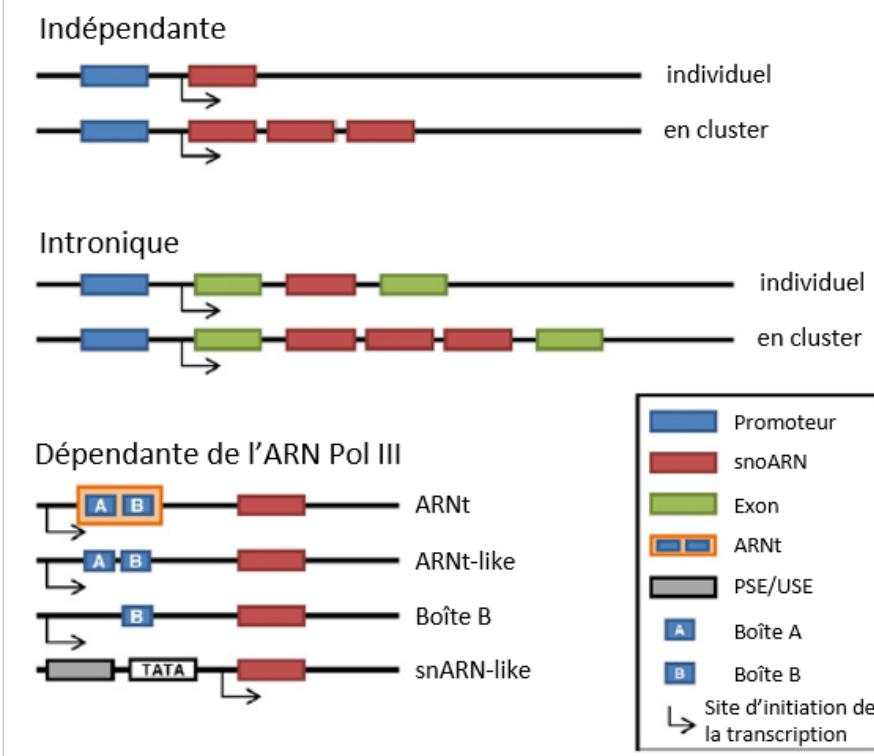


Figure 20 : Organisation génomique des séquences des snoARNs

Les snoARNs peuvent être exprimés sous leurs propres promoteurs, de façon intronique ou dépendante de l'ARN Pol III ; individuellement ou en cluster. *Adapté de Dieci, 2009⁹⁰.*

c – Expression et synthèse des snoARNs

La majorité des snoARNs est synthétisée par l'ARN Pol II mais quelques-uns sont synthétisés par l'ARN Pol III, notamment snR52 chez la levure et U3 chez les plantes. Les snoARNs peuvent être exprimés de façon indépendante sous leur propre promoteur ou ils peuvent être introniques. Ils peuvent également être sous forme monocistronique ou polycistronique. Dans certains de ces cas, ils nécessitent d'être épissés et/ou clivés. Ainsi, l'organisation des snoARNs au sein des génomes et leurs modes d'expression sont très diversifiés⁸⁵ (Figure 20).

Chez la levure, la plupart des snoARNs sont exprimés sous leurs propres promoteurs et monocistroniques, mais on retrouve aussi quelques snoARNs polycistroniques sous leurs promoteurs. Chez *S. cerevisiae*, huit snoARNs sont introniques, tous monocistroniques. Ceci est à corréler au fait qu'il n'y a que peu d'introns dans le génome de cette levure.

Chez les plantes, les snoARNs sont surtout organisés en groupes de gènes, que ce soit sous leurs propres promoteurs ou introniques. Les groupes dans les introns sont des snoARNs paralogues qui résultent de duplications. Les snoARNs peuvent aussi être transcrits par l'ARN Pol III sous forme dicistronique ARNt-snoARN.

Chez les mammifères, tous les snoARNs sont monocistroniques et la plupart sont introniques (90% chez l'Homme), mais les snoARNs essentiels et impliqués dans le clivage endonucléolytique des pré-ARNr sont exprimés sous leurs propres promoteurs (par exemple U3 (SNORD3A), U8 (SNORD118) et U13 (SNORD13))⁹⁰.

Chez toutes ces espèces, les snoARNs introniques sont retrouvés de façon enrichie dans les gènes codant des protéines ribosomiques et des protéines impliquées dans la traduction. Ainsi, chez *S. cerevisiae*, alors que seulement 3% des gènes codent des protéines ribosomiques, 37,5% des snoARNs introniques sont dans des gènes de protéines ribosomiques et 87,5% sont dans des gènes de protéines liées à la traduction. Chez l'Homme, seulement 0,4% des gènes codent des protéines ribosomiques, tandis que 22,2% des snoARNs sont dans des gènes de protéines ribosomiques et 26,3% dans des gènes de protéines impliquées dans la traduction. Ceci suggère une co-régulation de l'expression des snoARNs avec d'autres composants de la biogénèse des ribosomes⁹¹.

d – Effets d'une perte totale ou partielle des pseudouridylations et 2'-O-méthylation

Les pseudouridylations et 2'-O-méthylation sont conservées et concentrées dans les régions fonctionnellement importantes, ce qui suggère un rôle important dans le fonctionnement du

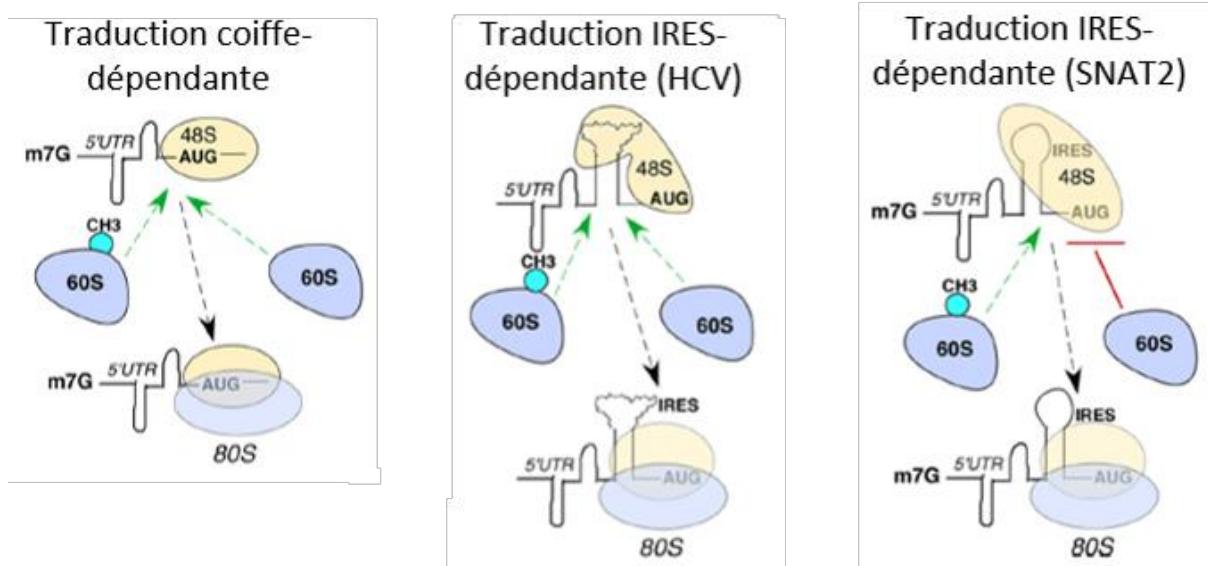


Figure 21 : L'absence de méthylation des ARNr empêche le recrutement de la 60S au niveau de l'IRES cellulaire SNAT2

Le recrutement du complexe 48S n'est pas empêché par l'absence de méthylation des ARNr, que ce soit par un mécanisme coiffe-dépendant ou IRES-dépendant (viral HCV ou cellulaire SNAT2). En revanche, le recrutement de la 60S non méthylée au niveau de la 48S est déficient dans le cas de l'IRES cellulaire SNAT2 mais pas de la coiffe ni de l'IRES viral HCV. *Adapté de Basu, 2011*⁹².

ribosome. Dans cette partie je vais m'intéresser aux effets d'une perte totale ou partielle de ces modifications sur la traduction, qui permettent de mieux comprendre la(es) fonction(s) de ces modifications.

❖ Perte globale des pseudouridylations ou des 2'-O-méthylations

La perte globale des pseudouridylations a été obtenue chez la levure en générant la mutation L94A dans le site catalytique de Cbf5p⁹³. Cette mutation abolit presque complètement les pseudouridylations à 37°C, mais elle n'affecte pas la stabilité des snoRNPs H/ACA ni la synthèse des ARNr 18S et 25S. L'absence quasi totale de pseudouridylation provoque une inhibition forte de la croissance. De manière surprenante, les ribosomes sont produits et assemblés correctement. Le fort défaut de croissance observé semble donc être dû à la perte des pseudouridines, très certainement à travers des défauts de traduction par les ribosomes affectés. Chez l'Homme, le gène *DKC1* qui code la dyskérine, peut présenter une mutation ponctuelle à l'origine de la dyskératose congénitale (DC), une maladie liée à l'X. La dyskérine présente une activité pseudouridine synthase mais elle est également un composant de l'holoenzyme télomérase, responsable du maintien des télomères⁹⁴. Il a été suggéré que l'apparition de la maladie est due aux défauts de synthèse et de fonction du ribosome et que les défauts de maintenance des télomères contribuent à la sévérité de la maladie^{95,96}. L'étude de patients atteints de la DC et d'un modèle de souris qui porte une mutation de la dyskérine montre que les ribosomes présentent un faible taux de pseudouridines associé à des défauts de traduction IRES-dépendante (les IRES sont des sites de recrutement interne du ribosome)⁹⁷, dus à une baisse de l'affinité de la sous-unité 40S pour les structures IRES⁹⁸. Des défauts de fidélité de l'elongation et de la terminaison de la traduction ont également été mis en évidence (décalages de cadre de lecture +1 et -1 et translecture du codon stop)⁹⁸.

Des conclusions similaires peuvent être apportées pour une perte globale des 2'-O-méthylations chez la levure. Un mutant thermo-sensible de Nop1p (Nop1-3) présente une absence de 2'-O-méthylation ainsi qu'un léger défaut de clivage des ARNr 18S et 25S. Il présente également des défauts de croissance, qui peuvent être dus au clivage déficient des ARNr ou à l'absence totale de 2'-O-méthylations. Cependant, le défaut de clivage est faible et n'affecte probablement que peu la biogénèse des ribosomes. Ainsi, les 2'-O-méthylations ne sont pas nécessaires pour la production des ribosomes chez la levure, mais sont nécessaires pour assurer une traduction et donc une croissance cellulaire⁹⁹. Contrairement à la levure, dans les cellules de mammifères, la méthylation des riboses est nécessaire pour le clivage des pré-ARNr¹⁰⁰. Dans des cellules

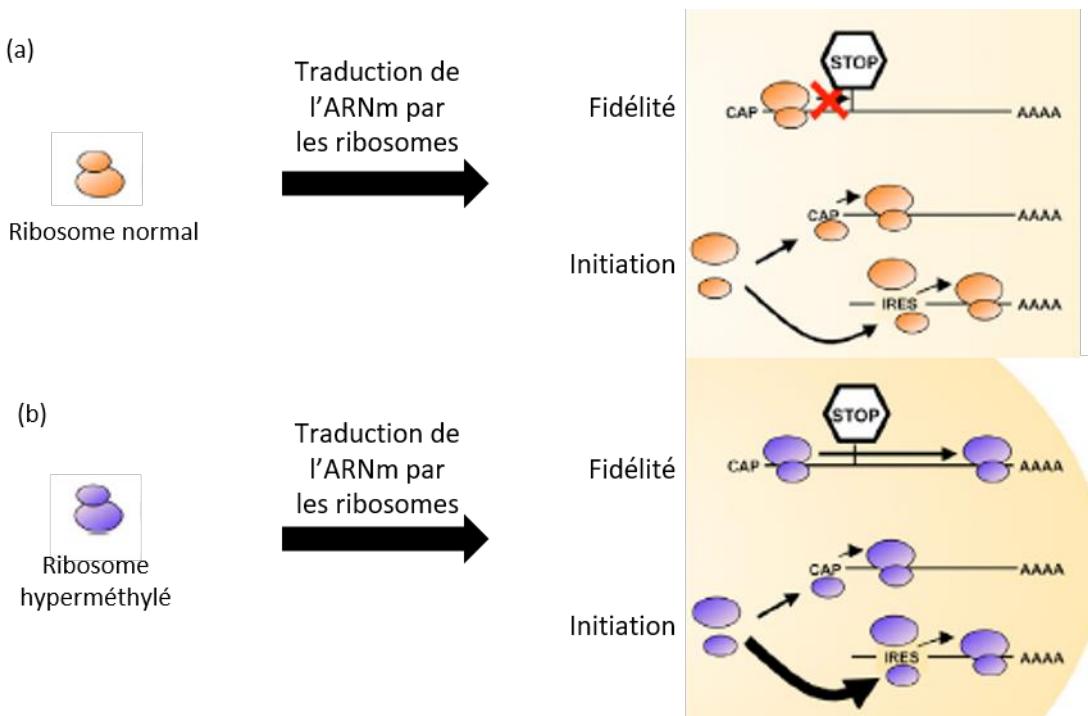


Figure 22 : Effets sur la traduction d'une hyperméthylation des ARNs ribosomiques

(a) lorsque la fibrillarine est présente en quantité normale dans la cellule, les ARNs ribosomiques sont méthylés normalement, la fidélité de la traduction est élevée et la balance entre l'initiation coiffé-dépendante et l'initiation IRES-dépendante est correcte. (b) si le taux de fibrillarine est augmenté, la méthylation globale des ARNs ribosomiques augmente. Cela entraîne une diminution de la fidélité de la traduction et un déséquilibre du mécanisme d'initiation en faveur de l'initiation IRES-dépendante. *Adapté de Marcel, 2013¹⁰¹.*

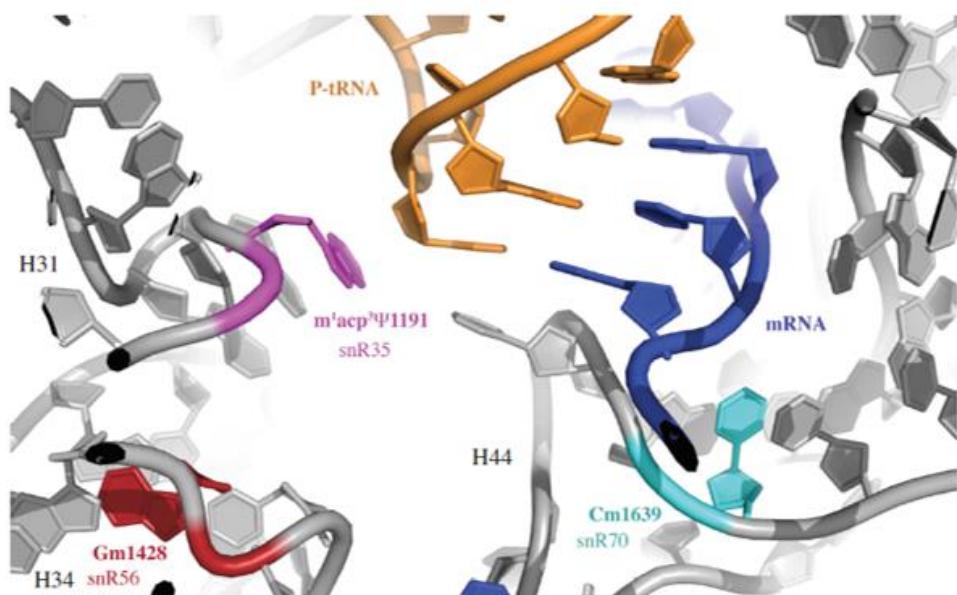


Figure 23 : Environnement proche du nucléotide hypermodifié $m^1apc^3\Psi1191$ au sein du site P

Représentation tridimensionnelle des bases de l'ARNr cartographiées au site P de *E. coli*. Le nucléotide hypermodifié $m^1apc^3\Psi1191$ est représenté en violet, le nucléotide Gm1428 en rouge, le nucléotide Cm1639 en bleu clair, l'ARNm en bleu foncé et l'ARNt au site P en orange. *Adapté de Baudin-Baillieu, 2009¹⁰².*

humaines, une déplétion de L13a, protéine ribosomique qui interagit avec la fibrillarine et régule probablement ainsi son activité méthyltransférase, a été effectuée et entraîne une réduction globale de la méthylation des ARNr. Cette réduction globale induit une diminution de l'initiation au niveau d'IRES cellulaires mais pas au niveau d'IRES viraux, due à un défaut de recrutement de la sous-unité 60S⁹² (Figure 21).

Dans le cas inverse, l'augmentation du taux de fibrillarine conduit à une augmentation du taux de méthylation de l'ARNr et est associée à une augmentation du taux de translecture au codon stop et à une augmentation de l'initiation IRES-dépendante de gènes cellulaires impliqués dans la tumorigénèse (*c-MYC*, *FGF1*, *FGF2*, *VEGFA* et *IGF-1R*)¹⁰¹ (Figure 22).

L'inhibition par des morpholinos antisens des snoARNs U26, U44 et U78 responsables de 2'-O-méthylation chez le poisson zèbre conduit à des défauts de développement et à une létalité embryonnaire. Ces défauts sont également observables lorsqu'un seul de ces snoARN est inhibé. Cet exemple illustre le fait que la délétion d'une seule 2'-O-méthylation peut conduire à un phénotype fort de létalité¹⁰³.

❖ Perte de modifications autour des sites de liaison des ARNt A, P et E

Autour des sites de liaison des ARNt au sein du ribosome, dans l'ARNr 18S, on retrouve huit nucléotides modifiés : quatre autour du site A (deux sont proches de l'interaction codon-anticodon et les deux autres sont localisées au-dessus de l'ARNt au site A), et les quatre autres à proximité des sites P et E. La perte d'une seule modification ou d'un nombre limité de modifications regroupées au sein d'une région fonctionnellement importante n'est généralement pas suffisante pour induire des défauts du taux de synthèse protéique et de fidélité de la traduction en conditions de laboratoire. Il existe cependant quelques exceptions, dont la délétion du snoRNA H/ACA snR35, qui guide la pseudouridylation de l'uridine hypermodifiée 1191 ($m^1apc^3\Psi$ 1191) au niveau du site P (Figure 23). La perte de ce snoRNA entraîne la perte de la pseudouridylation et de la méthylation de la position U1191, et conduit à une réduction de la croissance de 10% à 30°C ainsi qu'une baisse de 7% de l'incorporation de méthionine S³⁵ *in vivo*¹⁰⁴ et une augmentation du décalage de cadre de lecture +1 (le ribosome se décale de +1 nucléotide par rapport au cadre de lecture initial)¹⁰². La perte de snR35 induit des défauts significatifs dans la production de la sous-unité 40S, corrélés avec des défauts de maturation du précurseurs 20S de l'ARNr mature 18S¹⁰⁴, ce qui indique que la production de la sous-unité 40S nécessite la synthèse de la Ψ 1191 et/ou l'interaction de snR35 durant la maturation du pré-ARNr. On ne sait donc pas si les défauts observés sont dus à l'absence de la Ψ 1191, aux défauts de biogénèse des ribosomes dus à l'absence de la modification et/ou du snoARN ou aux deux.

D'autres délétions individuelles de modifications au niveau des sites A, P et E n'affectent pas la croissance cellulaire mais ont un effet sur certains paramètres de la traduction. Par exemple, la délétion du snoARN responsable de la méthylation de C1639, localisé au site P comme la ψ 1191, induit une réduction de 5% du taux d'incorporation de la méthionine S³⁵ in vivo, et des défauts de fidélité de l'elongation et de la terminaison de la traduction (augmentation des décalages de cadre de lecture +1 et -1 et augmentation de la translecture du codon stop, qui correspond à l'incorporation d'un acide aminé au codon stop)^{102,104}.

La délétion de deux ou trois modifications au niveau des sites de liaison des ARNt induit des défauts de croissance plus ou moins forts suivant les modifications touchées ainsi que des défauts de fidélité de la traduction. L'effet de la perte de pseudouridylations et 2'-O-méthylations est donc cumulatif et combinatoire⁸⁴.

❖ Perte de modifications dans le centre peptidyl-transférase

La mutation de snR10, responsable de la ψ 2923, localisée dans le PTC, induit une perte de cette modification chimique ainsi qu'une baisse du taux d'incorporation de la méthionine S³⁵. Bien que la production des deux sous-unités 40S et 60S soit correcte dans la souche mutante, le profil de polysome associé présente des épaulements appelés « halfmers », ce qui suggère un défaut de recrutement de la 60S au niveau du complexe de préinitiation 43S. Ceci explique la réduction du taux de synthèse protéique observée¹⁰⁵.

Six pseudouridines et six 2'-O-méthylations sont présentes dans l'ARNr 25S à proximité du centre peptidyl-transférase (PTC) mais aucune n'est assez proche pour participer directement aux réactions qui y ont lieu¹⁰⁵. La délétion des cinq snoARNs responsables des six pseudouridylations dans le PTC induit des défauts de croissance et une baisse du taux d'incorporation de la méthionine S³⁵, ce qui suggère que l'accumulation de la perte de plusieurs modifications affecte probablement la structure du PTC en le rendant moins efficace pour catalyser les liaisons peptidiques¹⁰⁵.

❖ Perte de modifications à l'interface entre les deux sous-unités

Plusieurs ponts ARN sont établis entre les deux sous-unités du ribosome en cours de traduction. Parmi ces ponts, le pont B2a est formé par une interaction entre l'hélice H69 de l'ARNr 25S, positionnée à proximité des sites A et P, et l'hélice h44 de l'ARNr 18S. L'hélice H69 contient cinq nucléotides modifiés : quatre pseudouridylations et une 2'-O-méthylation, guidées par quatre snoARNs. La délétion d'au moins trois de ces snoARNs conduit à des défauts de croissance et

d'incorporation de la méthionine S³⁵. Ces mutants montrent également une augmentation de la sensibilité à des antibiotiques ciblant le ribosome, comme la néomycine et la sparsomycine, qui visent respectivement le centre de décodage et le PTC. Ces données ainsi que des études structurales et le fait que les pseudouridines permettent de façon générale une stabilisation locale montrent que la perte des modifications correspondantes aux snoARNs délétés affecte la structure du ribosome¹⁰⁶. Suivant les combinaisons de trois snoARNs de l'hélice H69 délétés, on peut aussi observer des variations de la translecture et des décalages de cadre de lecture à la hausse ou à la baisse, de façon dépendante de la combinaison délétée⁸⁴.

❖ Hypothèses sur le rôle de ces modifications

Le fait que la délétion de la majorité des snoARNs seuls n'ait pas d'impact phénotypique est intrigant du fait que ces modifications ont pour la plupart été conservées de façon universelle au cours de l'évolution, ce qui suggère qu'elles pourraient conférer un avantage sélectif. Une hypothèse est que les conditions de culture testées en laboratoire sont trop différentes et moins complexes que les conditions naturelles, qui sont alors constamment soumises à des stress comme des changements fréquents et importants de la température, la dessiccation, les rayons UV, le stress oxydatif, la déprivation de nutriments ainsi que la compétition avec d'autres organismes. Dans ces cas, il est probable que l'absence d'une seule de ces modifications induise des défauts subtils dans la traduction qui sont suffisants pour entraîner un défaut de croissance, ce qui a conduit à une contre-sélection de ces mutants au cours de l'évolution⁸⁴. Cette hypothèse est soutenue par les résultats d'une expérience de compétition entre une souche sauvage et une souche dans laquelle le snoARN snR191 est délété. Ce snoARN est responsable des pseudouridylations des positions U2258 et U2260 localisées dans la boucle de l'hélice 69 de l'ARNr 25S qui forme un pont avec l'hélice 44 de l'ARNr 18S, formant une interaction entre les deux sous-unités. La délétion de snR191 n'induit pas de défauts de croissance à toutes les températures testées mais en compétition avec une souche sauvage, la souche mutante est fortement désavantagée¹⁰⁷.

Ces données suggèrent que chaque modification contribue modestement à la fonction du ribosome et permet plutôt une régulation fine de la traduction. Différentes études ont été réalisées sur des petits groupes de modifications par délétions des snoARNs correspondant. L'impact sur différents paramètres de la traduction a été testé, comme la fidélité de l'elongation par des mesures du taux de décalage du cadre de lecture, la fidélité de la terminaison par des mesures du taux de translecture du codon stop, le défaut d'assemblage des sous-unités et de la formation de polysomes.

e – Défaux des machineries snoRNPs et pathologies

Des variations de l'expression et de l'activité de différents composants des machineries snoRNPs ont été observées dans différentes pathologies, dont je vais donner quelques exemples dans cette partie.

La dérégulation de l'expression de certains snoARNs est associée à différents cancers¹⁰⁸. Ceci peut être mis en parallèle avec le fait que ces snoARNs, en grande majorité introniques, peuvent être associés à des protéines ribosomiques ou nucléolaires. Or on sait que dans les cancers, la biogénèse des ribosomes est dérégulée, via l'altération de nombreuses voies de signalisation¹⁰⁹.

Des études ont montré une surexpression globale des snoARNs à boîte C/D dans des cancers du sein chez la souris et chez l'homme, ainsi que dans des cancers de la prostate¹¹⁰. La perte ou la diminution de l'expression du snoARN à boîte C/D U50 a été observée dans des cancers colorectaux, du sein et de la prostate^{108,111,112}. La réexpression de U50 dans des cellules du cancer du sein et de la prostate entraîne une diminution de la formation de colonies^{111,112}, ainsi l'expression de U50 a un effet négatif sur la prolifération cellulaire. Dans des lymphocytes dont la prolifération est stimulée, le taux de U50 est plus faible que dans des lymphocytes dont la prolifération n'est pas stimulée, et s'accompagne d'une baisse du taux de 2'-O-méthylation du site C2848 qu'il cible. On observe en parallèle une légère altération de l'initiation IRES-dépendante¹¹³. Ces études ont montré l'implication de la dérégulation du snoARN U50 dans la tumorigénèse.

Les dérégulations de l'expression ou de l'activité des protéines des complexes snoRNPs peuvent également être associées à des pathologies. Comme il a déjà été mentionné précédemment, une mutation dans le gène codant la pseudouridyl-transférase dyskérine est à l'origine de la dyskératose congénitale, une pathologie dont les patients présentent une augmentation de la susceptibilité de développer des cancers, probablement lié à la baisse de l'initiation IRES-dépendante induite par la baisse des pseudouridylation des ARNr^{95,97}.

L'expression de la méthyltransférase fibrillarine est augmentée dans différents types de cancers comme des cancers du sein et de la prostate, ce qui peut être corrélé au fait que l'initiation IRES-dépendante de gènes impliqués dans la tumorigénèse augmente quand le taux de fibrillarine augmente^{101,114,115}. De plus, la fibrillarine interagirait avec les protéines virales du virus de la grippe Influenza A et la protéine Tat (transactivator of transcription) du virus HIV

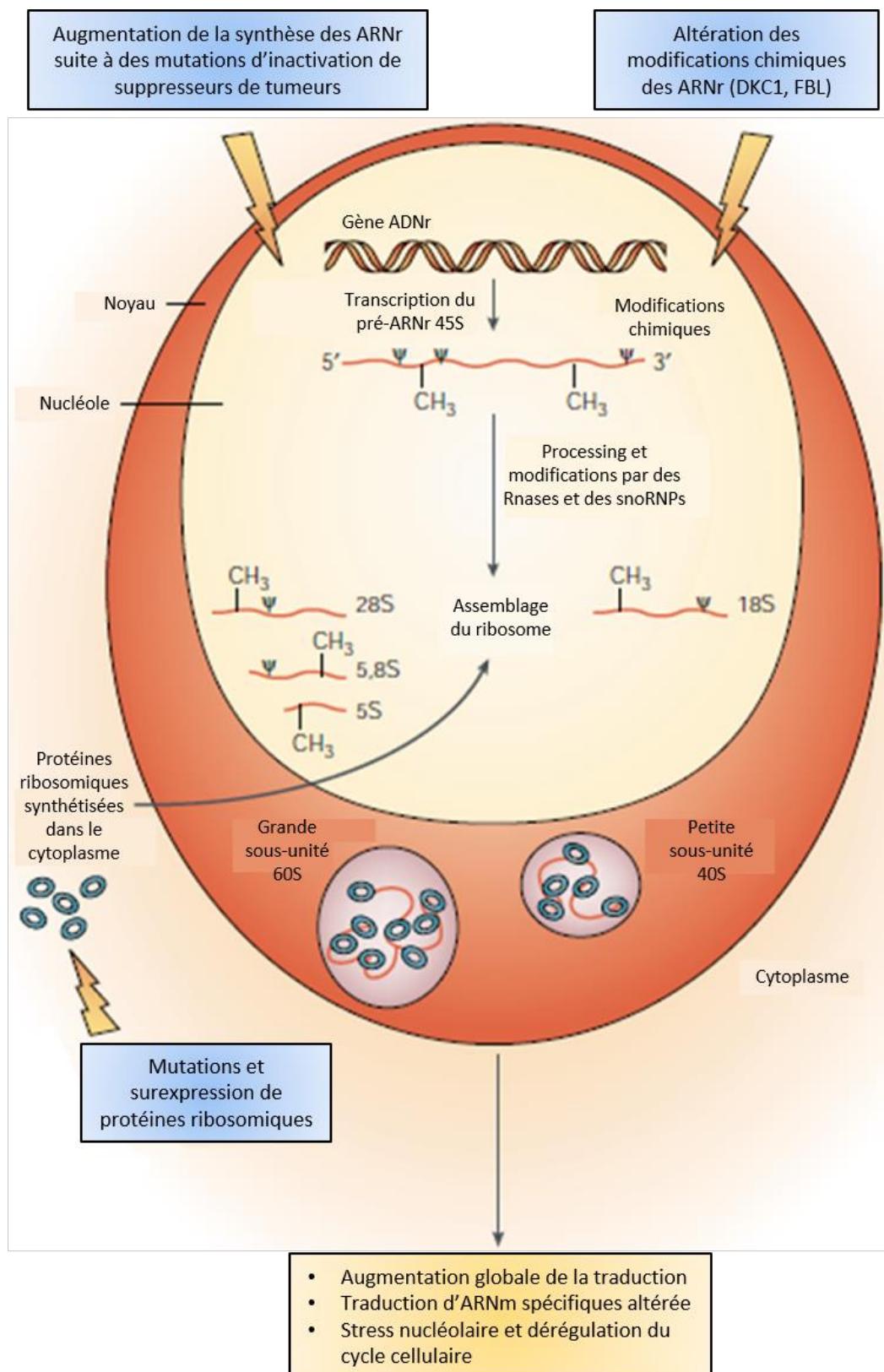


Figure 24 : Conséquences d'une dérégulation de la biogénèse des ribosomes

La dérégulation de la biogénèse des ribosomes peut survenir à différentes étapes et induire une altération globale et spécifique de la traduction. Ces altérations peuvent être à l'origine d'un état de stress cellulaire et une dérégulation de la progression dans le cycle cellulaire, ce qui peut conduire à des cancers. *Adapté de Ruggero, 2003*¹⁰⁹.

et la dérégulation de son expression pourrait jouer un rôle dans l'infection virale par ces virus^{116,117}.

Ces exemples illustrent des dérégulations de l'expression et de l'activité de différents composants des machineries snoRNPs qui sont impliquées dans différentes pathologies, majoritairement des cancers, probablement lié en partie à des variations des modifications des ARNr.

A plus large échelle, toute la biogenèse des ribosomes est un processus hautement régulé qui, en cas de dysfonctionnement, peut être à l'origine de dérégulation de la traduction et de la progression dans le cycle cellulaire, ce qui peut conduire au développement d'un cancer (Figure 24).

Dans ce chapitre, je viens de décrire le principal acteur de la traduction, le ribosome. Je vais à présent détailler les étapes de la traduction et comment le ribosome y participe activement.

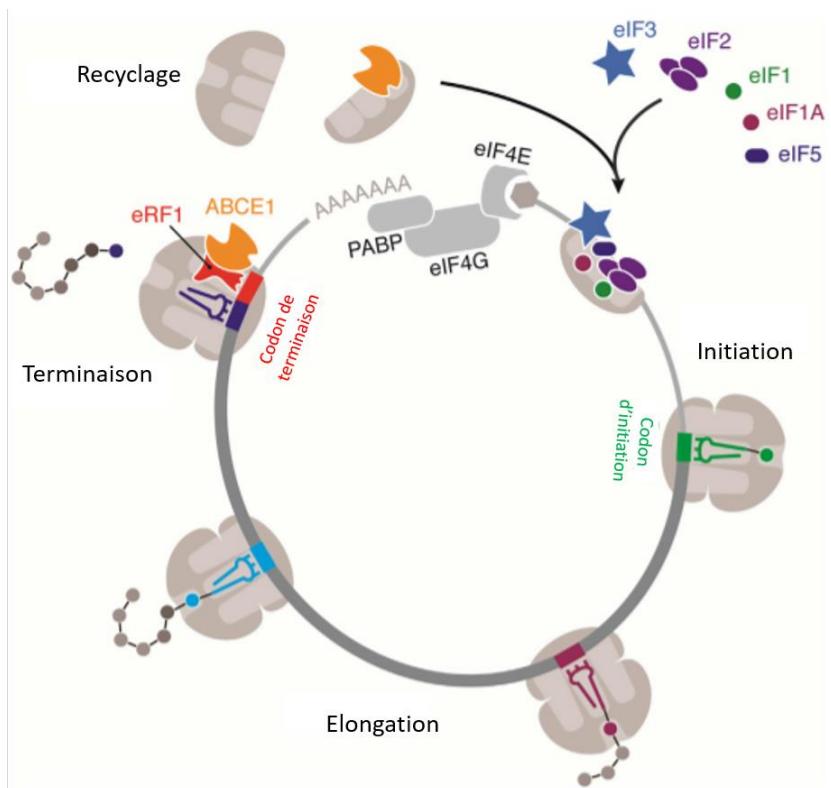


Figure 25 : Etapes de la traduction chez les Eucaryotes

La traduction est découpée en quatre étapes. (1) L'initiation, qui a lieu au niveau d'un codon d'initiation. (2) L'elongation, durant laquelle la chaîne peptidique naissante est synthétisée. (3) La terminaison, qui a lieu au niveau d'un codon de terminaison et au cours de laquelle la protéine néo-synthétisée est libérée. (4) Le recyclage des différents composants de la machinerie de traduction. *Adapté de Green, 2017¹¹⁸.*

II – La traduction chez les Eucaryotes

L'expression des gènes correspond à la lecture de l'information génétique présente sous forme d'ADN (AcideDésoxyriboNucléique) pour donner lieu à la synthèse de protéines, molécules aux rôles multiples dans les cellules. Elle fait intervenir plusieurs étapes : tout d'abord, l'ADN est transcrit en ARNs (AcideRiboNucléique) simple brin, qui peuvent être des ARNs codants ou non-codants. Ces ARNs subissent ensuite une maturation. Les ARNs codants, pré-messagers, deviennent des messagers matures qui sont ensuite exportés du noyau vers le cytoplasme pour être traduits.

Ces ARNm vont alors être pris en charge par le ribosome pour être traduits en protéines. Plusieurs ribosomes peuvent être en cours de traduction sur le même ribosome et forment ce que l'on appelle des polysomes.

La traduction peut être découpée en quatre étapes (Figure 25):

- L'initiation, qui correspond au scanning de la petite sous-unité 40S jusqu'au signal d'initiation de la traduction. La grande sous-unité 60S s'associe alors à la 40S pour former le ribosome complet 80S.
- L'elongation, au cours de laquelle la chaîne polypeptidique est synthétisée.
- La terminaison, qui a lieu au niveau d'un codon STOP et qui marque la fin de la synthèse protéique.
- Le recyclage des composants de la machinerie de la traduction et la libération de la protéine néo-synthétisée.

Au cours de ce chapitre je vais m'attacher à détailler ces différentes étapes.

1 – L'initiation de la traduction

L'initiation de la traduction fait intervenir de nombreux facteurs et correspond fréquemment à l'étape limitante de la traduction. Dans cette partie, je vais m'attacher à décrire l'initiation canonique de la traduction mais également plusieurs mécanismes particuliers, comme l'initiation IRES-dépendante.

A – L'initiation cap-dépendante

L'initiation de la traduction débute par la formation du complexe de pré-initiation (PIC) 43S qui est assemblé à partir du complexe ternaire eIF2-GTP-ARNt^{Met}, plusieurs facteurs d'initiation (notés eIFs) comme eIF1, 1A, 3 et 5, et la petite sous-unité 40S¹¹⁹ (Figure 26).

L'ARNt^{Met} initiateur n'est pas identique à l'ARNt^{Met} qui intervient durant l'elongation. En effet, il présente quelques variations nucléotidiques, structurales et de modifications chimiques qui lui permettent d'intégrer le complexe ternaire et de reconnaître le site d'initiation de façon très

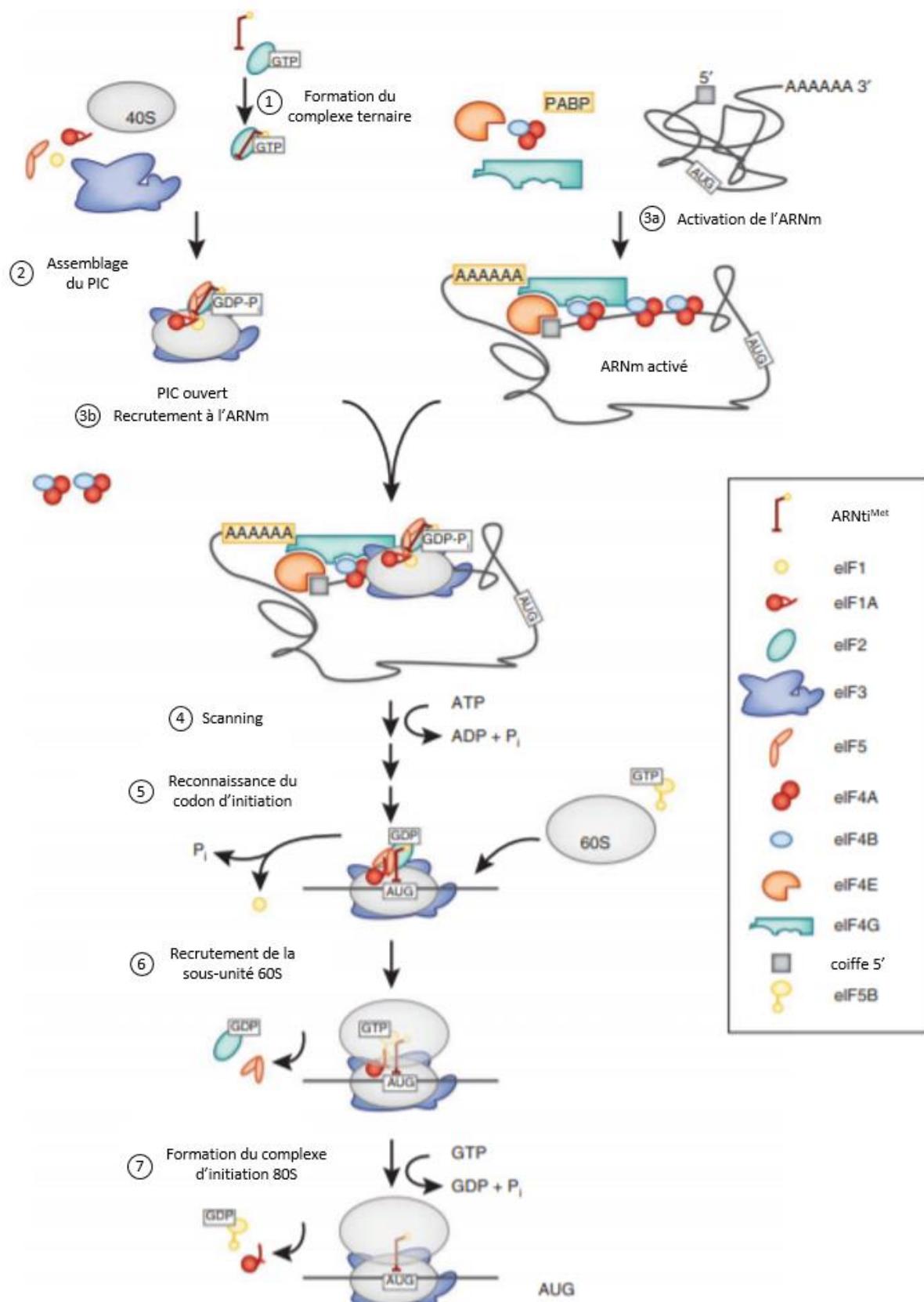


Figure 26 : Initiation canonique de la traduction

L'initiation commence avec la formation du complexe ternaire, composé du facteur eIF2 couplé au GTP et de l'ARNt initiateur (1). Le complexe ternaire est recruté à la sous-unité 40S avec l'aide des

fidèle¹²⁰. Le complexe ternaire se forme suite à la conversion du GDP associé au facteur d'initiation eIF2 en GTP. L'ARNt^{Met} initiateur est ensuite recruté. La formation du PIC 43S se fait de façon séquentielle : tout d'abord, les facteurs d'initiation eIF1, 1A et 3 se lient à la sous-unité 40S, suivi du complexe ternaire (eIF2-GTP-ARNt^{Met}) et enfin du facteur eIF5¹²⁰. Les facteurs eIF1 et eIF1A sont positionnés de telle sorte qu'ils stabilisent une conformation ouverte du PIC, qui est ainsi capable de s'assembler à l'ARNm et de le scanner jusqu'à trouver le codon initiateur¹²¹ (Figure 26).

L'activation de l'ARNm consiste en la liaison des protéines eIF4E, une protéine de liaison à la coiffe ; eIF4A, une hélicase qui déroule les structures secondaires à proximité de la coiffe ; et eIF4G, une protéine qui interagit avec les PAPBP (poly(A) Binding Protein) fixées à la queue poly(A), permettant ainsi de former une boucle entre la coiffe et la queue et de circulariser l'ARNm (Figure 26). eIF4G interagit également avec eIF4B, une protéine qui favorise l'activité d'eIF4A^{119,122}. L'activation des ARNm joue un rôle dans le recrutement du PIC 43S et dans le couplage des étapes de terminaison, de recyclage et de réinitiation du même ARNm¹²¹. La circularisation favorise la traduction mais n'est pas essentielle¹²³.

Dans les cellules de mammifères, le PIC 43S est recruté à l'ARNm par des interactions entre eIF3, qui fait partie du PIC, et eIF4G, lié à l'ARNm. Chez la levure, il semblerait que ce recrutement fasse intervenir un autre domaine d'eIF4G, qui interagit non pas avec eIF3 mais avec eIF5. Quoi qu'il en soit, le PIC 43S est recruté sur l'ARNm grâce à des interactions entre les différents facteurs d'initiation qui composent le PIC et ceux qui sont fixés à l'ARNm¹²⁰ (Figure 26).

Une fois lié à l'extrémité 5' de l'ARNm, le PIC 43S, avec l'ARNt^{Met} au site P, se déplace le long de l'ARNm en direction du 3' jusqu'à se trouver au niveau du codon de démarrage de la traduction AUG : c'est l'étape de scanning. La reconnaissance de ce codon se fait en grande partie par complémentarité avec l'anticodon de l'ARNt^{Met}¹²⁴. Quand le PIC rencontre un codon d'initiation, l'interaction codon-anticodon entraîne des mouvements au sein du PIC. L'ARNt^{Met}, qui n'occupait pas totalement le site P, se retrouve alors à occuper tout le site¹¹⁹ (Figure 27). Ces changements conformationnels conduisent à l'hydrolyse du GTP fixé à eIF2. Le PIC adopte alors une conformation fermée qui arrête le processus de scanning¹²⁵. Les facteurs eIF1, eIF2 lié au GDP et eIF5 sont dissociés du PIC. eIF2B recycle eIF2-GDP en eIF2-GTP pour un prochain cycle de traduction. Les facteurs eIF5B (une GTPase) et eIF5A favorisent l'association de la grande sous-unité 60S pour former le ribosome complet 80S. Après le départ des derniers facteurs d'initiation eIF3, eIF1A et eIF5B par hydrolyse du GTP associé à eIF5B, le ribosome est prêt pour la phase d'elongation de la traduction¹¹⁹ (Figure 26).

facteurs eIF1, 1A, 3 et 5 pour former le complexe de pré-initiation (PIC) 43S (2). En parallèle, le facteur eIF4 et les PABP se lient à l'ARNm au niveau de la coiffe et de la queue poly(A) respectivement, pour former l'ARNm activé sous forme circulaire (3a). Le PIC est ensuite recruté à l'ARNm activé (3b) où il commence à scanner (4) l'ARNm vers l'extrémité 3' jusqu'à reconnaître le codon d'initiation AUG (5). La reconnaissance du codon d'initiation entraîne la libération de eIF1 et la conversion du GTP lié à eIF2 en GDP. Le processus d'initiation s'arrête. Les facteurs eIF2-GDP et eIF5 se dissocient, permettant l'arrivée de eIF5B qui aide au recrutement de la sous-unité 60S (6) pour former le complexe d'initiation 80S (7) après hydrolyse du GTP. *Adapté de Aitken, 2012*¹²¹.

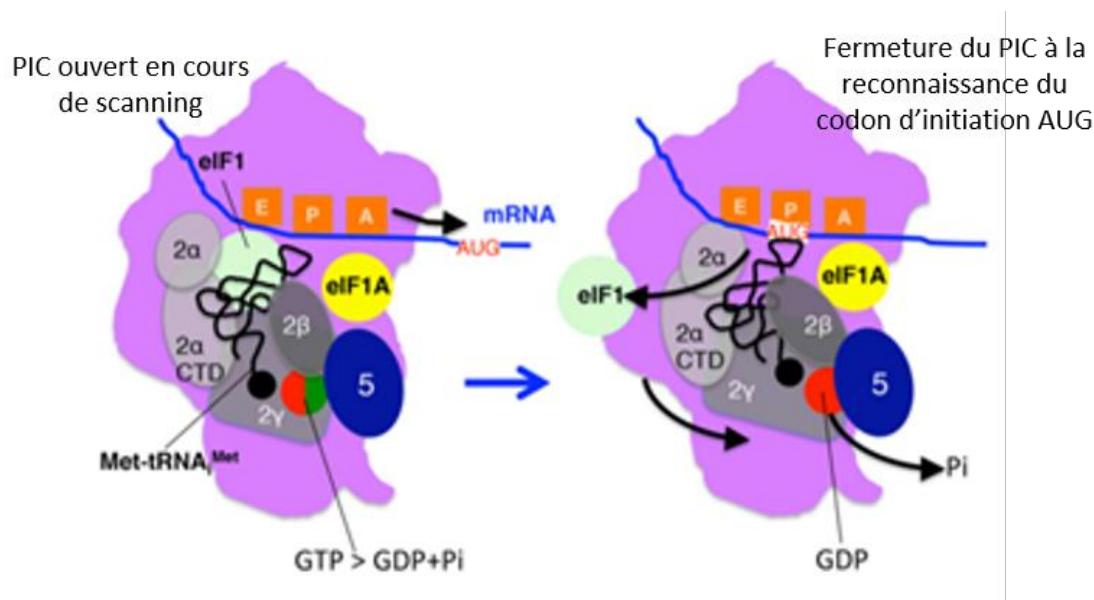


Figure 27 : Accommodation de l'ARNt initiateur lors de la reconnaissance du codon initiateur AUG

Schéma représentant les positions du PIC 43S en position ouverte et des différents facteurs d'initiation lors du scanning (à gauche) puis les changements conformationnels induits par la reconnaissance du codon d'initiation par l'ARNti^{Met}, conduisant à la fermeture du PIC, à la dissociation de eIF1 et à la conversion du GTP en GDP. *Adapté de Dever, 2016*¹²⁰.

Le plus souvent, l'initiation a lieu sur le premier codon AUG rencontré, mais ce n'est pas toujours le cas, notamment chez les mammifères. Ce phénomène s'appelle le « leaky scanning » et dépend du contexte nucléotidique de l'AUG et de la taille de la région 5'UTR. La séquence optimale pour initier la traduction est le contexte Kozak suivant : RCC**AUGG**, où R est une purine. Les positions les plus importantes sont les positions -3 et +4 par rapport au A de l'AUG initiateur désigné comme +1 (en gras dans la séquence au-dessus). Cette séquence stabilise probablement le PIC au moment de la reconnaissance du codon d'initiation¹²⁶. Cependant, l'initiation de la traduction peut également s'effectuer sur un autre codon que le codon AUG. Ces codons partagent pour la plupart deux nucléotides avec le codon AUG, parfois seulement un seul^{127,128}.

B – L'initiation IRES-dépendante

Il existe un autre mode d'initiation, indépendant de la coiffe, et qui fait intervenir le recrutement interne du ribosome au site d'initiation de la traduction. Le PIC 43S est recruté grâce à des éléments IRES (Internal Ribosome Entry Site) présents dans l'ARNm même, au niveau du codon initiateur. Ce sont des séquences de taille variable formant des structures secondaires qui vont interagir avec le PIC. Découverts chez les virus, ces éléments permettent de détourner la machinerie traductionnelle au profit de la traduction des ARNs viraux, qui ne possèdent pas de coiffe. Ces éléments également présents dans certains ARNm cellulaires permettent leur traduction lorsque certains facteurs d'initiation sont indisponibles comme lors d'un stress cellulaire par exemple.

Les IRES viraux sont classés en quatre classes suivant deux critères : leurs structures secondaires et tertiaires, et leur mode d'initiation de la traduction (Figure 28). Les IRES de classe I sont capables de recruter la sous-unité 40S en amont de la région codante - indépendamment de la coiffe - qui scanne ensuite l'ARN de 5' en 3' jusqu'au codon d'initiation. Ils nécessitent tous les facteurs d'initiation excepté eIF4E. Les IRES de classe II recrutent directement la sous-unité 40S au niveau du codon initiateur sans scanning. Ils nécessitent également tous les facteurs d'initiation excepté eIF4E. Les IRES de classe III présentent des structures secondaires et tertiaires plus complexes, comme des pseudonoeuds, et ne requièrent qu'un petit nombre de facteurs d'initiation, comme eIF2, eIF3 et eIF5. Ils recrutent la sous-unité 40S directement au codon initiateur sans scanning. Les IRES de classe IV présentent les structures les plus complexes et contiennent généralement plusieurs pseudonoeuds. Ils ne requièrent aucun facteur d'initiation, ni l'ARN^{Met}. Dans ce cas, on ne peut pas parler de PIC 43S car la sous-unité 40S n'est accompagnée d'aucun facteur.

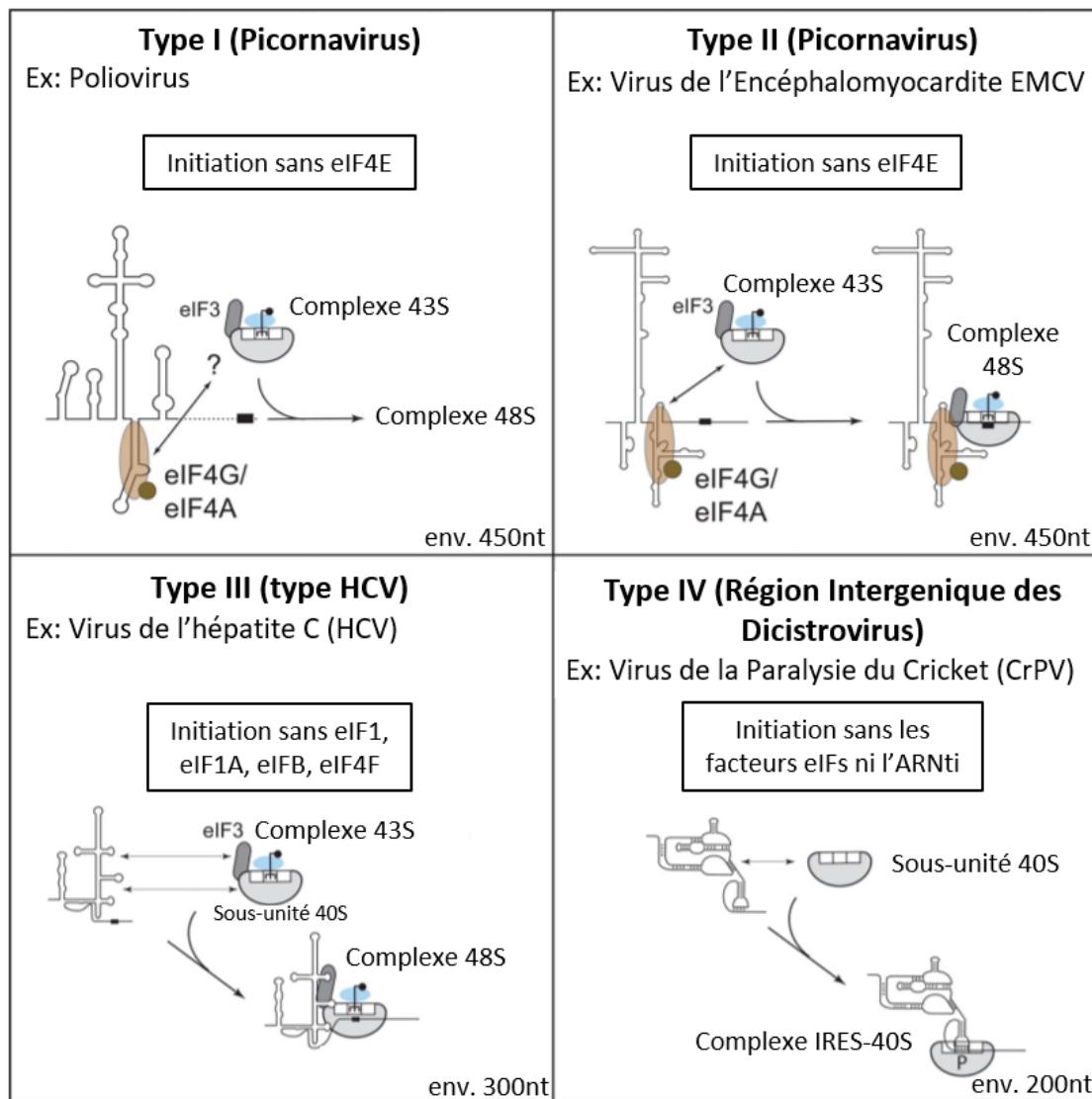


Figure 28 : Les quatre types d'IRES viraux

Il existe quatre types d'IRES chez les virus, qui sont classés suivant leurs mécanismes d'action. Les types I et II sont ceux qui nécessitent le plus de facteurs d'initiation tandis que les types IV ne nécessitent ni facteurs d'initiation ni ARNt initiateur. *Adapté de Jackson, 2010*¹²⁹.

Les IRES de classe I, II et III sont situés dans les régions 5'UTR tandis que les IRES de classe IV sont situés dans les régions intergéniques¹³⁰.

Le fonctionnement des IRES cellulaires n'est pas aussi bien connu. Néanmoins il est aujourd'hui admis que certains ARNm cellulaires possèdent des IRES particulièrement importants en cas de stress cellulaire, lorsque les facteurs d'initiation de la traduction sont indisponibles. Ces IRES sont localisés dans des gènes dont les protéines sont impliquées dans le cycle cellulaire, comme les facteurs de prolifération FGF-1 et FGF-2, le facteur d'angiogenèse VEGF ou encore la protéine p53, fortement impliquée dans le contrôle du cycle cellulaire. La dérégulation de l'expression de ces protéines est souvent impliquée dans les processus de cancérogenèse^{131,132}. On estime à environ 10% la proportion de 5'UTRs susceptibles de contenir une séquence IRES qui permettrait d'initier la traduction de manière cap-indépendante¹³³.

L'initiation au niveau d'IRES fait également intervenir la présence de facteurs agissant en *trans*, les ITAFs (IRES trans-acting factors)¹³⁴, dont certains sont partagés par les IRES viraux et cellulaires. Ces ITAFs sont des protéines interagissant avec l'ARN et qui participent à de nombreuses autres fonctions cellulaires. Toutefois ils ne sont pas impliqués dans l'initiation cap-dépendante de la traduction¹³⁵. Il est admis que les ITAFs, par leurs interactions avec les IRES, pourraient contribuer à stabiliser les structures instables ou à induire des changements conformationnels de l'IRES qui permettraient le recrutement et le positionnement correct du ribosome¹³⁶. Aujourd'hui seuls quelques ITAFs sont bien caractérisés. Un exemple dont le mécanisme a été révélé concerne l'ARNm d'APAF1. L'ITAF UNR se lie à cet ARNm au niveau d'une région riche en purines située dans une tige-boucle. Cette liaison ouvre deux structures tiges-boucles, ce qui permet la liaison de l'ITAF NPTB, rendant ainsi accessible le codon d'initiation de la traduction pour le recrutement de la sous-unité 40S¹³⁷ (Figure 29). La présence des ITAFs semble nécessaire à l'initiation au niveau de certains IRES, ces facteurs sont donc des éléments de régulation de l'initiation IRES-dépendante.

C – Autres modes d'initiation

Certains ARNm –très rares chez les mammifères - ont une région 5'UTR très courte (une vingtaine de nucléotides). L'initiation canonique de la traduction par scanning de la région 5'UTR ne peut pas avoir lieu. L'initiation de la traduction de ces ARNm fait donc intervenir d'autres mécanismes. Tout d'abord, le codon initiateur de ces ARNm est situé au sein d'une séquence particulière nommée TISU (translation initiator of short 5'UTRs), qui est la suivante : SAAS**AUG**GGCGC (ou S est une Guanine ou une Cytosine). Le recrutement du PIC fait intervenir les facteurs eIF1 et

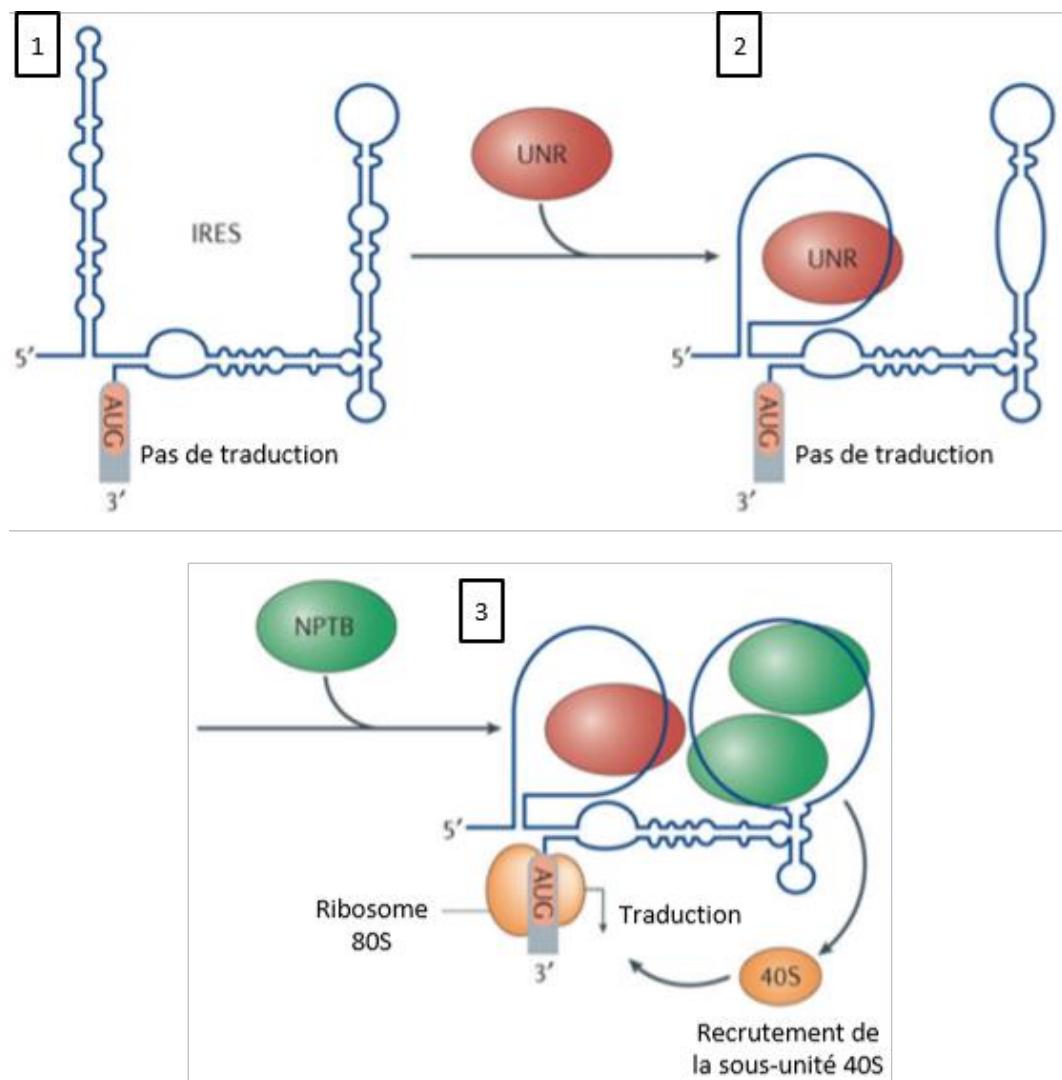


Figure 29 : Rôle des ITAFs dans la traduction de l'ARNm d'APAF1

La traduction de l'ARNm est initiée au niveau d'un IRES (1) et nécessite l'intervention d'ITAFs. Tout d'abord l'ITAF UNR se lie à cet ARNm au niveau d'une région riche en purines située dans une tige-boucle (2). Cette liaison ouvre deux structures tiges-boucles, ce qui permet la liaison de l'ITAF NPTB (3). Le codon d'initiation de la traduction est alors disponible pour le recrutement de la sous-unité 40S et l'initiation de la traduction peut avoir lieu. *Adapté de Leppek, 2018*¹³⁶.

elf4G. Une fois le PIC placé au niveau du codon initiateur et du fait de sa proximité avec la coiffe, les facteurs elf1A, elf2 et elf3 se détachent et la sous-unité 60S peut alors s'associer.^{131,138} (Figure 30).

Un autre type d'initiation, appelé « ribosome shunting », fait intervenir le contournement d'une partie du 5'UTR par le ribosome, qui vient alors directement se positionner au codon initiateur. Ce type d'initiation est dépendant de la coiffe et ne présente aucune activité IRES ni aucune structure de l'ARNm particulière. Un des gènes cellulaires dont la traduction s'initie par ce processus est le gène *BACE1*, qui code l'enzyme β -sécrétase, impliquée dans la formation des plaques A β -amyloïdes chez les patients atteints de la maladie d'Alzheimer. Le mécanisme de ce processus d'initiation reste aujourd'hui non compris¹³¹.

Il existe également un mode d'initiation cap-indépendant, proche de l'initiation IRES-dépendante mais qui en réalité ne fait pas intervenir un IRES mais un élément CITE (cap-independent translation enhancer). Cette initiation peut avoir lieu sans la coiffe mais nécessite l'extrémité 5' de l'ARNm (contrairement aux IRES). Les CITEs sont localisés au sein des UTR (5' ou 3') et attirent les facteurs d'initiation, ce qui permet l'assemblage des complexes d'initiation. Les facteurs elf4G et elf3 peuvent être recrutés directement ou indirectement au sein du 5'UTR via des interactions ARN-protéine, ce qui permet à d'autres composants de la machinerie d'initiation de la traduction d'être recrutés (Figure 30). Chez l'Homme, l'exemple de l'ARNm du gène *APAF1* (facteur impliqué dans l'apoptose induite suite à des dommages de l'ADN) a été bien décrit : en condition de stress, le facteur de liaison à la coiffe elf4E est indisponible et l'initiation de la traduction a donc lieu au niveau d'un élément CITE¹³². Il est intéressant de noter que plusieurs niveaux de régulation peuvent se superposer. C'est le cas du gène *APAF1* puisque comme je l'ai indiqué précédemment, l'initiation de la traduction de cet ARNm peut aussi s'effectuer au niveau d'un IRES.

Enfin, on peut trouver dans les régions 5'UTR la modification de base N⁶-méthyladénosine (m⁶A). Cette modification favorise l'initiation de la traduction à certains sites appelés MIRES (m⁶A-induced ribosome engagement sites) en se liant au facteur elf3, ce qui permet ensuite de recruter la sous-unité 40S. Elle permet l'initiation de la traduction même en absence du facteur elf4E, mais nécessite tout de même un processus de scanning dépendant de l'extrémité 5'. Le mécanisme de reconnaissance des m⁶A par les facteurs d'initiation de la traduction reste encore à déterminer¹³² (Figure 30).

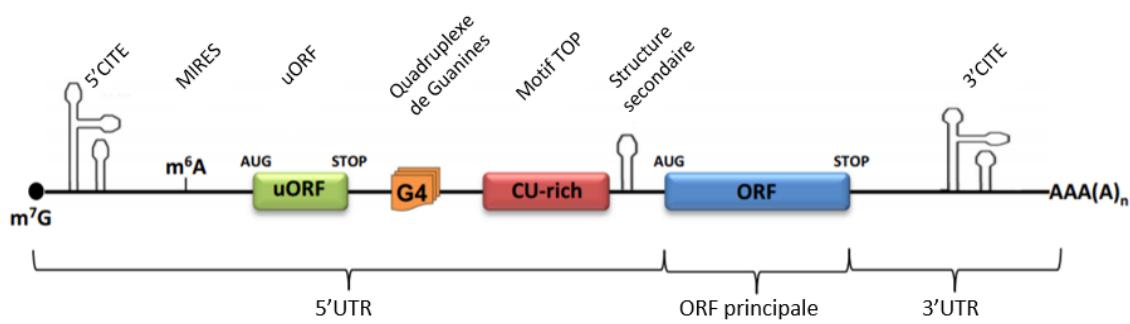


Figure 30 : Schéma d'un ARNm montrant différents éléments agissant en *cis* sur l'initiation de la traduction

Différents éléments qui agissent sont positionnés le long de l'ARNm : éléments CITE (cap-independent translation enhancer) 5' et 3', sites Mires ou sites du recrutement du ribosome induit par des N⁶-méthyladénosine m⁶A, uORFs, quadruplexes de guanines et motifs d'oligopyrimidine TOP. *Adapté de Lacerda, 2017*¹³².

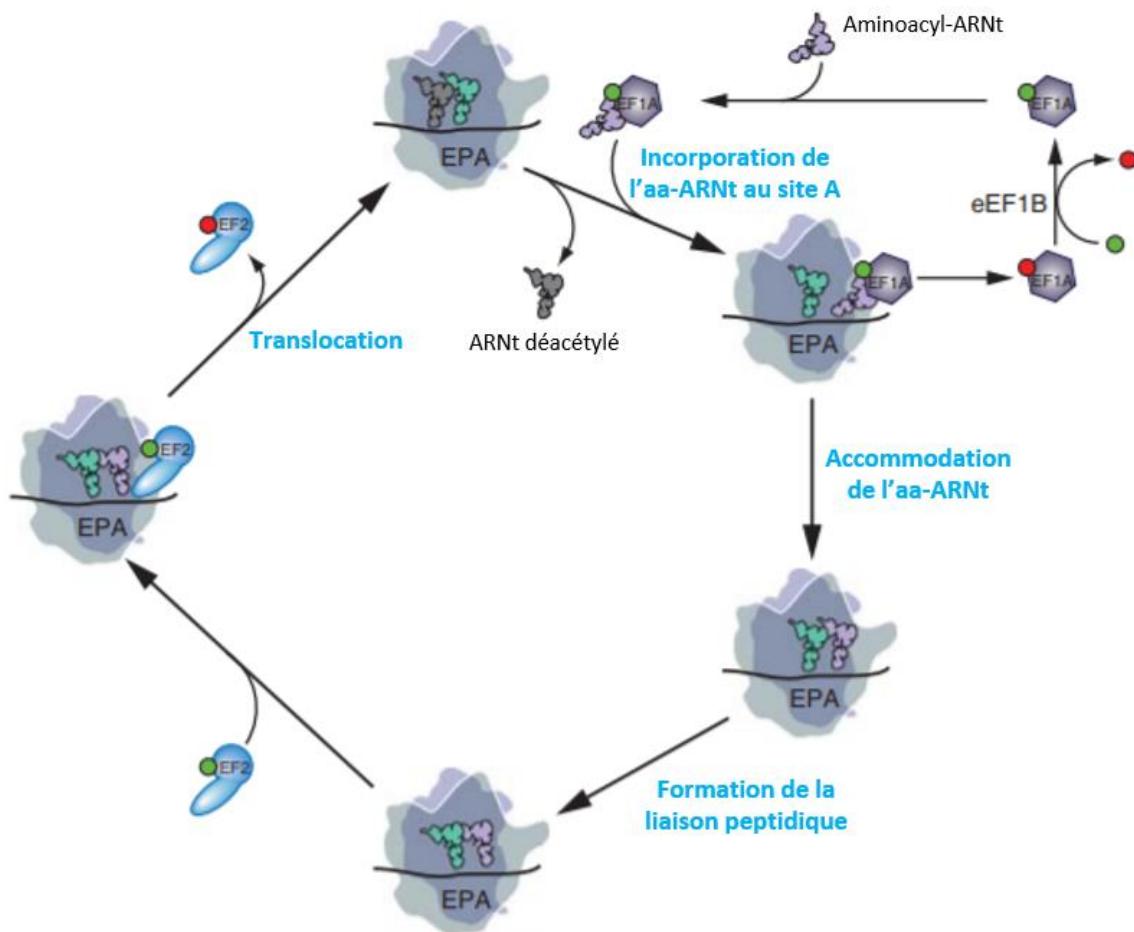


Figure 31 : Elongation de la traduction

Le cycle d'élongation de la traduction débute par l'incorporation de l'aminoacyl-ARNt lié à eEF1A-GTP au niveau du site A, par complémentarité entre codon de l'ARNm au site A et la boucle anticodon de l'ARNt. En parallèle, l'ARNt déacétylé au site E est expulsé. Le GTP est hydrolysé, eEF1A est libéré et l'aa-ARNt s'accommode au site A. eEF1A-GDP est recyclé en eEF1A-GTP par eEF1B. La formation de la liaison peptidique s'accompagne de mouvements des ARNt dans des états hybrides A/P et P/E. La liaison d'eEF2-GTP et l'hydrolyse du GTP en GDP induisent la translocation des ARNt aux sites P et E. Le ribosome est alors prêt pour un nouveau cycle d'élongation. Les molécules de GTP sont en vert et les molécules de GDP sont en rouge. *Adapté de Dever, 2012*¹³³.

2 – L’élongation de la traduction

L’étape d’élongation de la traduction correspond au décodage de l’ARNm en acides aminés suivant le code génétique et à la synthèse du peptide correspondant. Le code génétique établit une correspondance entre les codons et les acides aminés. Il est dit « redondant » ou « dégénéré » car plusieurs codons peuvent correspondre à un même acide aminé (il existe 64 codons pour 20 acides aminés). L’association entre chaque acide aminé et ses codons a été établie en 1965 par Nirenberg. Il existe également un codon initiateur AUG qui correspond à une méthionine et trois codons stop qui permettent l’arrêt de la traduction : UAG, UAA et UGA. Chez certains organismes ou dans les mitochondries humaines, le code génétique est différent de ce code canonique : certains codons sont dits « réassignés ». En plus des 20 acides aminés classiques, on trouve aussi deux acides aminés naturels rares - ces acides aminés ne sont retrouvés que chez quelques organismes et dans seulement quelques protéines et sont formés à partir d’acides aminés ordinaires -, la sélénocystéine et la pyrrolysine, dont le codon correspondant est en fait un codon stop (respectivement UGA et UAG). L’incorporation de ces deux acides aminés ne se produit qu’à un petit nombre de codons stops et correspond à un événement de recodage (voir paragraphe III-3 sur la fidélité de la terminaison de la traduction)¹⁴⁰.

L’élongation de la traduction fait intervenir trois étapes majeures : le décodage du codon de l’ARNm au site A, la formation de la liaison peptidique et la translocation de l’ARNm et de l’ARNT (Figure 31).

Lors du démarrage de la traduction, le site P du ribosome est occupé par l’ARNT^{Met} et le site A est vide, prêt à recevoir un nouvel ARNT qui va décoder le second codon de l’ORF. Il doit être sélectionné parmi un pool d’aa-ARNT disponibles. Chez les eucaryotes, les aa-ARNT sont recrutés au site A du ribosome au sein d’un complexe ternaire eEF1A-GTP-aa-ARNT, par appariement codon-anticodon entre l’ARNm au site A et l’aa-ARNT. La reconnaissance du codon par l’ARNT cognat (c’est-à-dire l’ARNT dont l’anticodon est complémentaire au codon) conduit à l’hydrolyse du GTP par eEF1A puis à la libération de eEF1A-GDP, ce qui permet à l’aa-ARNT de s’accommoder correctement au site A. Le centre peptidyl-transférase (PTC) participe au positionnement de l’aa-ARNT au site A et du peptidyl-ARNT au site P et permet la formation rapide de la liaison peptidique. L’ARNT au site A porte alors le peptide naissant, c’est le peptidyl-ARNT, et l’ARNT au site P est déacétylé. La formation de la liaison peptidique entraîne le cliquetage du ribosome, qui positionne respectivement l’ARNT déacétylé et le peptidyl-ARNT dans les états hybrides P/E et A/P avec les bras accepteurs dans les sites E et P et les boucles anticodon aux sites P et A.

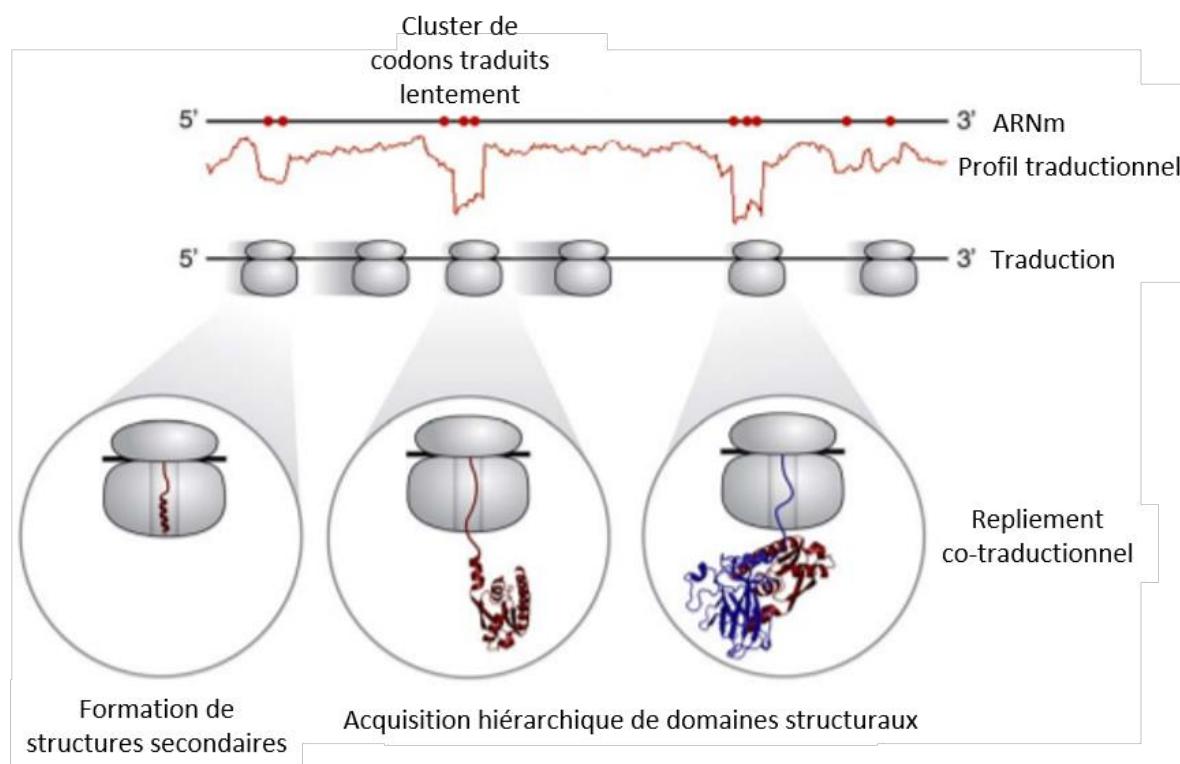


Figure 32 : Repliement co-traductionnel des protéines

Des discontinuités locales de la traduction permettent la coordination co-traductionnelle du repliement de la protéine naissante. Les clusters de codons associés à des ARNt peu abondants sont représentés par des points rouges sur l'ARNm et des minima dans le profil de traduction, reflétant le trafic non-uniforme des ribosomes le long de l'ARNm. Au niveau des minima locaux les plus bas, la traduction ralentit suffisamment pour permettre une coordination de la traduction avec le repliement hiérarchique de domaines protéiques, tandis qu'au niveau des minima locaux moins importants, la cinétique de la traduction permet la formation de structures secondaires. *Adapté de Zhang, 2011¹⁴¹.*

L'étape de translocation correspond au passage de l'ARNt déacétylé du site P au site E et du peptidyl-ARNt du site A au site P. Cette étape est étroitement liée au mouvement de cliquetage de la petite sous-unité par rapport à la grande sous-unité. Ce mouvement est induit par le facteur eEF2, qui se lie en complexe avec une molécule de GTP. C'est l'hydrolyse de ce GTP qui permet les mouvements du ribosome à l'origine de la translocation. A cause de l'appariement codon-anticodon, l'ARNm bouge par rapport au ribosome durant l'étape de translocation, libérant le site A, qui voit arriver un nouveau complexe eEF1A-GTP-aa-ARNt pour décoder le codon suivant et démarrer un nouveau cycle d'elongation. Ce cycle d'elongation se répète jusqu'au codon de terminaison de la traduction. Le recyclage d'eEF1A-GDP en eEF1A-GTP entre deux cycles d'elongation nécessite le facteur d'échange de guanine (GEF = guanine-nucleotide exchange factor) eEF1B^{120,142}.

Au cours de l'elongation, le PTC catalyse la formation de la liaison peptidique entre la chaîne naissante portée par le peptidyl-ARNt (site P) et l'acide aminé porté par l'aminoacyl-ARNt (site A). Cette liaison résulte d'une attaque nucléophile du groupe α -amino de l'aa-ARNt au site A sur le carbone carbonyl du peptidyl-ARNt au site P. L'aa-ARNt devient alors peptidyl-ARNt et le peptidyl-ARNt devient ARNt déacétylé. L'activité du PTC repose sur l'ARNr et son répertoire limité de groupes actifs¹⁴³.

La vitesse moyenne d'elongation est de 15-20 acides aminés par seconde chez les procaryotes et de 1-5 acides aminés par seconde chez les eucaryotes¹⁴¹ mais cette vitesse n'est pas constante tout le long de l'ARNm. Cela peut s'expliquer par différents facteurs : l'abondance des différents ARNt, la présence de structures secondaires au sein de l'ARNm, le contexte du codon, des pauses du ribosome ou encore des collisions entre ribosomes au sein des polysomes¹⁴³.

Le repliement de beaucoup de protéines a lieu de manière co-traductionnelle (Figure 32). Dans le tunnel de sortie du ribosome, le peptide naissant peut établir des interactions locales et stabiliser des structures α -hélicoïdales. Les contacts entre des régions éloignées ont plutôt lieu en dehors du ribosome, après la sortie de segments plus longs, ce qui peut conduire à l'adoption de conformations intermédiaires par le peptide. Le repliement co-traductionnel est limité par la vitesse de la sortie du peptide naissant - et donc de l'elongation de la traduction – qui impose des contraintes conformationnelles. Des études récentes suggèrent que la vitesse de la traduction régule finement le repliement co-traductionnel en fournissant un délai pour le repliement séquentiel de régions distinctes du peptide naissant.¹⁴¹ Ainsi, des modèles prédictifs ont montré que la vitesse à laquelle les segments du peptide naissant émergeaient du ribosome était cruciale pour le

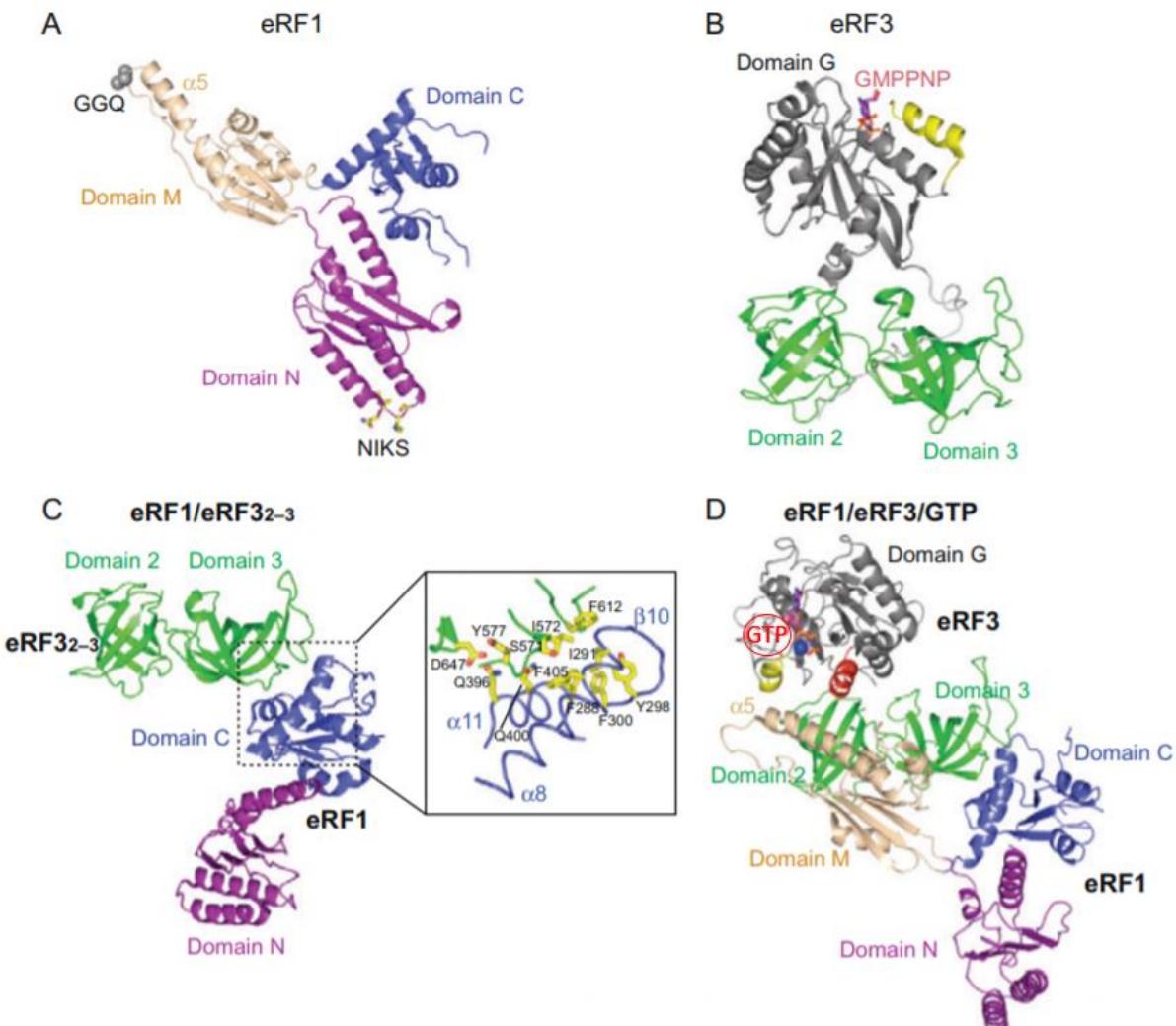


Figure 33 : Structures d'eRF1, eRF3 et des complexes eRF1-eRF3 et eRF1-eRF3-GTP

(A) Structure du facteur eRF1 de *S. pombe* avec les motifs reconnaissables NIKS et GGQ. (B) Structure d'eRF3 de *S. pombe*. (C) Structure du complexe eRF1-eRF3 chez *S. pombe*. (D) Structure du complexe eRF1-eRF3-GTP chez *S. pombe*. La molécule de GTP est représentée en rouge. *Adapté de Jackson, 2012*¹⁴⁴.

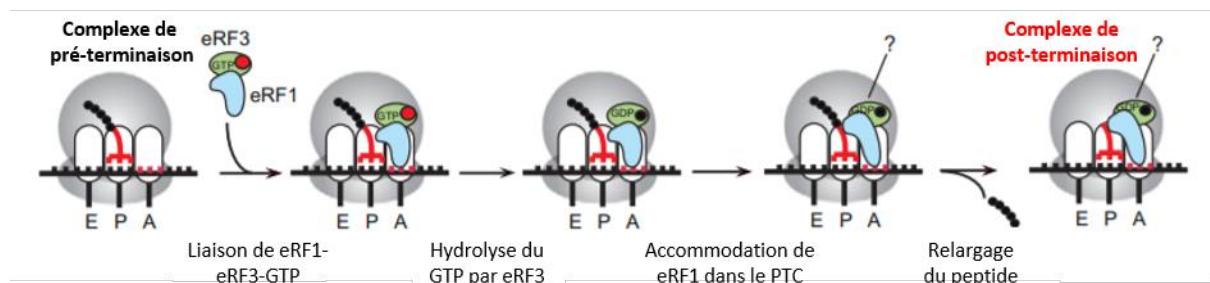


Figure 34 : Modèle de terminaison de la traduction

Le complexe de pré-terminaison comprend le peptidyl-ARNt au site P du ribosome. Les facteurs de terminaison eRF1 et eRF3 se lient au complexe de terminaison sous la forme d'un complexe ternaire eRF1-eRF3-GTP. Le codon stop est reconnu par eRF1. Après l'hydrolyse du GTP par eRF3, eRF1 induit la libération du peptide. eRF1 reste associé au complexe de post-terminaison. *Adapté de Jackson, 2012*¹⁴⁴.

repliement et pouvait influencer la conformation finale de la protéine¹⁴⁵. De façon intrigante, dans certains ARNm, il semble y avoir des régions entières qui présenteraient une traduction lente. En effet, on trouve des groupes de codons dont les ARNt correspondants sont peu abondants¹⁴⁶. La traduction de ces régions présenterait très probablement une vitesse réduite, ce qui laisserait le temps à la région du peptide qui vient de sortir du tunnel d'acquérir une conformation correcte.

Le repliement des protéines au cours de la traduction est assisté par des protéines chaperonnes. Ces chaperonnes permettent également aux protéines de conserver leurs conformations ou de les retrouver en cas de dénaturation à cause d'une augmentation de la température ou d'une variation des conditions physico-chimiques de la cellule.

La vitesse et la fidélité de la synthèse protéique sont des processus fondamentaux qui affectent la qualité du protéome - et donc la survie de la cellule - et sa potentielle compétitivité dans un environnement naturel. Je discuterai des notions de fidélité et de régulation de l'elongation de la traduction dans le chapitre III-2.

3 – La terminaison de la traduction

La terminaison de la traduction a lieu quand le ribosome arrive à la fin de la séquence codante de l'ARNm, c'est-à-dire quand le site A du ribosome est occupé par un codon stop UAA, UAG ou UGA.

Chez les eucaryotes, la terminaison nécessite deux facteurs protéiques eRF1 et eRF3. Le facteur de classe I, eRF1, est responsable de la reconnaissance hautement fidèle du codon stop et de l'hydrolyse du peptidyl-ARNt. Le facteur de classe II, eRF3, est une GTPase.

eRF1 présente une conformation qui mime un ARNt et est composée de trois domaines¹⁴⁷ (Figure 33A). Le domaine N-terminal participe à la reconnaissance du codon stop, qui sera traitée dans le paragraphe ci-dessous. Le domaine central (M) est fonctionnellement analogue à la tige acceptrice des ARNt et s'étend de la même façon dans le PTC pour contribuer à la libération du peptide¹⁴⁷. Ce domaine contient un motif universellement conservé Glycine-Glycine-Glutamine (GGQ), essentiel pour l'hydrolyse du peptide¹⁴⁸. Le domaine C-terminal a un rôle dans l'interaction d'eRF1 avec eRF3¹⁴⁹.

La reconnaissance du codon stop s'effectue à travers le motif hautement conservé NIKS du domaine N-terminal eRF1^{150,151}, ainsi que les motifs YxCxxxF¹⁵² et GTS¹⁵³ et la poche P1^{154,155}. Le motif NIKS participe à la reconnaissance de

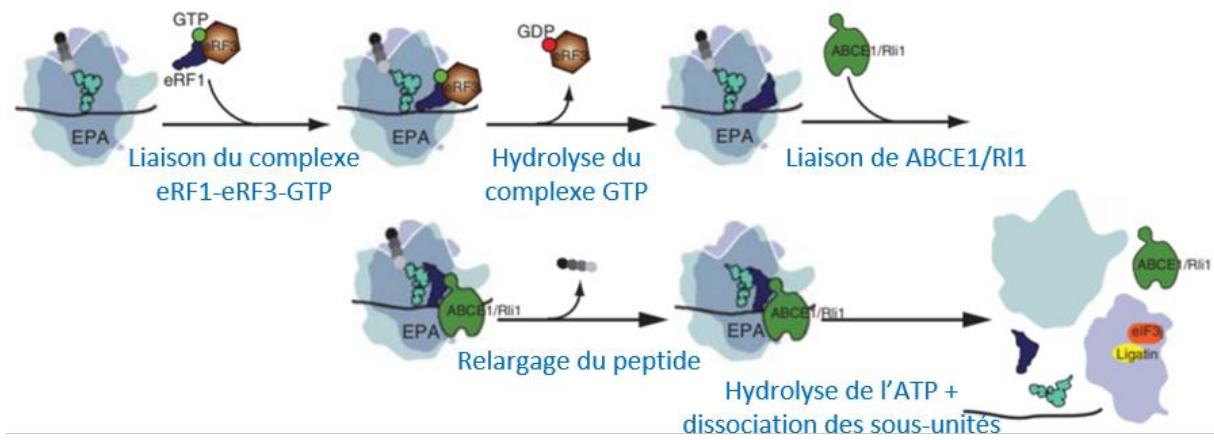


Figure 35 : Modèle de couplage des étapes de terminaison et de recyclage de la traduction

Lorsque le site A du ribosome arrive au niveau d'un codon de terminaison, le complexe eRF1-eRF3-GTP se lie au site A dans un état de pré-accommodation. Le GTP est hydrolysé et eRF3 est dissocié. Le facteur ABCE1/Ri1p se lie et facilite accommodation d'eRF1 au site A. Le peptide est libéré. L'hydrolyse de l'ATP est suivie de la dissociation des sous-unités. *Adapté de Dever, 2012¹³⁹.*

l'uridine en première position du codon stop tandis que les motifs YxCxxxF et GTS interviennent dans la reconnaissance des nucléotides en deuxième et troisième positions du codon stop. Il a également été montré que lors de cette étape, l'ARNm se trouve dans une configuration compactée, permettant ainsi l'accommodation de quatre nucléotides au site A et non trois. Cette compaction se produirait au moment de la reconnaissance du codon stop par eRF1 et serait importante pour la stabilisation d'eRF1 dans le site A, permettant une bonne fidélité de la terminaison de la traduction. Il a été montré que cette stabilisation est plus forte en présence d'une purine en position +4, ce qui expliquerait le fait qu'elles sont très largement majoritaires à cette position chez les eucaryotes¹⁵⁶.

eRF3 interagit avec les régions centrale et C-terminale d'eRF1 via son domaine C-terminal¹⁴⁹ (Figure 33C). L'interaction entre eRF1 et eRF3 stabilise la liaison au GTP, le complexe ternaire eRF1-eRF3-GTP est ainsi particulièrement stable¹⁴⁴ (Figure 33D). Le domaine N-terminal d'eRF3 est moins conservé (Figure 33B) et est impliqué dans les interactions avec la PAPB et Upf1p¹⁵⁷⁻¹⁵⁹. *In vitro*, eRF3 accélère la libération de la protéine néo-synthétisée et augmente l'efficacité de la terminaison de manière dépendante de l'hydrolyse du GTP¹⁶⁰.

Lorsque le site A du ribosome est positionné au niveau d'un codon de terminaison, le complexe ternaire eRF1-eRF3-GTP s'associe au ribosome, dans un état de pré-accommodation, par interaction d'eRF1 avec le codon stop. Cette association est similaire à l'association du complexe ARNt-eEF1A au ribosome lors de l'elongation¹³⁹. Le GTP est ensuite hydrolysé par eRF3 et eRF3 est dissocié¹⁶¹ (Figure 34). Il existe ensuite deux voies actuellement discutées pour permettre la libération du peptide naissant¹⁶². La première fait intervenir le facteur ABCE1 (Rli1p chez la levure), qui se lie au domaine C-terminal d'eRF1 et facilite son accommodation¹³⁹ (Figure 35). Cela consiste au déplacement du domaine M d'eRF1 dans le PTC - le motif GGQ est alors à proximité de la liaison ester entre l'ARNt et le peptide naissant – permettant ainsi l'hydrolyse du peptidyl-ARNt par une molécule d'eau. La deuxième ne nécessite pas le recrutement d'ABCE1 pour l'accommodation d'eRF1 et la libération de la protéine néo-synthétisée. ABCE1 ne se fixe alors à eRF1 qu'après cette étape.

4 – Le recyclage de la machinerie de traduction et la réinitiation

❖ Le recyclage de la machinerie de traduction :

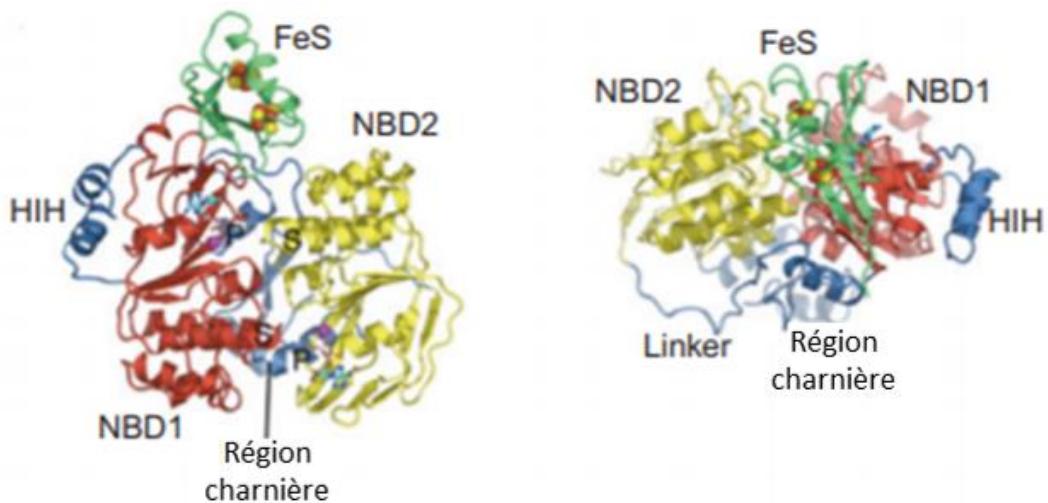


Figure 36 : Structure d'ABCE1 de *P. abyssi*

On reconnaît les domaines de liaison aux nucléotides NBD1 (avec son insertion hélice-boucle-hélice HIH) et NBD2, séparés par la région charnière. *Adapté de Jackson, 2012*¹⁴⁴.

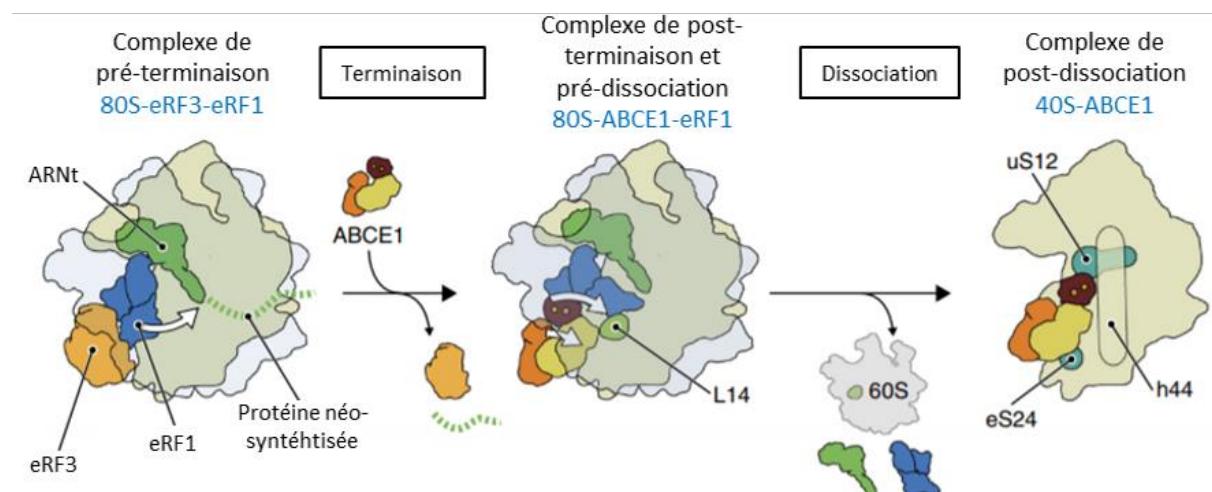


Figure 37 : Dissociation des sous-unités 40S et 60S par ABCE1

Après la reconnaissance d'un ribosome au codon stop par eRF1 et eRF3, et le départ d'eRF3, ABCE1 se lie à eRF1 dans un état intermédiaire, semi-fermé. La fermeture des domaines NBDs pousse le domaine FeS vers le facteur au site A, eRF1, ce qui déstabilise la structure et dissocie les deux sous-unités. *Adapté de Heuer, 2017*¹⁶³.

Après l'étape de terminaison de la traduction et la libération de la protéine néo-synthétisée, les sous-unités ribosomiques, l'ARNm, l'ARNt déacétylé et eRF1 doivent être dissociés pour être disponible pour un nouveau cycle de traduction. Il semblerait que les facteurs de terminaison eRF1 et eRF3 soient suffisants pour induire le recyclage des composants de la machinerie de traduction mais avec une efficacité très faible. La protéine ABCE1 augmente fortement cette efficacité¹⁴⁴.

La protéine ABCE1, pour ATP-binding cassette protein E1, est une ATPase qui convertit l'énergie générée par l'hydrolyse de l'ATP en mouvements mécaniques qui séparent les sous-unités 40S et 60S¹⁶⁴. Pendant la terminaison, elle stimule l'hydrolyse du peptidyl-ARNt par eRF1¹¹⁸. Elle contient deux domaines de liaison aux nucléotides (NBDs), entre lesquels se lie l'ATP¹⁶⁵, et un domaine N-terminal qui comprend quatre groupements fer-soufre (FeS)¹¹⁸ (Figure 36).

Les deux domaines NBDs sont positionnés dans une conformation ouverte grâce à une région charnière hautement conservée qui lie les domaines NBDs à l'extrémité C-terminale. Cette région charnière pourrait agir comme un pivot lors des changements conformationnels induits par l'hydrolyse de l'ATP¹⁴⁴.

Le mécanisme de dissociation des sous-unités du ribosome par ABCE1 a été résolu en 2017 grâce à une étude structurale. Tout d'abord, le facteur ABCE1 se lie à la région C-terminale d'eRF1 (où eRF3 était lié précédemment) à travers son domaine FeS. Puis une molécule d'ATP vient se lier et les domaines NBDs se ferment, poussant le domaine FeS vers le facteur au site A, eRF1, ce qui déstabilise les ponts entre les sous-unités 40S et 60S et entraîne leur dissociation¹⁶³ (Figure 37). Après la dissociation des sous-unités par ABCE1, l'ARNt déacétylé et l'ARNm restent liés à la sous-unité 40S. La sous-unité 60S est prise en charge par le facteur d'initiation eIF6 tandis que les facteurs d'initiation eIF3, eIF1, eIF1A se lient au complexe 40S-ARNm-ARNt pour permettre la libération de l'ARNt. L'ARNm est libéré après recrutement du facteur eIF3j¹⁶⁵. Un autre modèle fait intervenir la ligatine pour permettre la dissociation de l'ARNt déacétylé et de l'ARNm¹⁴⁴ (Figure 38). La réassocation des deux sous-unités est empêchée par le domaine FeS d'ABCE1 qui bloque toute interaction de la sous-unité 40S avec la protéine uL14 de la sous-unité 60S¹⁶³. Après le départ d'ABCE1, ce sont les facteurs eIF1 et eIF6 respectivement liés aux 40S et 60S qui empêchent la réassocation des sous-unités^{166,167}.

❖ La réinitiation de la traduction :

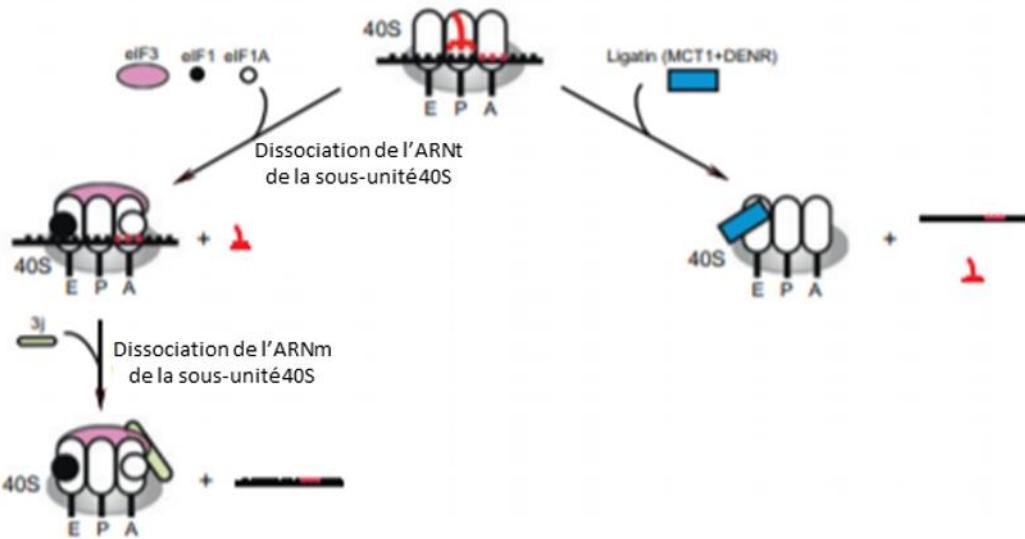


Figure 38 : Dissociation de l'ARNm et de l'ARNt de la sous-unité 40S

Modèle de dissociation de l'ARNt et de l'ARNm de la sous-unité 40S, soit par les facteurs d'initiation eIF3, 3j, 1 et 1A (à gauche), soit par la ligatine (à droite). *Adapté de Jackson, 2012*¹⁴⁴.

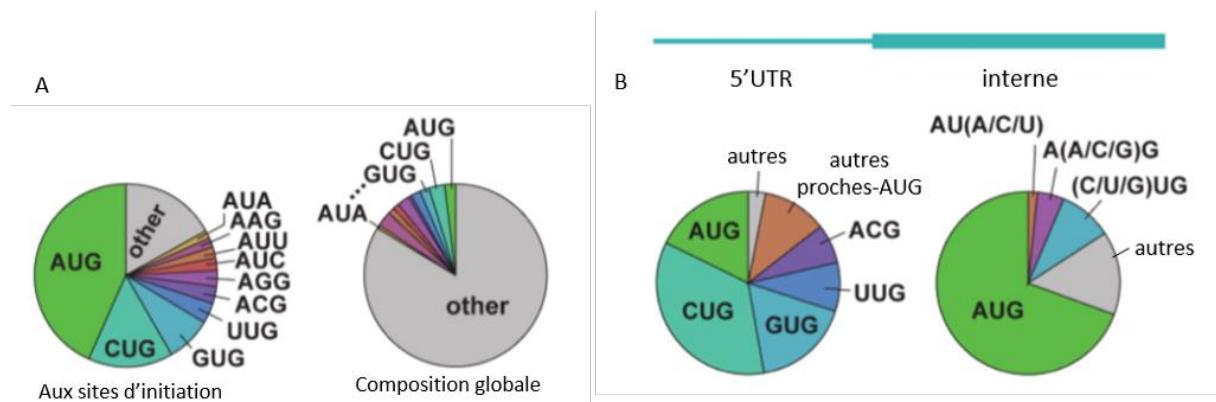


Figure 39 : Distribution des codons d'initiation dans des cellules souches embryonnaires de souris

A : Distribution des codons AUG et proches AUG aux sites d'initiation comparé à la composition globale des codons. B : Distribution des codons d'initiation dans les 5'UTR (uORFs) et internes. *Adapté de Ingolia, 2011*¹²⁸.

Dans certains cas, la dissociation du ribosome de l'ARNm n'est pas totalement efficace et conduit à des événements de réinitiation de la traduction. Ces événements de réinitiation peuvent avoir lieu en aval ou en amont de l'ORF principale, le scanning de la sous-unité 40S n'étant alors pas dirigé en 3' mais plutôt guidé par l'absence de structures secondaires entre le codon de terminaison et l'AUG¹⁶⁸.

Environ 13% des ARNm chez la levure et 50% des ARNm chez l'Homme contiennent des petites ORFs en amont de la séquence codante, appelées uORFs (upstream ORF)¹⁶⁹. Par définition, une uORF est une courte phase ouverte de lecture située dans la région 5' d'un ARNm contenant une longue ORF codante et qui est composée d'un codon d'initiation et d'un codon de terminaison, séparés par au moins un codon sens¹⁶⁹. En raison de leurs courtes longueurs, les uORFs ne sont généralement pas considérées comme des séquences codantes d'une protéine fonctionnelle mais plutôt des éléments régulateurs de l'ORF principale agissant en *cis*. L'initiation canonique de la traduction faisant intervenir le mécanisme de scanning du PIC 43S, elle débute au premier codon d'initiation AUG rencontré, qui peut être celui d'une uORF. Dans ce cas, la traduction de l'uORF semble retenir les ribosomes et inhiber la traduction de l'ORF principale. La plupart des uORFs régulent négativement l'expression de l'ORF principale mais ce n'est pas toujours le cas¹⁶⁹. Cet aspect de régulation sera discuté plus en détails dans la partie III.

La capacité d'une uORF à induire la réinitiation de la sous-unité 40S au codon d'initiation suivant dépend de quatre facteurs principaux : (1) les caractéristiques nucléotidiques et structurales de l'ARNm autour de l'uORF agissant en *cis* ; (2) la longueur de la séquence et la présence éventuelle de structures secondaires. Il a en effet été montré que l'efficacité de la réinitiation diminuait assez fortement avec l'augmentation de la longueur de l'uORF¹⁷⁰ ; (3) les facteurs d'initiation impliqués dans l'initiation au niveau du codon d'initiation de l'uORF ; (4) la distance entre l'uORF et l'ORF principale. Cette dernière caractéristique détermine la probabilité qu'un nouveau complexe ternaire associé à eIF2 soit recruté par la sous-unité 40S au moment où elle atteint le codon d'initiation de l'ORF principale¹⁶⁹.

Contrairement aux codons d'initiation d'ORFs principales, qui sont en grande majorité des codons AUG (70% dans des cellules souches embryonnaires de souris ESC), il y a un nombre important d'uORFs dont le codon d'initiation n'est pas un AUG (80% de non-AUG dans des ESC)¹²⁸ (Figure 39).

Presque toutes les uORFs des ARNm mammifères semblent permettre un nouveau scanning en aval et la réinitiation de la traduction. Cela suggère qu'il n'y a pas de séquence spécifique pour permettre la réinitiation. En revanche, il existe une séquence spécifique qui empêche fortement la

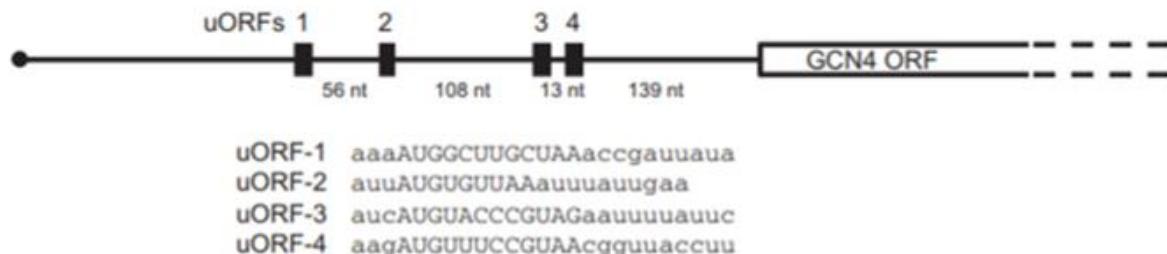


Figure 40 : uORFs et ORF principale de l'ARNm de *GCN4*

L'ARNm de *GCN4* présente quatre uORFs en amont de l'ORF principale, dont la traduction régule la synthèse de la protéine. Cette régulation est décrite dans la partie III-1. *Adapté de Jackson, 2012*¹⁴⁴.

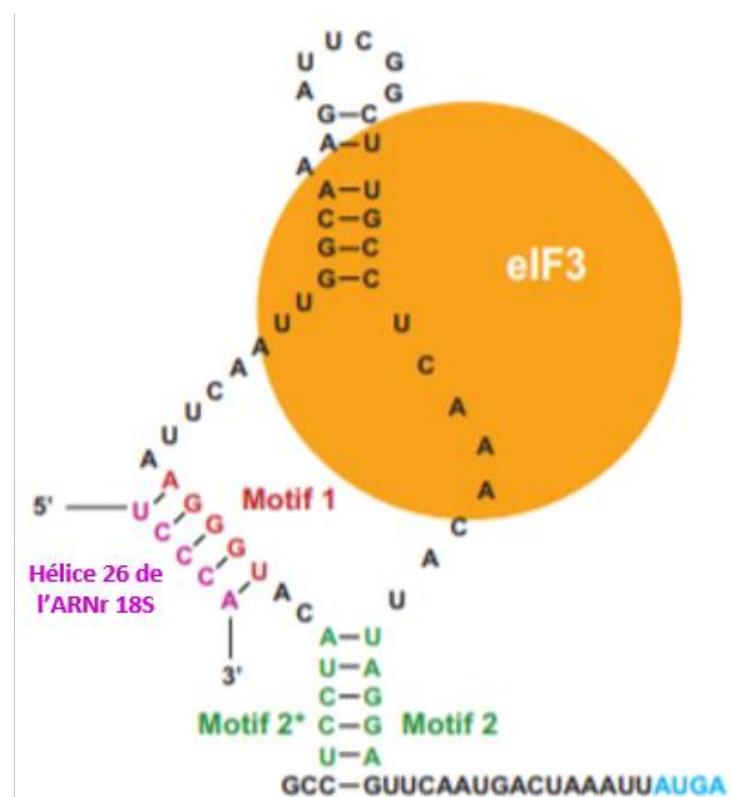


Figure 41 : Elément TURBS permettant la réinitiation chez les calicivirus

Structure secondaire de l'élément TURBS du FCV (Feline Calicivirus) dans laquelle on retrouve les interactions par complémentarité des motifs 2 et 2* ainsi que du motif 1 et de l'hélice 26S de l'ARNr 18S. L'élément TURBS interagit avec le facteur d'initiation de la traduction eIF3 afin de recruter le ribosome pour la réinitiation de la traduction. *Adapté de Gunisova, 2018*¹⁶⁹.

réinitiation, au moins dans certaines conditions¹⁴⁴. Cela semble plutôt être dû à l'identité des acides aminés qu'à la séquence de l'ARN.

Chez la levure, la réinitiation de la traduction est permise de manière dépendante d'une séquence spécifique dans l'ARNm. L'exemple le mieux caractérisé est la traduction de l'ARNm de *GCN4*, qui contient quatre uORFs (Figure 40). La première uORF permet la réinitiation de la traduction, que la cellule soit en condition normale ou de stress, et cela dépend de séquences en amont et en aval de cette première uORF. La séquence en aval est située après le codon stop et comprend environ 10 nucléotides riches en AU, et la séquence en amont s'étend sur environ 180 nucléotides et interagit avec eIF3a¹⁶⁸.

Il peut également y avoir réinitiation après la traduction d'une longue ORF, mais cela reste un événement rare qui a lieu principalement dans des ARNs viraux et qui dépend de séquences spécifiques¹⁴⁴. L'exemple le mieux caractérisé est l'ARNm bicistronique chez les calicivirus de mammifères. La première ORF code la protéine de capsid majeure (VP1) et la seconde ORF code la protéine de capsid mineure (VP2), essentielle pour la capacité d'infection. Le ratio d'expression de VP2/VP1 varie de 0,05 à 0,2 suivant le virus. Les deux séquences se chevauchent : cela peut être un chevauchement **AUGA** (codon d'initiation en rouge, codon de terminaison en italique) ou un chevauchement tel que le codon stop UAG de la première ORF se trouve plusieurs nucléotides après le codon AUG de la seconde ORF. L'événement de réinitiation nécessite la présence en amont d'une séquence TURBS (Termination Upstream Ribosomal Binding Site) constituée de plusieurs motifs (Figure 41). Le motif 1 UGGGA, qui est conservé et contenu dans la boucle d'une tige-boucle formée par la complémentarité entre les motifs 2 et 2*, est complémentaire de l'hélice h26 de l'ARNr 18S¹⁷¹. Après la dissociation de la sous-unité 60S et de l'ARNt, la sous-unité 40S est capturée par la séquence TURBS via le motif 1 complémentaire à ARNr 18S et via eIF3¹⁴⁴.

Au cours de ce chapitre, j'ai détaillé les étapes de la traduction. La synthèse protéique est un processus qui doit être fidèle et qui est fortement régulé. Dans le chapitre suivant j'aborderai des mécanismes de maintien de la fidélité et de régulation de la traduction.

III – La régulation et la fidélité de la traduction

Le contrôle de la traduction s'effectue en réponse à des signaux endogènes et exogènes, comme la disponibilité des nutriments, le stress, un changement de l'état physiologique de la cellule ou encore des hormones, et joue un rôle important particulièrement dans la maintenance de l'homéostasie, le contrôle de la prolifération cellulaire, la croissance et le développement. De nombreuses pathologies résultent d'une mauvaise régulation de la synthèse protéique¹⁷².

La régulation de la traduction peut affecter un seul ARNm, un groupe plus ou moins grand d'ARNm ou tous les ARNm¹⁷². Une régulation globale de la traduction a souvent lieu quand un ou plusieurs composants de la machinerie de traduction est activé ou inhibé tandis qu'une régulation spécifique a lieu grâce à l'action de protéines agissant en trans ou de microARNs qui se fixent à des éléments cis dans l'ARNm.

1 – L'initiation est l'étape la plus régulée

L'initiation de la traduction, et plus précisément le recrutement de la sous-unité 40S sur l'ARNm, est l'étape limitante de la traduction. Au cours de l'initiation, le recrutement de l'ARN_i^{Met} est aussi fréquemment régulé et les étapes de scanning et de reconnaissance du codon d'initiation peuvent aussi être affectées. De façon globale, la disponibilité des différents facteurs d'initiation module l'efficacité de l'initiation de la traduction. L'activité des facteurs d'initiation peut également être régulée par l'ajout de modifications post-traductionnelles comme la phosphorylation, la méthylation, l'ubiquitylation et la glycosylation mais ce sont les phosphorylations qui sont les plus caractérisées¹⁷³. L'exemple le plus étudié est la phosphorylation d'elF2 α , qui entraîne une inhibition de la liaison de l'ARN_i^{Met} au ribosome¹⁷⁴. La dérégulation de l'expression de facteurs d'initiation de la traduction est à l'origine de cancers¹⁷⁵.

Différents mécanismes permettent de réguler l'initiation de la traduction, et donc la traduction de façon plus générale. Je vais vous en présenter quelques-uns dans ce chapitre.

Une voie de régulation très importante est la voie mTORC1 (m pour mammalian). mTORC1 est un complexe de protéines qui comprend le facteur mTOR (target of rapamycin), une Ser/Thr kinase conservée de la levure à l'homme (Figure 42). Au sein du complexe mTORC1, mTOR est en association avec les protéines Raptor (regulatory associated protein of mTOR) et LST8.

L'activation de la voie de signalisation mTORC1 se fait en réponse à la disponibilité en nutriments et à des stimuli de facteurs de croissance (Figure 42). Les facteurs de croissance ou

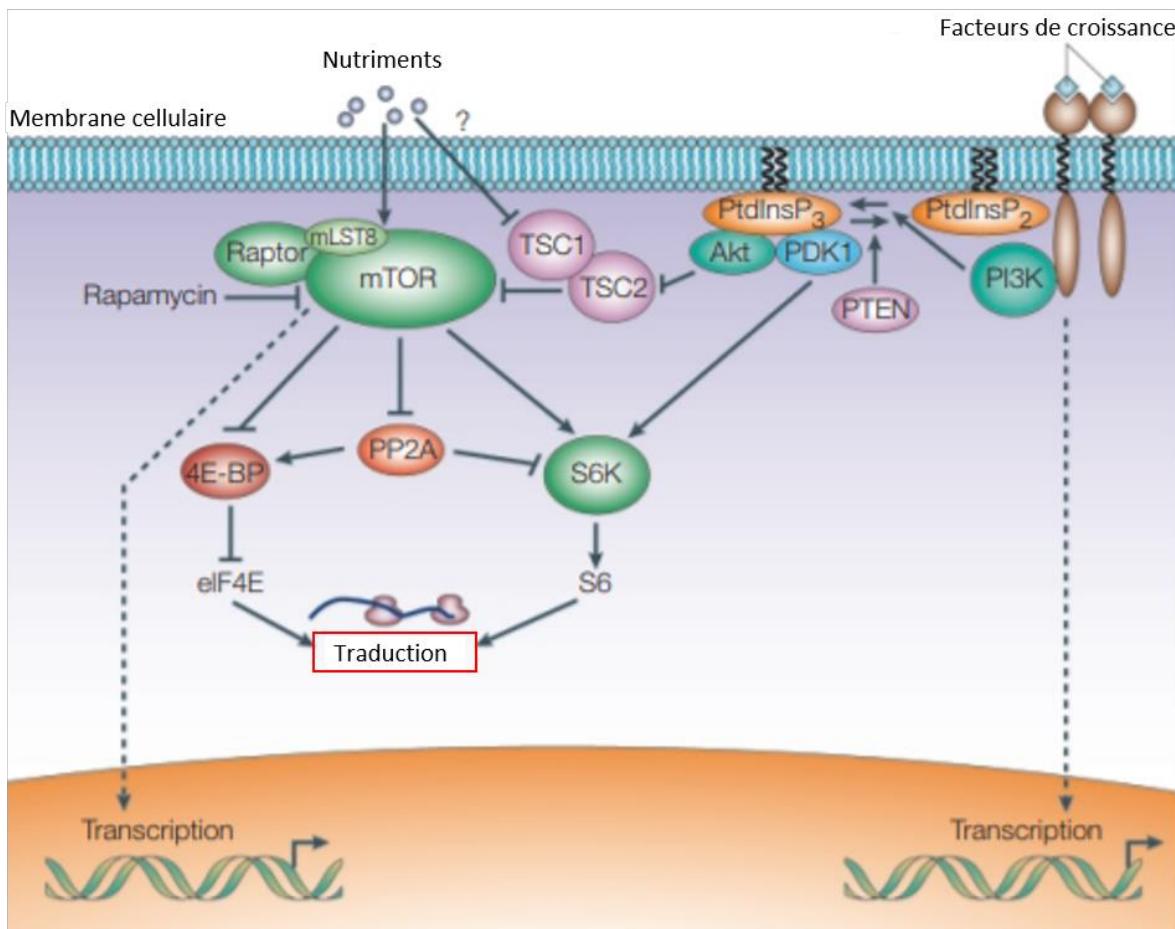


Figure 42 : Le complexe mTORC1 intègre les signaux des nutriments et des facteurs de croissance pour assurer la croissance cellulaire

Les facteurs de croissance sont intégrés par mTORC1 via PI3K, qui active Akt, qui lui-même inhibe TSC2, un inhibiteur de mTORC1. Les nutriments disponibles sont intégrés par mTORC1 directement ou induisent l'inhibition de TSC1-TSC2, levant ainsi l'inhibition sur mTORC1. Le complexe mTORC1 active alors la traduction, directement par phosphorylation de 4E-BP (un inhibiteur du facteur d'initiation eIF4E) et de S6K ou indirectement en inhibant PP2A. *Adapté de Ma, 2009*¹⁷⁶.

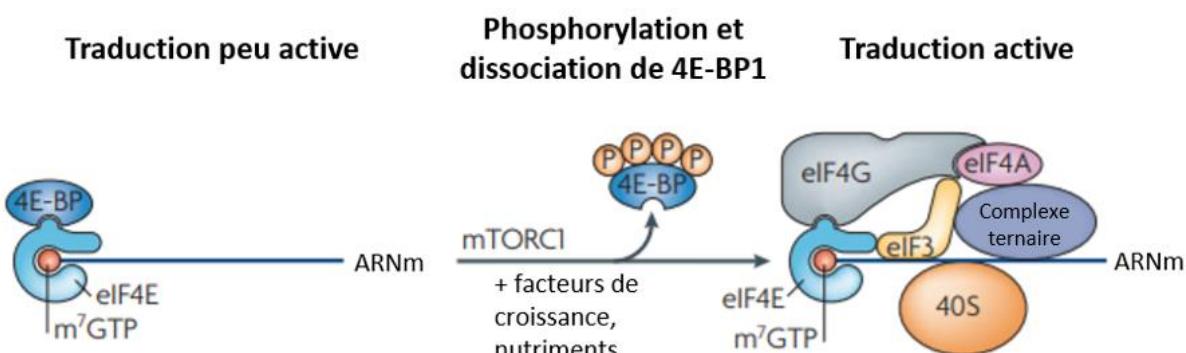


Figure 43 : Le complexe mTORC1 phosphoryle les facteurs 4E-BPs

Des événements de phosphorylation médiés par des signaux de transduction régulent l'initiation de la traduction via eIF4E. Les facteurs 4E-BP se lient à eIF4E, empêchant son interaction avec eIF4G et inhibant la traduction. Le complexe mTORC1 phosphoryle 4E-BP, ce qui le dissocie de eIF4E. eIF4E peut alors être recruté sur l'ARNm et permettre l'initiation de la traduction. *Adapté de Ma, 2009*¹⁷⁶.

certaines hormones activent des récepteurs Tyrosine kinases et des récepteurs couplés aux protéines G qui à leur tour activent plusieurs voies de signal de transduction. Parmi ces voies, les voies PI3K-AKT (phosphoinositide 3-kinase – AKT) et Ras-ERK (Extracellular signal-regulated kinase) stimulent la voie mTORC1 en inhibant les complexes suppresseurs de tumeurs TSC1-TSC2 - principalement par la phosphorylation de TSC2 par différentes kinases comme AKT ou EKR -, qui sont des régulateurs négatifs de mTORC1. De plus, en cas de baisse de l'énergie ATP disponible dans la cellule, la kinase AMPK (AMP-activated protein kinase) perçoit l'augmentation de l'AMP ou du ratio AMP/ATP dans la cellule. En réponse, AMPK phosphoryle TSC2, qui inhibe mTORC1, ou phosphoryle directement le composant Raptor de mTORC1, ce qui inhibe le complexe¹⁷⁷.

En captant la présence de facteurs de croissance et la présence suffisante de nutriments, mTORC1 régule directement la synthèse protéique¹⁷⁸ par son activité phosphotransférase dont les cibles principales sont les composants de la machinerie de la traduction, comme les facteurs d'initiation eIF4G, eIF4B et 4E-BP1¹⁷⁶.

Il existe aussi un autre complexe similaire, mTORC2, qui ne présente pas exactement les mêmes fonctions que mTORC1. Le complexe mTORC2 contient, comme mTORC1, les protéines mTOR et LST8 mais aussi Rictor et SIN1, que ne contient pas mTORC1. L'activité de mTORC2 est insensible à la quantité de nutriments disponibles mais est régulée par des facteurs de croissance, probablement par la voie PI3K. Cette régulation n'est toutefois pas complètement connue^{176,179}. mTORC2 semble avoir pour fonction principale la phosphorylation de AKT, qui active mTORC1, ainsi que d'autres kinases¹⁸⁰. mTORC2 est ainsi considéré comme agissant en amont de mTORC1. Cependant mTORC2 est aussi régulé négativement par mTORC1. mTORC2 est impliqué dans la régulation de l'organisation du cytosquelette¹⁸¹ mais aussi dans les pathologies cancéreuses¹⁷⁹.

Certains des mécanismes de régulation par mTORC1 seront abordés dans la partie III-4. Ici je vais présenter un mécanisme de régulation de l'initiation de la traduction via le facteur eIF4F (Figure 43). Le complexe eIF4F est composé de la protéine de liaison à la coiffe eIF4E, de la protéine plateforme eIF4G et de l'hélicase eIF4A, qui permet de dérouler les structures secondaires présentes dans les régions 5'UTR des ARNm. Les 4E-BPs (4E Binding Protein) sont des petites protéines qui répriment la traduction en interférant dans l'assemblage du complexe eIF4F par liaison à eIF4E¹⁸². mTORC1 est responsable de la phosphorylation de 4E-BP1, ce qui conduit à la dissociation de la 4E-BP d'eIF4E et permet ainsi l'assemblage du complexe eIF4F¹⁸². eIF4F peut alors ouvrir les structures secondaires dans le 5'UTR des ARNm et le passage de la sous-unité 43S.

En parallèle de son action sur le complexe eIF4F, mTORC1 a également une activité d'activation de l'assemblage du complexe ternaire à travers la phosphorylation de la sous-unité

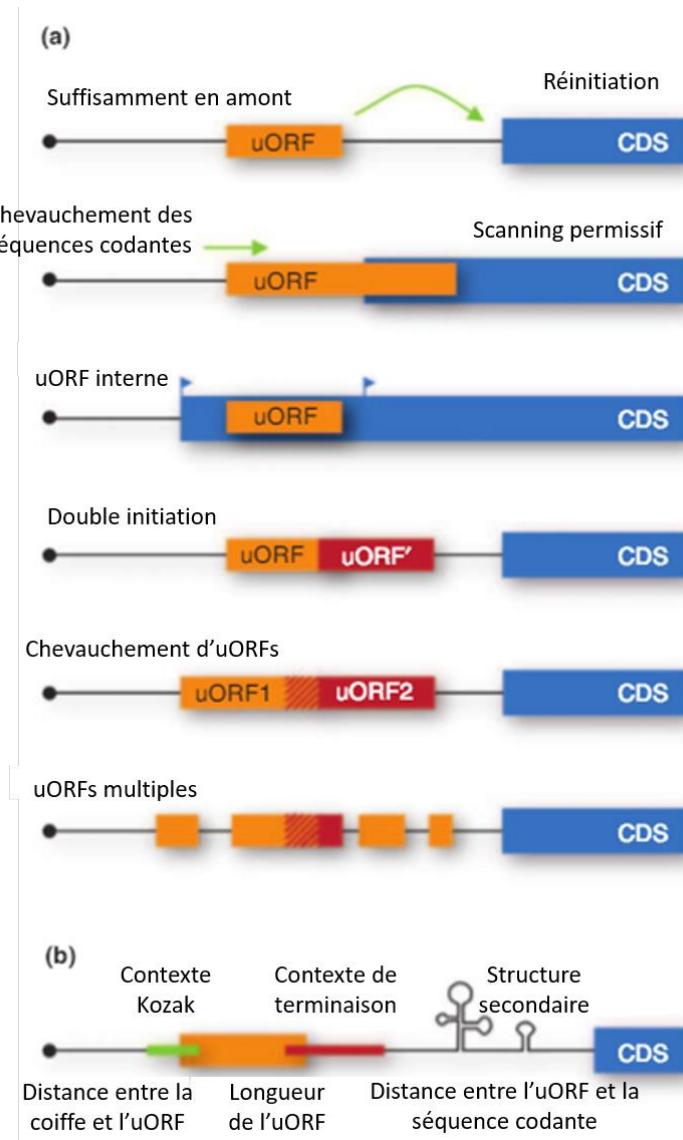


Figure 44 : Propriétés structurales régissant le potentiel de régulation des uORFs

(a) Les uORFs présentent de nombreuses variations de structures au sein des régions 5'UTR des ARNm. La coiffe est représentée par un rond noir, la séquence codante (CDS) par une barre bleue, les uORFs en orange et en rouge. Les flèches vertes représentent les possibilités de traduction de la CDS en présence d'uORF. (b) La fonction régulatrice des uORFs est affectée par différentes caractéristiques : le contexte Kozak de l'uORF (en vert), son contexte de terminaison (en rouge), sa longueur, la présence de structure secondaire entre l'uORF et la CDS et la longueur entre la coiffe et l'uORF et entre l'uORF et la CDS. *Adapté de Wethmar, 2014*¹⁸³.

elf2 β . Ainsi, en réponse à des stimuli de croissance cellulaire, mTORC1 coordonne l'assemblage du complexe d'initiation elf4F et du complexe ternaire elf2-ARNtMet-GTP¹⁸⁴.

Il a été établi que l'altération du taux de elf4E et/ou de son activité n'affectait pas la traduction globale mais la traduction d'un groupe spécifique d'ARNm « elf4E-sensibles ». Ces ARNm « elf4E-sensibles » codent majoritairement des protéines impliquées dans la prolifération et la survie cellulaire (comme les cyclines, c-Myc, VGF et EGF) et possèdent des 5'UTRs longs et fortement structurés. Leur traduction est donc fortement dépendante de l'activité hélicase de la sous-unité elf4A du complexe elf4F. A l'inverse, les ARNm « elf4E-insensibles » codent plutôt des protéines de ménage et présentent des 5'UTRs courts et peu structurés¹⁸⁵.

Dans le chapitre précédent, j'ai discuté de la présence d'uORFs dans les régions 5'UTR des ARNm. Ces uORFs sont connues pour réguler la traduction de l'ORF principale située en aval. Le plus souvent, la traduction des uORFs réprime la traduction de l'ORF par la rétention ou la dissociation de la machinerie de traduction¹⁸⁶. Mais le ribosome peut aussi passer l'uORF par balayage ou réinitier la traduction après la traduction de l'uORF^{187,188}.

Du fait de la grande diversité d'organisation des uORFs dans les régions 5'UTRs et des mécanismes de leur traduction, le contrôle traductionnel par les uORFs est loin d'être uniforme et totalement compris (Figure 44). Les propriétés de régulation des uORFs sont fortement liées à leurs positions au sein des 5'UTRs, des structures secondaires et tertiaires de l'ARNm à proximité et dans certains cas¹⁸³, de certaines caractéristiques spécifiques du peptide généré par la traduction de l'uORF¹⁸⁹ ou de la séquence de l'ARNm¹⁹⁰. Le contexte de terminaison de l'uORF ainsi que des facteurs agissant en trans peuvent également affecter les capacités de régulation de l'uORF¹⁹¹.

L'épissage alternatif et la transcription au niveau de promoteurs alternatifs peuvent être source de variations de la présence de certaines uORFs en 5' de l'ORF principale. Lorsqu'une ou plusieurs uORFs est(sont) présente(s) en 5'UTR de l'ORF principale, elle(s) a(ont) un rôle de régulation de la traduction de l'ORF. La présence d'uORFs peut être spécifique d'un ou plusieurs tissus ou encore d'une condition environnementale¹⁸³. Par exemple, l'oncogène *MDM2* a deux transcrits de tailles différentes : une forme longue et une forme courte. La forme longue possède deux uORFs qui répriment l'expression de *MDM2* d'un facteur dix. La forme courte permet une synthèse élevée de *MDM2*, qui inhibe l'apoptose médiée par p53¹⁹². *MDM2*, par son interaction avec p53, est une oncoprotéine. Ainsi, des mutations dans les deux uORFs de sa forme courte peuvent être à l'origine de la levée de la répression de la traduction et ainsi entraîner la formation d'un cancer. L'apparition de mutations dans les uORFs d'oncogènes a été retrouvée dans des tumeurs cancéreuses et pourrait être impliquée dans la carcinogenèse¹⁸⁸.

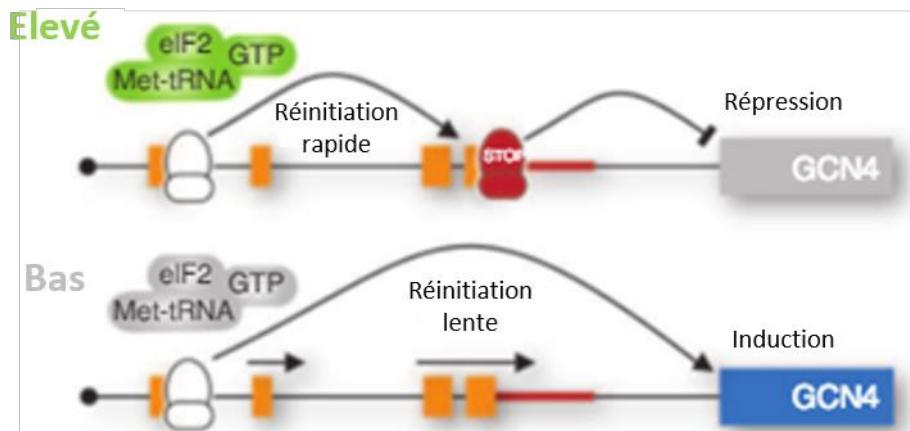


Figure 45 : La traduction de l'ORF de l'ARNm de *GCN4* est régulée par ses quatre uORFs

Lorsque l'état traductionnel de la cellule est favorable, le facteur est disponible et permet une réinitiation des ribosomes aux uORFs 2-4 après la traduction de l'uORF1, ce qui entraîne une inhibition de la traduction de l'ORF principale de *GCN4*. En cas de stress cellulaire, le facteur eIF2A est très peu disponible. Après la traduction de l'uORF1, le ribosome scanne la région 5'UTR sans initier aux uORFs 2-4 et peut alors initier la synthèse de la protéine *GCN4*. *Adapté de Wethmar, 2014*¹⁸³.

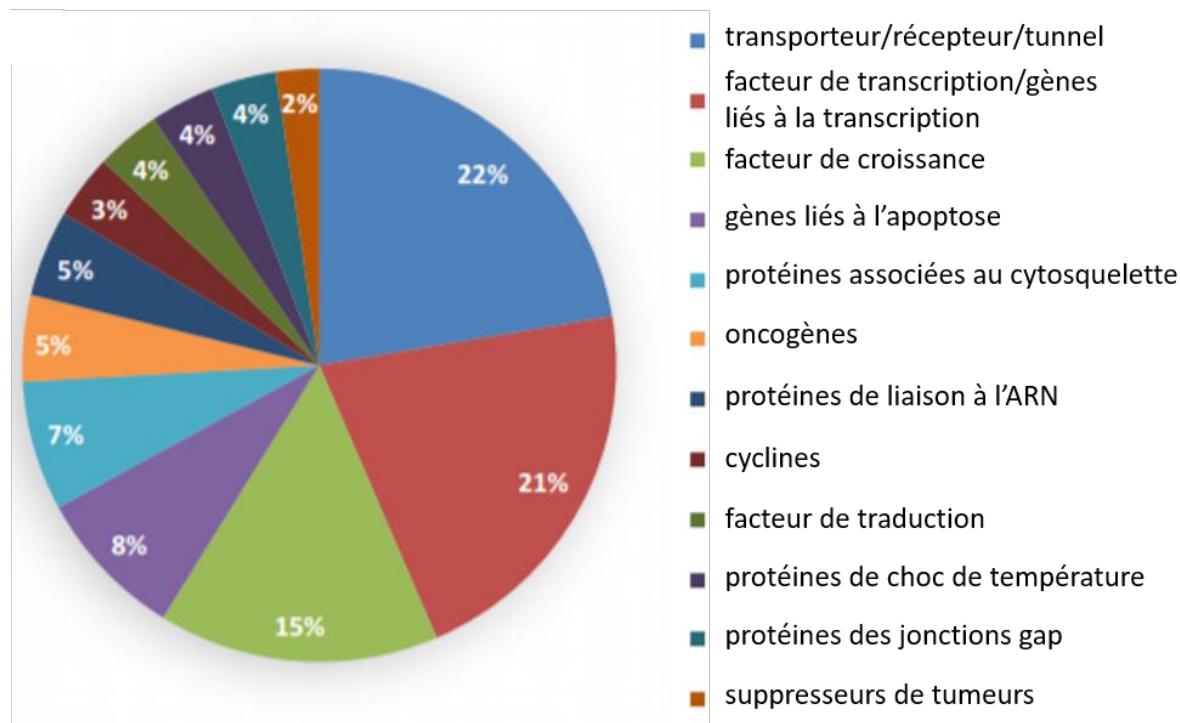


Figure 46 : Distribution des transcrits connus pour contenir des IRES suivant leurs fonctions

La plupart des transcrits contenant des IRES codent pour des transporteurs, récepteurs et tunnels, des facteurs de transcription ou des gènes liés à la transcription et des facteurs de croissance. D'autres classes fonctionnelles sont aussi représentées. Toutes les protéines correspondantes nécessitent une régulation fine de leur synthèse, dans la mesure où elles sont impliquées plus ou moins directement dans des processus capitaux de survie cellulaire et d'adaptabilité. *Adapté de Lacerda, 2017*¹³².

La régulation de la traduction par des uORFs peut aussi être liée au statut traductionnel de la cellule, c'est-à-dire la disponibilité des facteurs d'initiation de la traduction. Un exemple représentatif est la régulation de la traduction du transcrit de *GCN4* chez la levure, qui comporte quatre uORFs (Figure 45). En cas de stress cellulaire, le facteur eIF2A est phosphorylé, ce qui l'empêche de participer au complexe ternaire nécessaire à l'initiation de la traduction. Lorsque l'état traductionnel de la cellule est favorable, eIF2A n'est pas phosphorylé et permet une réinitiation des ribosomes aux uORFs 2-4 après la traduction de l'uORF1, ce qui entraîne une inhibition de la traduction de l'ORF principale. En cas de stress, eIF2A est très peu disponible. Après la traduction de l'uORF1, le ribosome scanne la région 5'UTR sans initier aux uORFs 2-4 et peut alors initier la traduction de l'ORF principale¹⁹³.

Un autre exemple est la régulation de la traduction du transcrit du gène *AMD1* (protéine adénosylméthionine décarboxylase ou AdoMetDC) chez l'homme, qui code une enzyme importante dans la biosynthèse des polyamines. Cette enzyme catalyse la décarboxylation de la S-adénosylméthionine qui fournit l'aminopropyl nécessaire à la synthèse de la spermidine et de la spermine¹⁹⁴. Son transcrit présente une uORF qui code l'hexapeptide MAGDIS, qui réprime la traduction de l'ORF principale par un mécanisme de pause du ribosome. Dans la cellule, lorsque le taux de polyamines est élevé, cela induit une interaction stable entre le ribosome en traduction et le peptide MAGDIS naissant. Le ribosome se trouve bloqué, la terminaison n'a pas lieu et la voie de dégradation NGD est mise en place. Lorsque le taux de polyamines est bas, l'interaction entre le ribosome et le peptide MAGDIS n'est pas stable et la réinitiation au niveau de l'ORF de *AMD1* est permise¹⁸⁹.

La traduction est globalement régulée négativement sous une variété de conditions cellulaires comme la mitose (division cellulaire), les chocs de température, l'hypoxie, en cas de dommages de l'ADN, de chocs osmotiques, de privation de nutriments et d'apoptose. Cette régulation négative passe notamment par une inhibition du facteur d'initiation eIF4E. Afin de permettre aux cellules de survivre, certaines protéines essentielles à la survie et à la prolifération cellulaires doivent continuer à être synthétisées. Les ARNm de ces protéines comportent des éléments IRES, dont j'ai déjà discuté dans la partie II-1-B. La traduction IRES-dépendante se met donc principalement en place en cas de stress cellulaire.

Puisque la plupart des ARNm qui possèdent des IRES codent des protéines impliquées dans la survie et la prolifération cellulaire, il n'est pas surprenant que la traduction IRES-dépendante soit impliquée dans la formation de tumeurs et la progression tumorale^{131,195} (Figure 46).

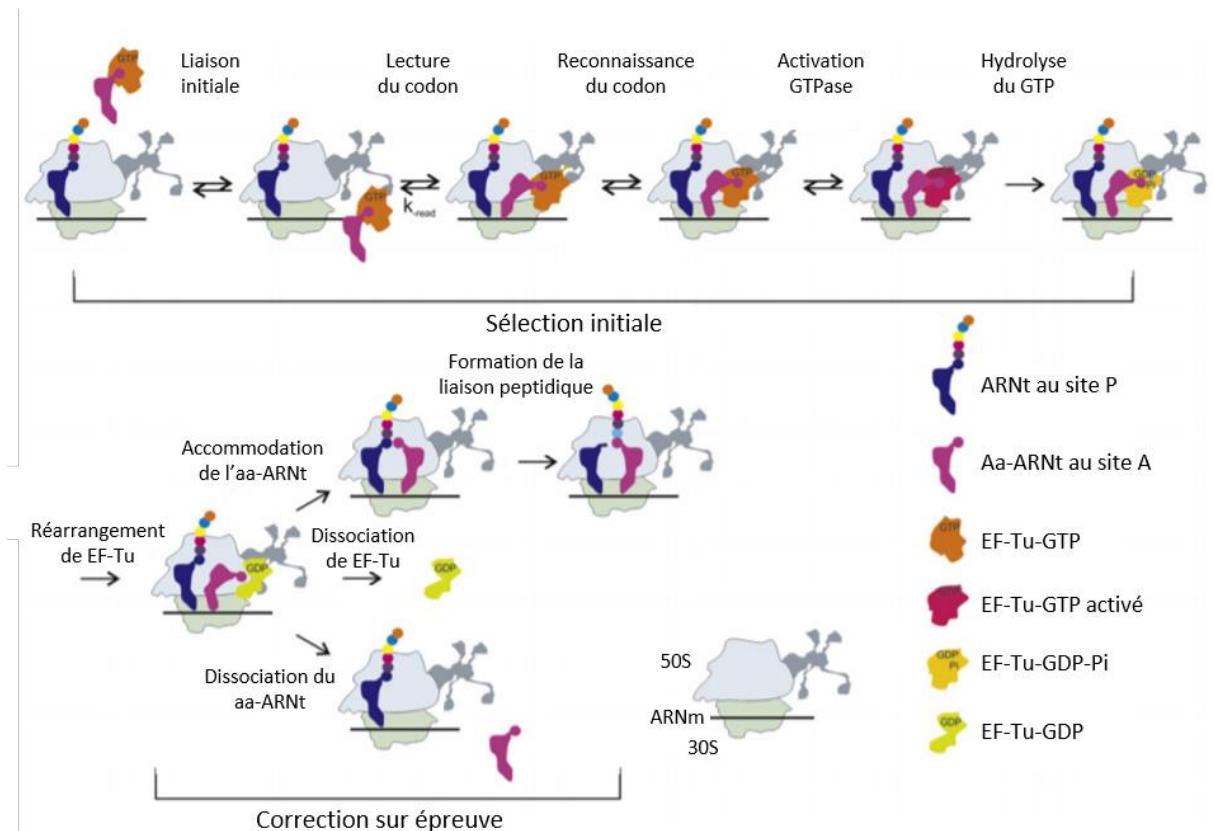


Figure 47 : Etapes de la sélection de l'ARNt chez *E. coli*

La fidélité de la traduction est assurée par deux étapes de sélection : (1) la sélection initiale permet le rejet des complexes ternaires incorrects avant l'hydrolyse du GTP, (2) l'étape de correction sur épreuve, qui a lieu après l'hydrolyse du GTP, entraîne la dissociation des aminoacyl-ARNt incorrects du ribosome, avant l'incorporation du nouvel acide aminé au sein du peptide en cours d'élargissement.

Adapté de Rodnina, 2016¹⁴³.

La régulation de la traduction peut aussi se faire par l'action de microARNs (miARNs) qui régulent l'expression génique à différents niveaux en réprimant la traduction, stimulant la déadénylation des ARNm cibles et en induisant leur dégradation prématuée¹⁹⁶. La répression de la traduction par les miARNs se fait au début de l'étape d'initiation, avant le recrutement de la sous-unité 60S, et nécessite des facteurs liés à la coiffe directement ou indirectement, comme eIF4G ou PAPBP¹⁹⁶. Les mécanismes exacts d'inhibition de la traduction ne sont pas encore complètement compris et varient selon les miARNs.

Certaines modifications au niveau de la coiffe ou de la région 5'UTR peuvent réguler l'initiation de la traduction. Par exemple, la triméthylation de la coiffe m⁷G en 2,2,7-triméthylguanosine des ARNm des sélénoprotéines (protéines qui comportent une ou plusieurs sélénocystéines) par la protéine TGS1 empêche le recrutement d'eIF4E (et donc du complexe eIF4G) à la coiffe⁵. Une autre modification dans le 5'UTR, la méthylation m⁶G, permet le recrutement du ribosome via une interaction avec eIF3, indépendamment de la coiffe et des facteurs qui y sont associés comme eIF4E¹⁹⁷.

2 – Régulation et fidélité de l'elongation

La fidélité de la traduction est assurée par deux étapes de sélection : (1) la sélection initiale permet le rejet des complexes ternaires eEF1A-GTP-aa-ARNt incorrects avant l'hydrolyse du GTP par eEF1A, (2) l'étape de correction sur épreuve, qui a lieu après l'hydrolyse du GTP, entraîne la dissociation des aminoacyl-ARNt incorrects du ribosome, avant l'incorporation du nouvel acide aminé au sein du peptide en cours d'elongation (Figure 47).

Chez la bactérie *E. coli*, il a été montré que si un acide aminé incorrect était incorporé dans la chaîne peptidique, le taux d'incorporation d'acides aminés incorrects par la suite augmentait. Ainsi le nombre de mauvais acides aminés dans le peptide se multiplie et conduit à la terminaison prématuée de la traduction et à la libération du peptide naissant. Ce mécanisme permettrait donc un contrôle qualité de la fidélité de la synthèse protéique après l'incorporation d'un acide aminé incorrect¹⁹⁸. Cependant, il n'a été décrit que chez les bactéries et il n'est pas certain qu'il existe chez les eucaryotes.

Afin d'assurer l'efficacité et la fidélité de la traduction, il est nécessaire d'avoir une balance entre la vitesse d'incorporation et une sélection correcte des acides aminés incorporés. La fidélité de la traduction requiert une discrimination des ARNt non-cognats (dont l'anticodon n'est pas

complémentaire au codon au site A) et proche-cognats (dont l'anticodon présente deux appariements avec le codon au site A) des ARNt cognats.

Des études chez la bactérie *E. coli* ont montré que lors de la sélection initiale, qui permet principalement le rejet des ARNt non-cognats mais aussi des ARNt proche-cognats, un dialogue structural s'établit entre le centre de décodage et les facteurs d'elongation. Quand un aa-ARNt se présente au centre de décodage du ribosome, l'ARN autour subit un changement conformationnel. Ce changement, appelé « fermeture du domaine », implique des interactions entre les résidus universellement conservés de l'ARNr A1492, A1493 et G530 et le petit sillon de l'hélice codon-anticodon¹⁹⁹. Ces interactions dépendent fortement des appariements Watson-Crick entre les deux premières bases du codon et les deux dernières bases de l'anticodon. L'interaction au niveau de la troisième base du codon, la base Wobble, peut supporter de ne pas être dans une conformation Watson-Crick : c'est la dégénérescence du code génétique. Les appariements codon-anticodon non-cognats et proche-cognats n'induisent pas de changement conformationnel correct, ce qui prévient l'hydrolyse du GTP et entraîne l'expulsion de l'aa-ARNt incorrect. La discrimination entre un ARNt proche-cognat et un ARNt cognat dépend également de domaines de la grande sous-unité, dont l'ARNr 23S qui participe à travers la base universellement conservée A1913²⁰⁰. Bien que ces études aient été réalisées sur des ribosomes procaryotes et n'aient pas encore été confirmées chez les eucaryotes, l'importante conservation de la structure du ribosome dans cette région et plus particulièrement des bases impliquées suggèrent fortement que ce mécanisme est similaire chez les eucaryotes.

L'expulsion des aa-ARNt non-cognats et proches-cognats peut aussi avoir lieu après l'hydrolyse du GTP, lors de l'étape de correction sur épreuve. Le facteur d'elongation eEF1A, qui a alors perdu son affinité pour l'aa-ARNt, se dissocie du ribosome. L'aa-ARNt s'accommode au site A et son bras accepteur se place correctement au niveau du PTC. Alors que les aa-ARNt cognats s'accommode rapidement et efficacement, les ARNt non-cognats et proche-cognats sont plus lents à s'accommode et les interactions ribosome-ARNt sont moins stables, ce qui entraîne leurs rejets¹⁴².

Au cours de la translocation, il est nécessaire de maintenir la phase de lecture. Une étude chez *E. coli* a montré la participation du facteur EF-G (eEF1A) à ce maintien. L'extrémité de son domaine IV s'insère dans le petit sillon de l'hélice codon-anticodon. Ce contact est maintenu durant la translocation de l'ARNt au site A vers le site P, maintenant ainsi la phase de lecture²⁰¹.

Pour permettre une translocation correcte, il est nécessaire que les mouvements qui s'effectuent au cours de cette étape soient unidirectionnels. Plusieurs éléments semblent imposer l'unidirectionnalité de la translocation²⁰². Tout d'abord, au cours de la translocation, les

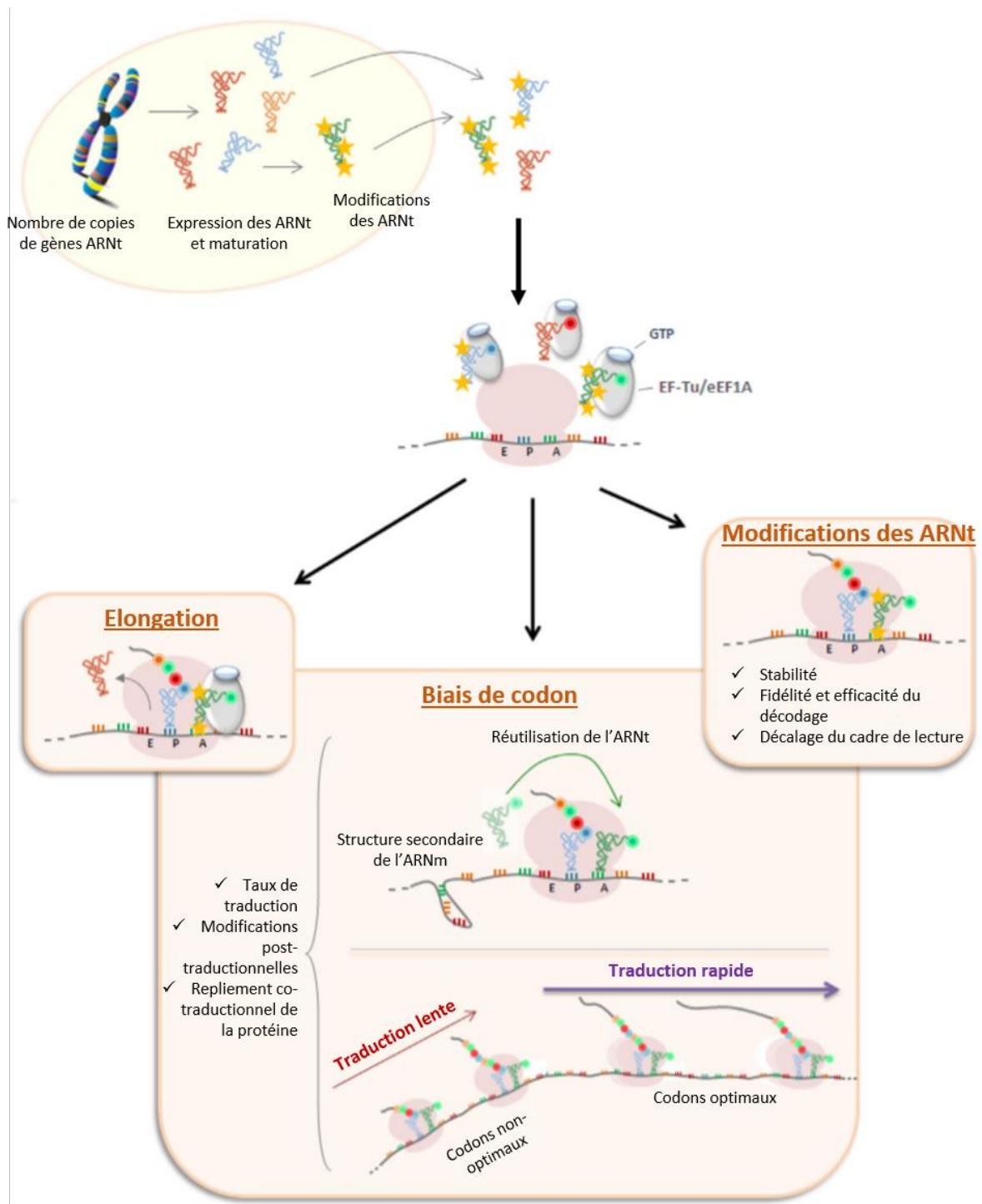


Figure 48 : Paramètres intervenant dans la fidélité et la régulation de l'elongation

La synthèse et la maturation des ARNt ainsi que le biais de codon sont les principaux paramètres qui influencent la fidélité et la régulation de l'elongation. *Adapté de Saint-Léger, 2015*¹⁴².

propriétés chimiques du bras accepteur des ARNt changent, modifiant leur spécificité de liaison pour les sites A, P et E. Le peptidyl-ARNt devient ARNt déacétylé, son bras accepteur est alors capable de se lier au site E, qui n'est capable que de se lier à un ARNt déacétylé. Le site P est alors vacant et présente une forte affinité pour le bras accepteur du peptidyl-ARNt au site A. Ces propriétés d'affinité sont à l'origine de la formation des états hybrides P/E et A/P et imposent l'unidirectionnalité du mouvement des bras accepteurs des ARNt. De plus, au sein de la sous-unité 30S l'insertion de l'extrémité du domaine IV de EF-G dans le petit sillon de l'hélice codon-anticodon rompt sa liaison au site A et joue un rôle de barrière anti-retour.

La redondance des codons ainsi que les différences de concentrations des ARNt jouent un rôle important dans l'efficacité et la fidélité de la traduction et participent finement à leur régulation. La variation dans la composition des codons peut également induire des changements dans la structure secondaire de l'ARNm et leur stabilité¹⁴² (Figure 48).

Il existe plusieurs codons qui codent le même acide aminé, appelés codons synonymes. Mais certains de ces codons sont plus fréquents dans les ARNm que d'autres. C'est ce que l'on appelle le biais de codon. Quand le biais de codon est optimal, le codon le plus représenté au sein de la séquence parmi un groupe de codons synonymes correspond à l'ARNt le plus abondant dans la cellule parmi ceux associés au même acide aminé. L'efficacité de la traduction est alors optimale. Cela permet également de réguler la vitesse d'elongation de la traduction d'un ARNm. Ainsi les gènes les plus exprimés présentent les codons les plus optimaux, et les ARNm peuvent présenter des régions de codons non-optimaux pour permettre une diminution de la vitesse d'elongation et donc un repliement correct de la protéine naissante (Figure 48).

Le ribosome a un taux d'erreur d'environ 10^{-3} à 10^{-5} (une erreur tous les 1000 à 100000 acides aminés incorporés)^{203,204}. Les erreurs de traduction les plus communes sont les mauvaises incorporations d'acides aminés mais il existe aussi d'autres erreurs : la translecture de codon stop, discutée dans le paragraphe suivant, et le décalage du cadre de lecture. Le ribosome se déplaçant de trois nucléotides en trois nucléotides, un décalage du cadre de lecture correspond à un décalage du ribosome d'un nucléotide en 5' (frameshift ou décalage de cadre de lecture -1) ou d'un nucléotide en 3' (frameshift +1) et à la poursuite de la traduction de l'ARNm dans cette nouvelle phase de lecture. L'enchaînement d'acides aminés incorporés au peptide en cours d'elongation se trouve alors totalement modifié (Figure 49a). Le décalage de cadre de lecture est une erreur de la traduction qui se produit très rarement et la plupart des événements de décalage du cadre de lecture sont des événements programmés²⁰⁵. Ils sont retrouvés pour la plupart dans les ARNs viraux (ex : la traduction de l'ARN Gag-Pol du virus HIV-1 fait intervenir un décalage du cadre de lecture -1).

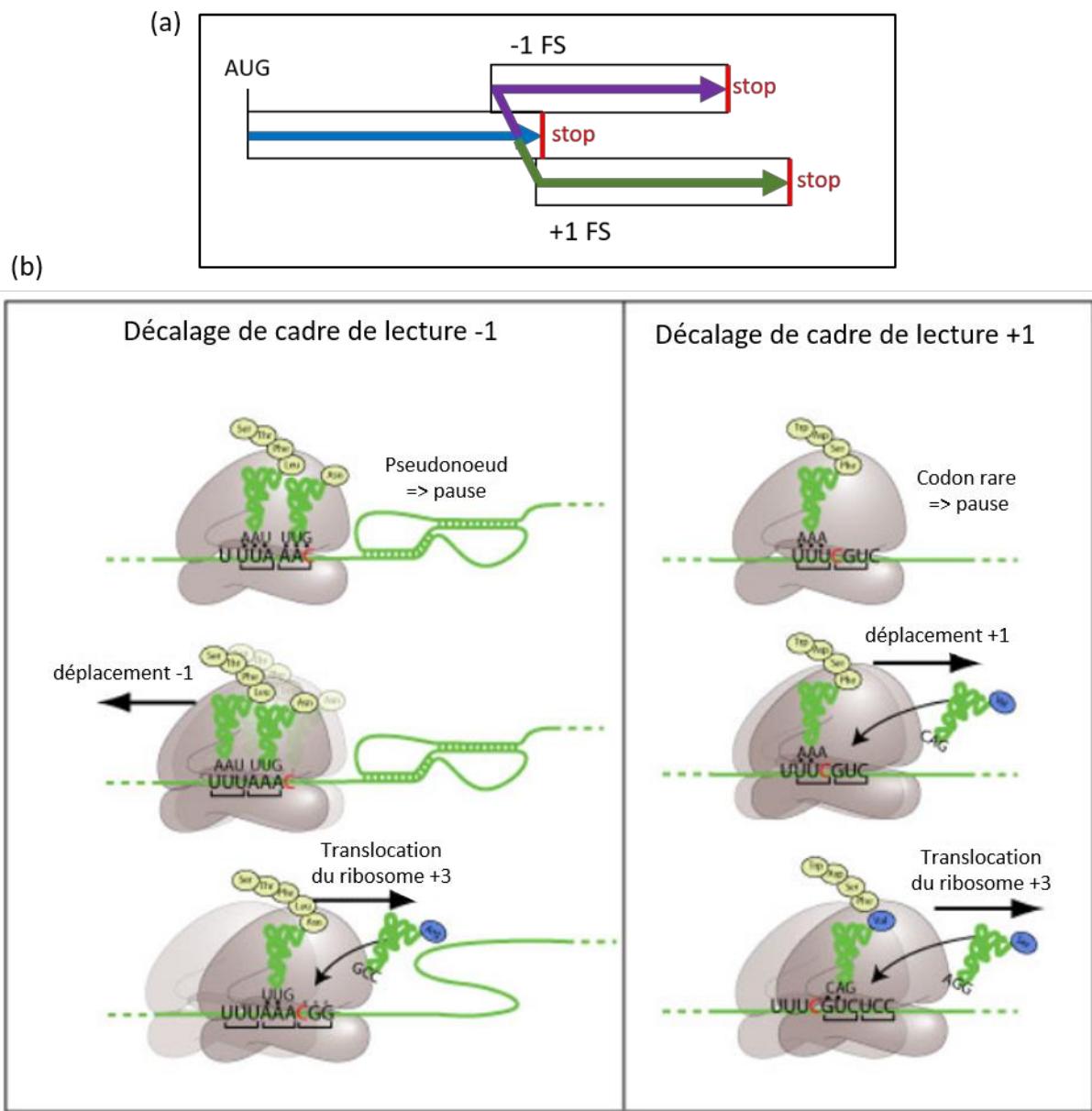


Figure 49 : Le décalage de cadre de lecture

(a) Schéma représentant le décalage de cadre de lecture (ou frameshift, FS). Le décalage du cadre de lecture correspond au décalage du ribosome d'un nucléotide en 5' ou en 3' lors de l'elongation, et à la poursuite de la lecture dans cette nouvelle phase de lecture. La chaîne peptidique alors synthétisée est complètement différente de celle correspondante à la séquence de la phase de lecture initiale. (b) Le décalage de cadre de lecture a lieu lorsque le ribosome rencontre certains éléments particuliers, comme un pseudonoeud et une séquence glissante pour le décalage en -1, ou un codon rare induisant une pause de la traduction pour le décalage en +1. Le ribosome se décale alors d'un nucléotide en amont ou en aval et reprend sa synthèse dans la nouvelle phase de lecture.

Adapté du site *ViralZone*.

Dans ces cas de décalage de cadre de lecture programmée, certains éléments sont nécessaires dans l'ARNm (Figure 49b). Les éléments typiques d'un décalage -1 sont une séquence glissante, c'est-à-dire une suite de nucléotides identiques, suivie d'une structure secondaire dans l'ARNm, comme une simple tige-boucle ou une structure plus complexe comme un pseudonoeud. Cette structure doit se trouver à 5-9 nucléotides de la séquence glissante pour être efficace²⁰⁶. Récemment, des facteurs protéiques et miARNs ayant un effet inducteur en *trans* sur le décalage de cadre de lecture ont été découverts^{207,208}. Ainsi, le taux de décalage de cadre de lecture peut être contrôlé et entraîne une variation du taux de protéine synthétisée²⁰⁹.

Tout comme pour les facteurs d'initiation, la dérégulation de l'expression des facteurs d'elongation est souvent retrouvée dans les cellules tumorales de différents types de cancer²¹⁰. De plus, l'expression des ARNt et leurs modifications ont un rôle dans la régulation de l'elongation. D'une part, l'ARN Pol III, qui synthétise les ARNt, peut être contrôlée par des facteurs impliqués dans la tumorigénèse, comme c-Myc. Une dérégulation des taux des ARNt est ainsi retrouvée dans des cellules cancéreuses²¹¹. D'autre part, les modifications des ARNt ont un rôle important dans la fidélité de la traduction et la sélection des ARNt cognats au site A et une altération de ces modifications entraîne des défauts de fidélité de la traduction²¹².

3 – Régulation et fidélité de la terminaison

La terminaison de la traduction, qui a lieu au niveau d'un codon de terminaison UAA, UAG ou UGA, est une étape très efficace. Elle présente un taux d'erreurs de moins de 0,1% chez les mammifères et 0,3% chez la levure²¹³. Néanmoins, certains facteurs peuvent réduire l'efficacité de la reconnaissance de ces codons et permettre l'incorporation d'un acide aminé. Cet événement est appelé translecture du codon stop et sera décrit dans cette partie.

❖ La translecture du codon stop :

La translecture du codon stop correspond à l'ajout d'un acide aminé au codon de terminaison par un ARNt proche-cognats (Figure 50a). A chaque terminaison de la traduction, il y a compétition entre les facteurs de terminaison et les ARNt proche-cognats au site A du ribosome. Mais l'événement de terminaison est plus efficace que l'incorporation d'un acide aminé car le facteur de terminaison eRF1 est plus affin pour le site A qui contient un codon stop que les ARNt²¹⁴.

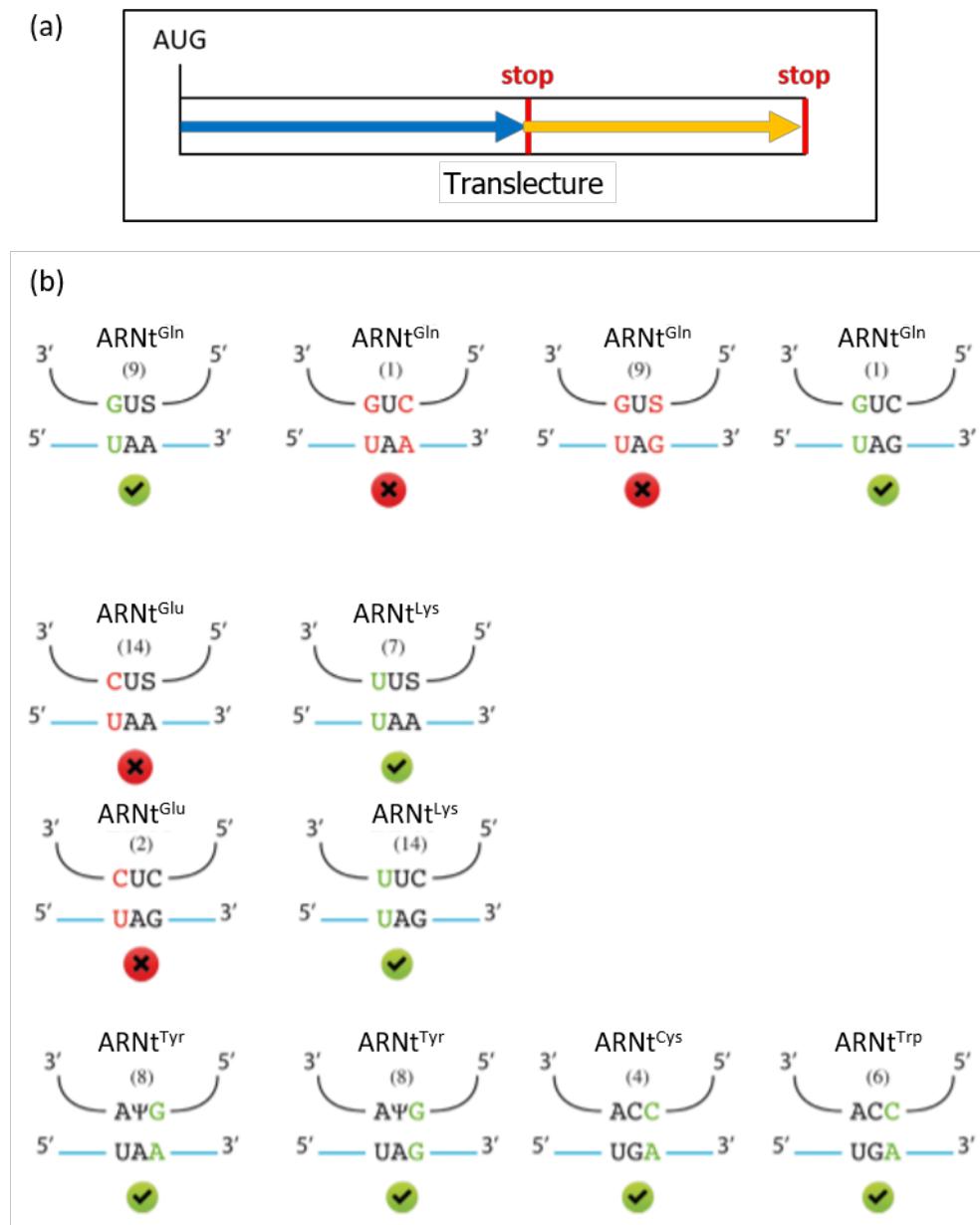


Figure 50 : La translecture du codon de terminaison

(a) Schéma de la translecture du codon stop. Le ribosome incorpore un acide aminé au codon de terminaison et poursuit la traduction jusqu'au codon stop suivant. (b) Cette figure montre les appariements possibles entre les potentiels ARNt suppresseurs et les codons stop. Pour chaque ARNt, le nombre de copies du gène est indiqué entre parenthèses. Les appariements non-canoniques sont en vert lorsqu'ils sont acceptés et en rouge lorsqu'ils ne le sont pas. La base S représente une uridine modifiée par mcm^5S^2 et Ψ correspond à une uridine isomérisée en pseudouridine. *Adapté de Blanchet, 2014*²¹⁵.

L'efficacité de la terminaison varie suivant le codon stop : le codon UAA présente la meilleure efficacité, suivi par le codon UAG puis UGA¹⁵². Il peut également varier suivant d'autres facteurs : le contexte nucléotidique en 3' et en 5', des structures secondaires dans l'ARNm - comme des pseudonoeuds -, des ARNt suppresseurs, la présence de pseudouridines et certains antibiotiques.

Tout d'abord, il a été montré que la présence de certains nucléotides en 3' du codon stop pouvait augmenter le taux de translecture. La présence d'une cytosine en position +4 (+1 étant l'uracile du codon stop) est ce qui permet l'augmentation de la translecture la plus importante et semble la plus conservée entre les espèces. Cette caractéristique a permis l'émergence de l'hypothèse d'une lecture du signal de terminaison à quatre nucléotides (le codon de terminaison et la cytosine en 3')²¹⁶, hypothèse fortement appuyée par l'observation en cryo-EM de la présence au site A, à la fois du codon de terminaison et de la position +4¹⁵⁶. Les nucléotides suivants ont aussi une influence sur l'efficacité de la terminaison, ainsi différents consensus permettant une translecture optimale ont été déterminés chez différentes espèces : chez la levure, ce consensus serait CAR NBA (avec R = A ou G, N = n'importe quelle base, B = U ou C ou G)²¹³ et chez l'homme cela serait CUA²¹⁷. L'influence de la séquence en 5' est beaucoup plus modérée, bien que la présence d'adénines en position -1 et en position -2 permettrait des taux de translecture élevés^{218,219}.

Les ARNt suppresseurs peuvent être divisés en deux catégories. Les ARNt suppresseurs naturels sont des ARNt dont les anticodons s'associent aux codons de terminaison avec un mésappariement, permettant ainsi l'incorporation d'un acide aminé dans la chaîne peptidique (Figure 50b). Ils sont à distinguer des ARNt suppresseurs non naturels qui sont des ARNt qui présentent une mutation dans leur anticodon, ce qui leur permet de lire un codon stop comme un codon sens.

L'isomérisation de l'uridine +1 des codons stop en pseudouridines conduit à l'incorporation d'ARNt au site A. Ainsi, les codons ψ AG et ψ AA sont décodés par la sérine et la thréonine et le codon ψ GA est décodé par la tyrosine et la phénylalanine. L'hypothèse est que les pseudouridines stabilisent l'interaction codon-anticodon avec les ARNt correspondants²²⁰.

Les antibiotiques sont des molécules qui, pour la plupart, ciblent le ribosome chez les bactéries. Chez les eucaryotes, du fait d'un certain nombre de différences structurales du ribosome, ces molécules perdent leurs effets inhibiteurs de la traduction mais peuvent toutefois se fixer au ribosome avec moins d'efficacité. De cette manière, ils peuvent avoir un effet mesurable sur l'efficacité de la traduction. La classe d'antibiotiques des aminoglycosides est aujourd'hui bien caractérisée comme inducteur de la translecture du codon stop chez les eucaryotes²²¹.

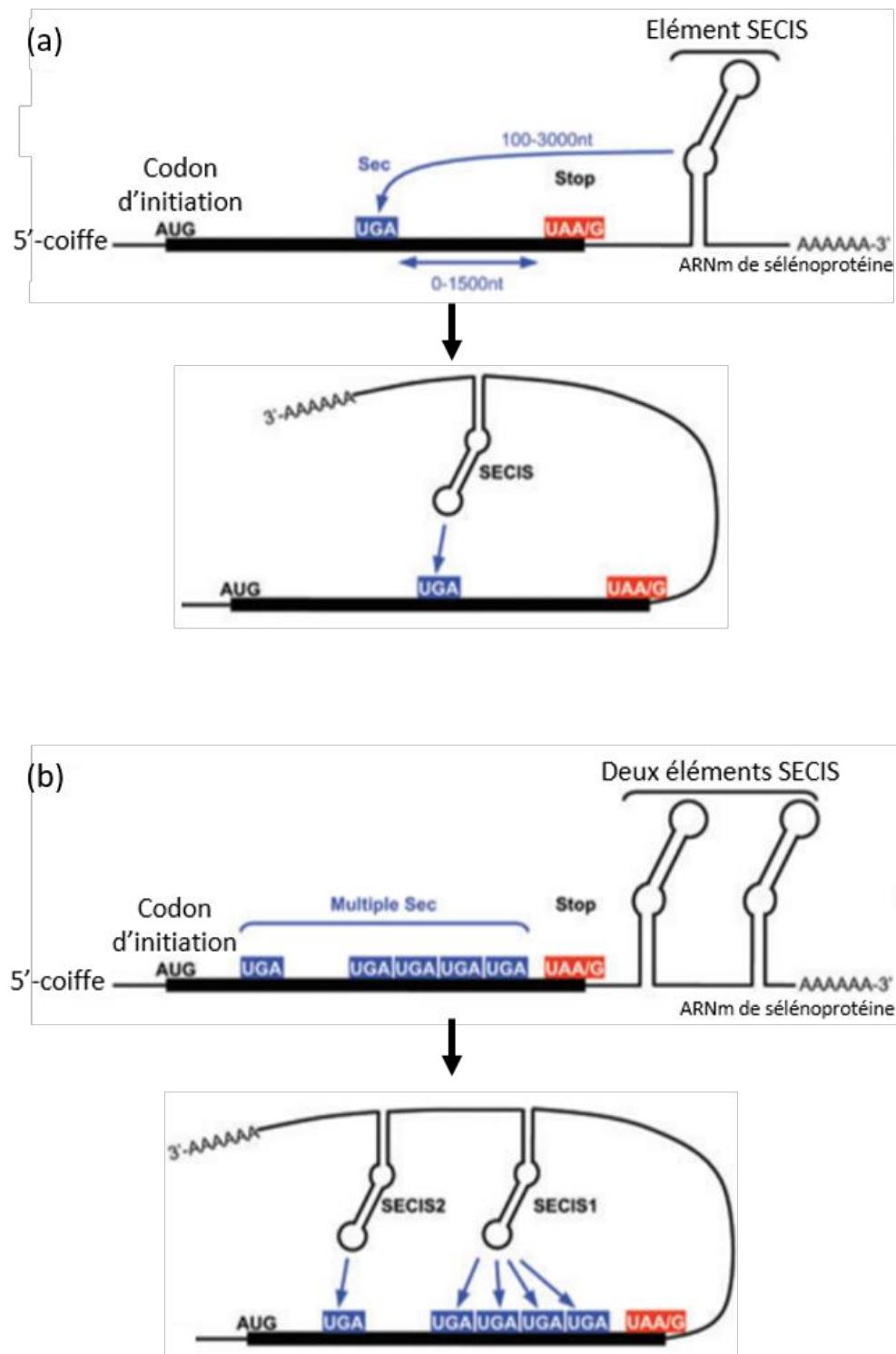


Figure 51 : Eléments SECIS et représentation schématique d'ARNm de sélénoprotéines

(a) La majorité des ARNm de sélénoprotéines contiennent un codon de terminaison UGA en phase et un élément SECIS. (b) Dans le cas des sélénoprotéines P et MsrBp de *M. senile*, de multiples codons UGA sont présents au sein de l'ARNm. Dans ce cas, le modèle proposé est que les deux éléments SECIS présents dans le 3'UTR ont différents rôles quand l'ARNm se replie : l'élément SECIS 2 contrôle l'efficacité du recodage du premier codon UGA, et l'élément SECIS 1, qui est plus efficace, est responsable du recodage des multiples codons UGA dans la partie C-terminale de la protéine. En bleu sont représentés les codons UGA qui vont être lus et en rouge sont représentés les codons de terminaison UAA et UAG. *Adapté de Bulteau, 2015*²²².

La nature des acides aminés incorporés aux codons de terminaison a été identifiée. Ainsi une étude réalisée dans l'équipe a permis de définir que, chez la levure, les codons UAA et UAG étaient décodés par la glutamine, la tyrosine ou la lysine et le codon UGA était décodé par le tryptophane, la cystéine ou l'arginine²¹⁵ (Figure 50b).

D'après l'OMS 20% des maladies génétiques humaines sont dues à la présence d'une mutation non-sens dans un gène, ce qui entraîne la production d'une protéine tronquée non fonctionnelle ou délétère. L'étude de la translecture du codon stop et des facteurs qui l'influencent a permis l'émergence de stratégies thérapeutiques fondées sur l'augmentation de la translecture des codons stop prématurés afin de restaurer la synthèse d'une protéine complète. Aujourd'hui, ces stratégies sont diverses et sont fondées notamment sur l'utilisation de certains des facteurs inducteurs de la translecture, comme les aminoglycosides ou les ARNt suppresseurs²²³.

❖ L'incorporation de sélénocystéine :

La sélénocystéine (Sec) est le 21e acide aminé naturel du code génétique. C'est un analogue de la cystéine, dont l'atome de soufre est remplacé par du sélénium, mais elle dérive en réalité de la sérine. Les modifications de la sérine en sélénocystéine s'effectuent alors que l'acide aminé sérine est accroché à un ARNt^{Ser[Sec]}. Le sélénium est un élément trace essentiel pour la plupart des organismes et est impliqué dans les fonctions musculaires, thyroïdiennes, immunes, cardiaques, reproductive ou encore cérébrales. Dans la majeure partie des cas, le sélénium est présent au sein même des protéines, sous la forme de sélénocystéines. Les protéines présentant des sélénocystéines sont appelées des sélénoprotéines.

L'ajout de sélénocystéines au sein des protéines se fait au niveau du codon UGA, qui est normalement un codon de terminaison de la traduction (Figure 51). Ce cas de translecture fait intervenir un ARNt spécifique, l'ARNt^{Ser[Sec]}, ainsi qu'une machinerie spécifique. L'ARNm dont le codon stop UGA est décodé en sélénocystéine présente une structure secondaire spécifique tige-boucle-tige-boucle nommée élément SECIS (Selenocysteine Insertion Sequence), située à au moins 50 nucléotides du codon UGA, dans la région 3'UTR chez les eucaryotes et dans la séquence codante chez les procaryotes. Chez les eucaryotes, les protéines EFSec (facteur d'elongation) et SBP2 agissent en complexe, permettent la reconnaissance de l'élément SECIS et interviennent dans l'incorporation de la sélénocystéine au sein du peptide naissant²²².

Lorsqu'un ARNm présente plusieurs codons UGA, il peut y avoir incorporation de la sélenocystéine à chacun de ces codons. La terminaison de la traduction de ces ARNm a lieu généralement au niveau des codons UAA et UAG²²².

❖ L'incorporation de la pyrrolysine :

La pyrrolysine (Pyl) est le 22e acide aminé naturel du code génétique. Elle est synthétisée à partir de la lysine et est uniquement présente chez les espèces d'archées méthanogènes *Methanosarcina* et chez les espèces de bactéries *Desulfitobacterium*²²⁴. La pyrrolysine est retrouvée dans des protéines impliquées dans la biosynthèse du méthane, comme les protéines MttBp, MtbBp et MtmBp chez les archées²²⁴. La pyrrolysine est synthétisée à partir de la lysine par une suite de réactions chimiques puis chargée sur son ARNt²²⁴. Elle est incorporée au codon UAG, qui est normalement un codon stop²²⁴. L'ARNt^{Pyl} présente une structure secondaire inhabituelle et une structure tertiaire plus courte bien que typique²²⁴ (Figure 52). Il n'y a pas d'éléments spécifiques à l'incorporation de pyrrolysine, bien que la présence d'une tige-boucle PYLIS (pyrrolysine insertion elements) en 3' du codon stop augmenterait le taux d'incorporation de pyrrolysine²²⁵. En revanche, il n'y a aucun facteur spécifique et l'elongation de la traduction fait intervenir le facteur classique²²⁴.

4 – Régulation de la traduction et croissance cellulaire

La traduction nécessite une quantité importante de matériel cellulaire et d'énergie. Pour croître et se multiplier, les cellules ont besoin de vérifier que les ressources sont suffisantes et disponibles pour assurer la synthèse protéique. Quand l'énergie ou les acides aminés diminuent, la synthèse protéique a besoin d'être régulée négativement pour que les cellules utilisent leurs ressources limitées pour survivre. Ainsi, les cellules ont mis en place des mécanismes de contrôle de la traduction qui, pour la plupart, sont sensibles à la ressource en nutriments, à l'énergie de la cellule, au stress, aux hormones et aux facteurs et stimuli de croissance¹⁷⁶.

Comme je l'ai déjà mentionné dans la partie III-1 sur la régulation de l'initiation de la traduction, le complexe mTORC1 est un important régulateur de la traduction en réponse à la disponibilité en nutriments et à des stimuli de facteurs de croissance, à travers des événements directs ou indirects de phosphorylation¹⁷⁶.

Par exemple eIF4G, composant du complexe eIF4F, est polyphosphorylé par mTORC1. Ces sites de phosphorylation sont regroupés dans une région charnière qui joint deux domaines

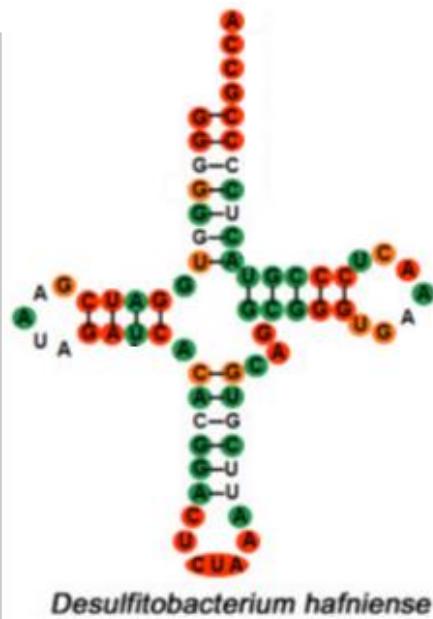


Figure 52 : Structure bidimensionnelle de l'ARNt^{Pyl} de *D. hafniense*

Les résidus colorés en rouge sont identiques au sein des ARNt^{Pyl} d'archées et de bactéries, en orange sont identiques au sein des ARNt^{Pyl} de bactéries et en vert sont conservés (plus de 70% d'identité) au sein des ARNt^{Pyl} de bactéries. *Adapté de Gaston, 2011*²²⁴.

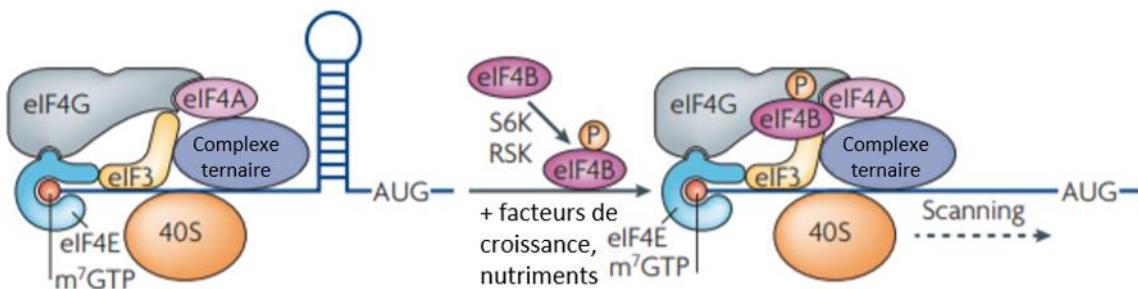


Figure 53 : mTORC1 régule l'interaction entre eIF4B et eIF4A

eIF4B est un facteur qui s'associe à eIF4A, favorisant son activité ARN hélicase. eIF4B ne s'associe à eIF4A que lorsqu'il est phosphorylé par la kinase S6K (ou RSK), permettant le déroulement des structures secondaires dans la région 5'UTR de l'ARNm. *Adapté de Ma, 2009*¹⁷⁶.

structuraux et il a été prédit que les modifications pourraient induire un changement de conformation qui affecte l'activité de la protéine²²⁶.

Les protéines S6Ks, dont S6K1 et S6K2, sont d'autres cibles importantes de mTORC1. Elles sont phosphorylées par plusieurs protéines et complexes, mais c'est le complexe mTORC1 qui initie les phosphorylations²²⁷. Les protéines S6Ks modulent l'activité des facteurs d'initiation de la traduction mais également la biogénèse du ribosome, permettant ainsi une traduction plus efficace²²⁸.

Le facteur d'initiation eIF4A voit aussi son activité fortement modulée par la voie mTORC1. eIF4A est nécessaire pour son activité ARN hélicase au niveau des structures secondaires dans les régions 5'UTR des ARNm. Il a été montré que seul, son activité est faible et qu'elle est stimulée par le cofacteur eIF4B²²⁹. eIF4B favorise l'affinité d'eIF4A pour l'ATP, ce qui augmente la processivité de l'hélicase. La protéine eIF4B peut être phosphorylée, entre autres, par des facteurs de croissance. Ainsi S6K1 phosphoryle eIF4B sur la Ser422, localisée dans la région de liaison à l'ARN nécessaire pour la stimulation de l'activité hélicase d'eIF4A. Cela augmente alors considérablement l'association eIF4B-eIF4A et donc l'activité d'eIF4A (Figure 53). De plus, la phosphorylation d'eIF4B corrèle avec la traduction d'ARNm avec des 5'UTRs longs et structurés, ce qui suggère un rôle important de la phosphorylation d'eIF4B dans la réponse aux facteurs de croissance^{176,230}. La protéine PDCD4 (Programmed Cell Death 4), qui a une fonction suppresseur de tumeurs, se lie également à eIF4A mais cela a pour effet d'inhiber son activité²³¹. eIF4G et PDCD4 présentent un motif conservé qui permet la liaison avec eIF4A, ainsi PDCD4 empêche l'incorporation d'eIF4A au sein du complexe eIF4F par compétition avec eIF4G pour se lier à eIF4A²³². En cas de stimulation par des facteurs de croissance, PDCD4 peut rapidement être phosphorylée par S6K1 sur la Ser67, ce qui entraîne sa dégradation et prévient ses effets inhibiteurs²³³.

Le cycle cellulaire est composé de plusieurs phases de croissance, G1, S et G2 au cours desquelles la cellule croît et duplique son matériel génétique, et d'une phase où la cellule se divise en deux cellules filles identiques, la mitose.

Tout au long du cycle cellulaire, le taux de synthèse protéique varie, la traduction est fortement régulée et régule également la poursuite du cycle cellulaire. Dans ce chapitre, nous allons voir quelques exemples de régulation de la traduction au moment des étapes clés du cycle cellulaire.

❖ Transition G1-S :

Chez *S. cerevisiae*, des mutants des facteurs d'initiation Cdc33p (homologue d'eIF4E) et Prt1p (homologue d'eIF3η) présentent des défauts de transition de la phase G1 en phase S. Cdc33 est requise en phase G1 pour stimuler la synthèse de Cln3p, une protéine très instable qui présente un

taux faible de traduction. Cln3p est une cycline de la phase G1 qui couple la croissance cellulaire à la progression dans le cycle. L'ARNm de *CLN3* présente une uORF dans sa région 5'UTR. En cas de croissance faible, la traduction de cette uORF retient les complexes de pré-initiation 43S et empêche la synthèse de la protéine. A l'inverse, en milieu riche, le nombre de complexes de pré-initiation augmente, ce qui permet le passage de l'uORF de certains complexes et donc une synthèse de la protéine Cln3p. La protéine permet alors l'entrée en phase S^{234,235}.

Dans les cellules HeLa, la transition G1-S est facilitée par DDX3, une ARN hélicase membre de la famille DEAD-box. L'inhibition de cette protéine entraîne une pause en phase G1. Cette pause peut être levée par l'expression de la cycline E1, ce qui suggère fortement que l'ARNm de E1 serait la cible principale de DDX3²³⁶. La région 5'UTR de l'ARNm de E1 est riche en GC et forme très probablement une structure secondaire stable. Ainsi, la traduction de cet ARNm nécessite un déroulement de la structure secondaire présente dans le 5'UTR, probablement par l'ARN hélicase DDX3^{236,237}.

❖ Transition G2-M et progression mitotique :

Chez *S. pombe*, Cdc25p, une phosphatase qui active Cdc2, est nécessaire pour le passage en mitose. La région 5'UTR de l'ARNm de *CDC25* contient plusieurs uORF et tige-boucles. Sa traduction nécessite l'hélicase eIF4A et est modulée en fonction de sa disponibilité²³⁸.

L'exemple de Cdc25p ainsi que celui de Cln3p illustrent le fait que certains transcrits de cyclines qui interviennent tout au long du cycle cellulaire présentent des uORFs. La régulation de la traduction de ces transcrits à travers des uORFs est ainsi un moyen de coordonner la poursuite du cycle cellulaire avec les signaux de croissance extérieurs²³⁹.

Au cours de la mitose, la synthèse d'eIF4E est fortement inhibée. Cependant, certains ARNm sont toujours traduits. Ces ARNm possèdent des IRES, ce qui permet le recrutement du complexe de pré-initiation en l'absence d'eIF4E. Ce sont en grande partie des ARNm nécessaires à la progression mitotique, comme les ARNm des protéines ornithine décarboxylase, c-Myc et Cdk11/p58PITSRE²⁴⁰.

5 – Voies de surveillance NGD, NSD et NMD

La production de protéines aberrantes peut être délétère pour la cellule. Ces protéines aberrantes peuvent résulter d'une mauvaise conformation ou de la traduction d'un ARNm aberrant. La production d'ARNm aberrant peut être la conséquence de mutations dans la séquence d'ADN ou avoir lieu au moment de sa transcription ou au cours des étapes de maturation. Dans ce cas, il existe des machineries de contrôle qualité des ARNm qui dégradent les messagers aberrants. Il existe trois

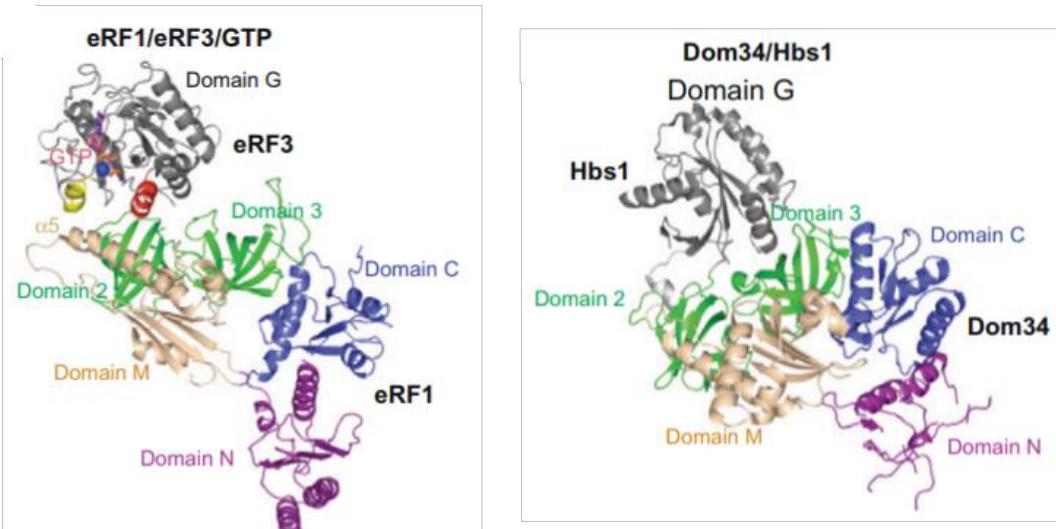


Figure 54 : Comparaison des structures d'eRF1-eRF3-GTP et Dom34-Hbs1

Dom34p est un homologue d'eRF1 et Hbs1p est un homologue d'eRF3. Ainsi, les complexes eRF1-eRF3-GTP et Dom34-Hbs1 sont structurellement similaires. *Adapté de Jackson, 2012*¹⁴⁴.

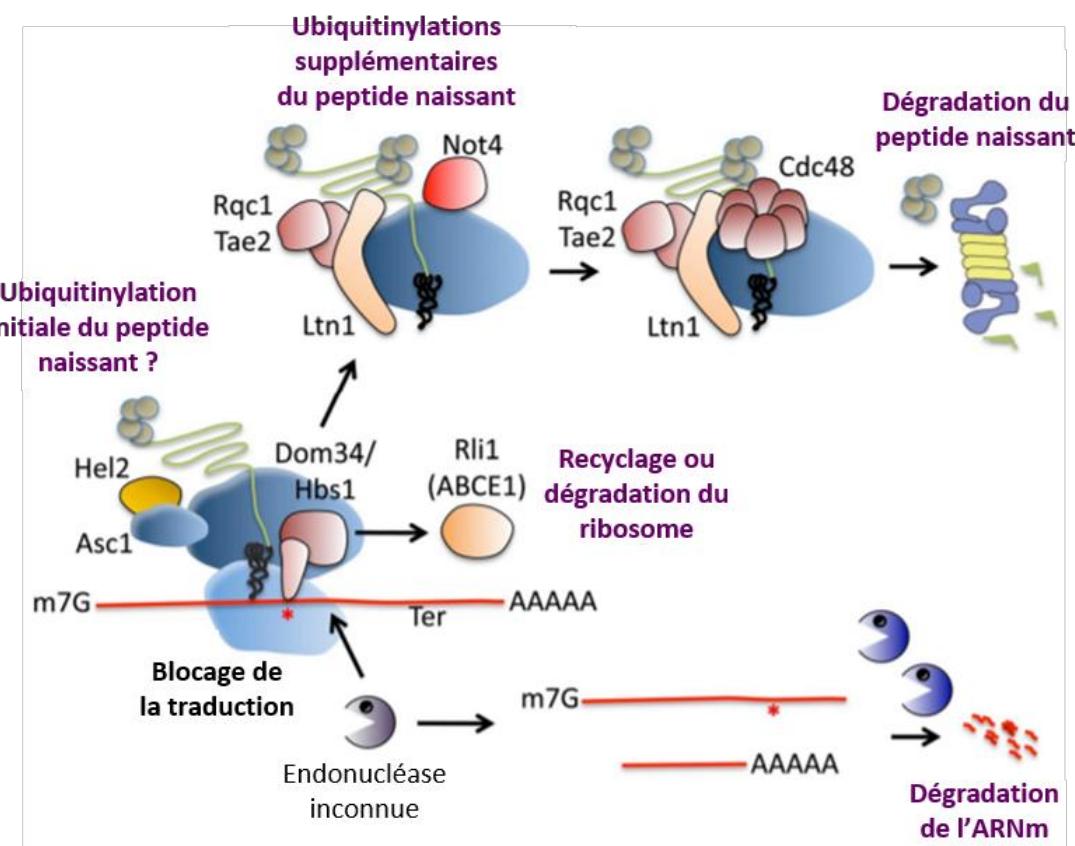


Figure 55 : Acteurs de la voie No-Go Decay (NGD)

Les blocages du ribosome au cours de la traduction sont pris en charge par des facteurs dédiés à la dégradation des ARNm et protéines aberrantes. Le clivage initial de l'ARNm est réalisé par une endonucléase. Puis les facteurs Dom34p et Hbs1p se lient au site A du ribosome, ce qui permet son recyclage assisté par Rli1p (ABCE1). L'ARNm est ensuite dégradé par des exonucléases. Le peptide naissant est ubiquitinylé par Not4p et le complexe RQC comprenant Ltn1p. *Adapté de Lykke-Andersen, 2014*²⁴¹.

voies de contrôle qualité des ARNm. La voie NGD (No-Go Decay) permet le contrôle des ARNm sur lesquels les ribosomes se retrouvent bloqués, en pause prolongée. La voie NSD (Non-Stop Decay) permet le contrôle des ARNm qui ne présentent pas de codon stop. La voie NMD (Nonsense-Mediated Decay) permet le contrôle des ARNm présentant un stop prématuré et qui sont responsables de la synthèse d'une protéine tronquée. Je vais vous détailler ces trois voies de contrôle et de dégradation des ARNm aberrants.

❖ La voie NGD :

La voie No-Go Decay prend en charge les ARNm sur lesquels les ribosomes se trouvent bloqués. Ce blocage des ribosomes peut être dû à différents facteurs : une suite d'acides aminés lysines et/ou arginines dans le peptide naissant qui interagissent avec le tunnel de sortie, bloquant le ribosome ; une suite de prolines ; une structure secondaire dans l'ARNm ; la dépurination de l'ARNm ; un ARNt non chargé ; un codon non-optimal ou encore un ARNm tronqué²⁴².

La voie NGD débute par le clivage endonucléolytique et la dégradation de l'ARNm en 5' du ribosome bloqué²⁴¹.

Chez les eucaryotes, les ribosomes en pause sont reconnus par des protéines similaires aux facteurs de terminaison appelées Dom34p (Pelota chez les mammifères) et Hbs1p. Dom34p est un homologue d'eRF1 qui ne présente pas de motif de reconnaissance du codon stop ni de motif catalytique GGQ pour faciliter l'hydrolyse du peptide. Hbs1p est un homologue de la GTPase eRF3 (Figure 54). Le complexe Dom34:Hbs1 est recruté au niveau du site A du ribosome bloqué et favorise la dissociation des sous-unités de manière codon-indépendante²⁴³. Lorsque le complexe Dom34:Hbs1 est recruté, Hbs1p est liée à une molécule de GTP. Après le clivage endonucléolytique de l'ARNm par une endonucléase encore inconnue, stimulé par Dom34:Hbs1²⁴¹, le GTP est hydrolysé, Hbs1p est libérée et Dom34p change de conformation, ce qui conduit à la dissociation des sous-unités du ribosome. La protéine Rli1p (ABCE1 chez les mammifères) intervient aussi dans cette dissociation. Le complexe Dom34:Hbs1 stimule également la dégradation de l'ARNm aberrant par l'exosome (Figure 55).

En parallèle, la protéine naissante est reconnue par le complexe RQC (Ribosome-bound Quality Control)²⁴⁴ et ubiquitinylée par les ubiquitines ligases Not4p et Ltn1p (Ltn1p fait partie du RQC). L'ajout d'ubiquitines est un signal d'adressage au protéasome, où les protéines aberrantes sont dégradées. Il semblerait que la protéine RACK1 (Asc1p chez la levure) ait également un rôle dans la dégradation de la protéine aberrante mais ce rôle n'est pas encore défini²⁴¹ (Figure 55).

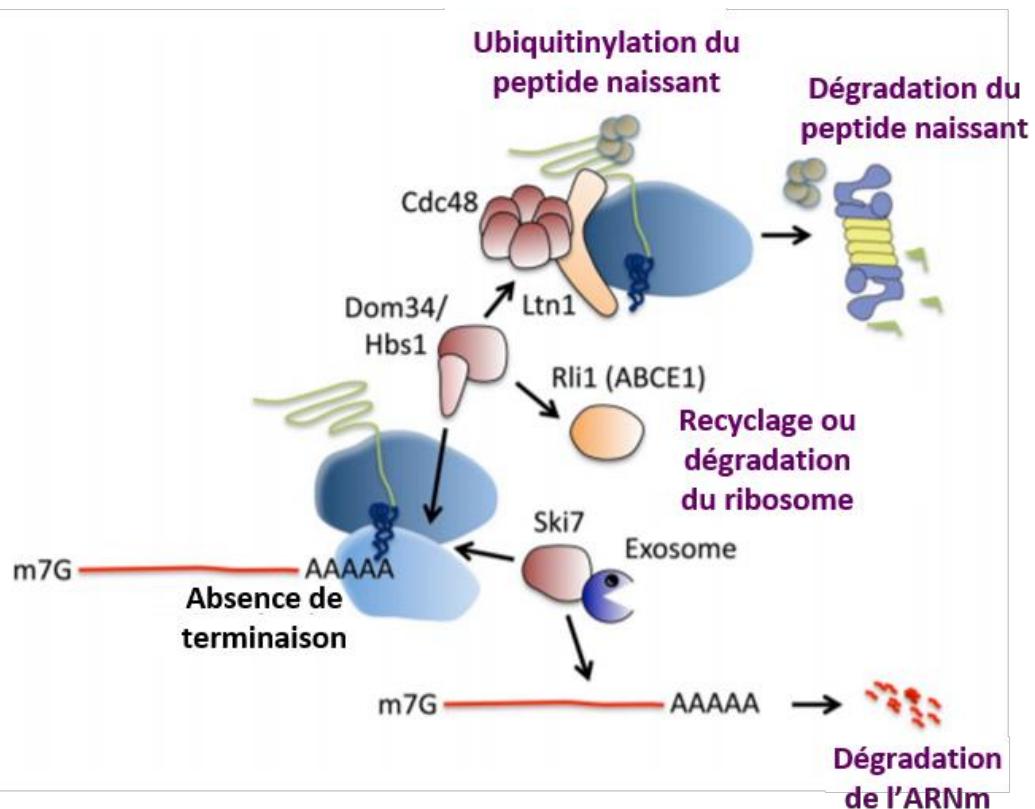


Figure 56 : Acteurs de la voie Non-Stop Decay (NSD)

Les ARNm sur lesquels les ribosomes en traduction atteignent l'extrémité 3'UTR sans rencontrer un codon de terminaison sont, soit détectés par le facteur Ski7p, homologue d'eRF3, et marqués par celui-ci pour être dégradés par l'exosome, soit reconnus par le complexe Dom34-Hbs1 et recyclés par l'intermédiaire de Rli1p (ABCE1). Les peptides naissants sont ubiquitinylés par un complexe contenant Ltn1p, et dégradés. *Adapté de Lykke-Andersen, 2014*²⁴¹.

❖ La voie NSD :

La voie Non-Stop Decay prend en charge les ARNm qui ne présentent pas de codon de terminaison de la traduction. Parfois, l'ajout de la queue poly(A) ne s'effectue pas au bon endroit, ce qui conduit à la production d'un ARNm aberrant sans codon de terminaison. Certains ARNm sans codon stop résultent aussi d'un clivage endonucléolytique effectué durant la voie NMD (voir paragraphe suivant). Lorsque le ribosome traduit un ARNm « sans stop », il ne s'arrête pas à la fin de la séquence codante et traduit la queue poly(A). Les polylysines qui sont alors ajoutées au peptide naissant interagissent avec le tunnel de sortie, bloquant ainsi le ribosome²⁴¹.

Ceci aboutit à l'induction de la voie NSD (No-Stop Decay), ce qui conduit à la dégradation de l'ARNm aberrant et de la protéine tronquée^{241,245}. Dans la plupart des organismes, ce sont les facteurs Dom34p et Hbs1p qui, de façon similaire à la voie NGD, reconnaissent les ribosomes sur un transcrit sans stop et qui sont responsables de sa dissociation avec l'ATPase Rli1p (ABCE1)²⁴⁶. Mais la voie NSD ayant principalement été étudiée chez la levure, il a été révélé dans un petit nombre de levures le rôle très important de la GTPase Ski7p, un homologue du facteur de terminaison eRF3 et de Hbs1p. En plus du domaine C-terminal GTPase homologue à ceux d'eRF3 et de Hbs1p, Ski7p présente un domaine N-terminal important pour l'activité catalytique exonucléase 3' vers 5' de la forme cytoplasmique du complexe exosome²⁴⁷. L'hypothèse est que Ski7p interagit avec le ribosome bloqué et recrute l'exosome, via le complexe Ski, afin d'activer la dégradation de l'ARNm aberrant^{247,248}. Cependant, le site catalytique de Ski7p présente certains résidus qui ne sont pas conservés alors qu'ils le sont dans toutes les autres GTPases. Plusieurs hypothèses ont alors été émises : 1) ces différences n'empêchent pas l'activité GTPase de Ski7p, et Ski7p nécessite, comme eRF3 et Hbs1p, un cofacteur qui active l'hydrolyse du GTP et qui ne serait pas Dom34p ; 2) Ski7p a une activité GTPase sans cofacteur ; 3) Ski7p ne présente pas d'activité d'hydrolyse du GTP et agit suivant un mécanisme différent pour permettre la dissociation du ribosome²⁴⁹. Il semblerait également que dès la mise en place de la voie NSD, la traduction de l'ARNm aberrant soit inhibée²⁴⁸ (Figure 56).

❖ La voie NMD :

L'apparition d'un codon stop prématûr (PTC) dans l'ARNm conduit à la production d'une protéine tronquée et à l'activation de la voie Nonsense-Mediated Decay²⁵⁰.

La terminaison prématûrée de la traduction peut être reconnue par l'absence d'une région 3'UTR normale juste en aval du codon de terminaison. Certains facteurs associés à la région 3' de l'ARNm, comme la PABPC (PABP cytoplasmique), permettraient une terminaison correcte et

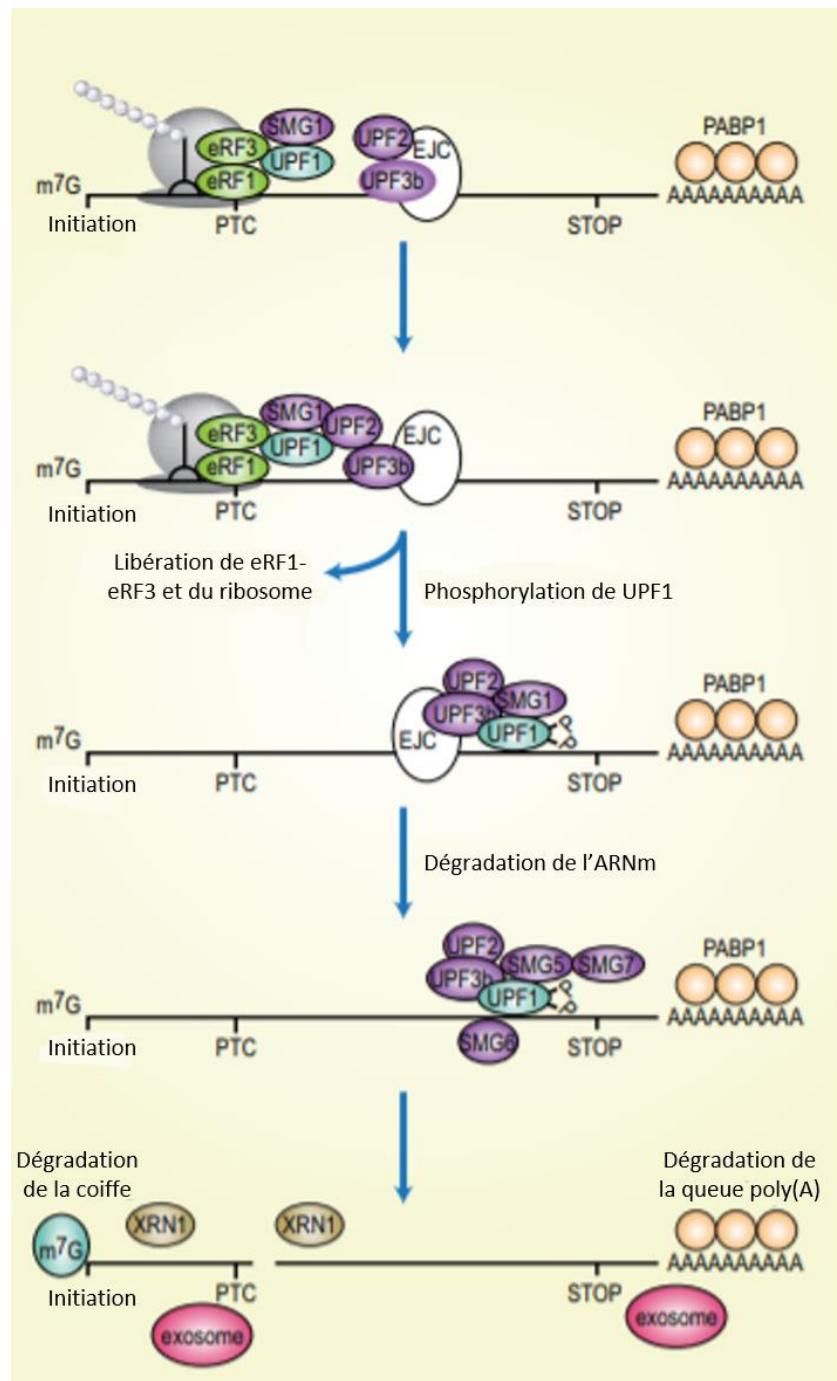


Figure 57 : Mécanisme de la voie Nonsense-Mediated Decay (NMD)

Le facteur UPF1 et la kinase SMG1 se lient aux facteurs de terminaison eRF1 et eRF3 au niveau d'un codon stop prématûr (PTC). Les protéines UPF2 et UPF3b sont liées au complexe EJC (Exon Junction Complex). Les deux groupes de protéines se joignent, ce qui permet la phosphorylation d'UPF1 et la libération d'eRF1-eRF3 et du ribosome. Les protéines SMG5, SMG6 et SMG7 sont recrutées, ainsi que des facteurs de dégradation de l'ARNm comme XRN1 et l'exosome. *Adapté de Hug, 2016*²⁵¹.

antagoniseraient la voie NMD. Si la région 3'UTR est trop éloignée du codon où se termine la traduction, l'effet antagoniste serait inefficace^{241,252,253}. Chez les métazoaires, la voie NMD est principalement stimulée quand la terminaison a lieu avant la dernière jonction exon-exon. Au niveau des jonctions exon-exon, des complexes appelés EJC (Exon-Junction Complex) sont déposés durant l'épissage des ARNm. Lors de la traduction de l'ARNm par le ribosome, ces complexes sont enlevés de l'ARNm par le passage du ribosome. En cas de terminaison prématuée de la traduction, tous les complexes EJC ne sont pas retirés et le(s) complexe(s) restant(s) est(sont) reconnu(s) par les facteurs UPF de la voie NMD²⁴¹. Chez les mammifères, la voie est plutôt bien caractérisée. Tout d'abord, UPF1, le facteur central de la voie NMD, et SMG1 s'associent aux facteurs de terminaison eRF1 et eRF3, au niveau du ribosome en cours de terminaison sur un PTC. En parallèle, UPF2 et UPF3b se lient au complexe EJC immédiatement en 3' du PTC. Ces deux complexes de protéines se lient entre eux, conduisant à la libération des facteurs eRF1 et eRF3 et à la phosphorylation de UPF1. UPF1 sous sa forme phosphorylée permet alors le recrutement de SMG5, SMG6 (une endonucléase spécifique de la voie NMD) et SMG7 qui recrutent à leur tour des facteurs qui vont dégrader l'ARNm²⁵¹ (Figure 57). La dissociation de l'ARNm du ribosome dépend de l'activité ATPase de UPF1. La dégradation de l'ARNm débute par un clivage endonucléolytique et se poursuit par l'enlèvement de la coiffe ou de la queue poly(A) suivant l'ARNm et l'organisme²⁴¹.

La voie NMD est une voie de contrôle qualité des ARNm mais également de régulation de l'expression des gènes. En effet, plusieurs caractéristiques des ARNm peuvent être des marqueurs de la dégradation de ces ARNm par la NMD. Ainsi, la NMD régule le taux de transcrits qui présentent des uORFs²⁵⁴, un PTC apparu par épissage alternatif^{255,256} ou décalage de cadre de lecture^{208,257}, une longue région 3'UTR²⁵⁸⁻²⁶⁰, de transcrits de sélénoprotéine en cas d'insuffisance de sélénium et si le codon stop UGA est situé en amont d'une jonction exon-exon²⁶¹ ou encore de long non-coding RNAs (lncRNAs)²⁶². Par ces mécanismes, la NMD a un rôle dans de nombreux processus physiologiques, dont le développement embryonnaire^{263,264} et la différenciation en différents types cellulaires, ainsi que dans des processus pathologiques, comme la β-thalassémie²⁶⁵.

6 – Hétérogénéité des ribosomes et traduction

Les ribosomes peuvent voir leur composition en protéines ribosomiques varier, ainsi que la composition en ARNs ribosomiques. L'hétérogénéité de la composition en protéines peut résulter d'une variation de la stoechiométrie des protéines au sein du ribosome²⁶⁶ ou encore de

changements dans les modifications post-traductionnelles des protéines ribosomiques. L'hétérogénéité des ARNr peut provenir de variations de la séquence primaire ou des modifications post-transcriptionnelles²⁶⁷. Les ARNr sont fortement modifiés, en particulier à des positions présentant un rôle fonctionnel, principalement par des 2'-O-méthylations et des pseudouridylations. Cependant, tous les sites modifiés d'un ARNr ne le sont pas forcément dans tous les ribosomes. Les ribosomes varient donc dans leur composition en modifications chimiques des ARNr. Toutes ces variabilités peuvent être une source pour la spécialisation des ribosomes, c'est-à-dire la traduction spécifique de groupes d'ARNm en fonction de la composition des ribosomes²⁶⁸.

A – Hétérogénéité dans la composition des protéines

L'hétérogénéité de la composition des protéines ribosomiques (PR) au sein des ribosomes est établie depuis plus de 40 ans mais était peu étudié du fait de la limitation des techniques disponibles²⁶⁷. Avant cela, il était plutôt établi que les protéines ribosomiques étaient produites de manière coordonnée afin de s'assurer que chaque ribosome contenait tous ces composants, nécessaires pour une synthèse protéique correcte. Cette hypothèse a été remise en cause par plusieurs observations où la stoechiométrie des protéines ribosomiques variait suivant les conditions ou le type cellulaire^{269,270}. Dans une étude de 1981 réalisée chez *Dictyostelium discoideum*, il a été montré que les protéines ribosomiques qui composaient les ribosomes étaient différentes entre des cellules à l'état végétatif et des spores²⁷⁰.

Plus récemment, des variations de synthèse de plusieurs PR ont été rapportées dans six tissus humains différents et plusieurs régions de l'embryon de souris^{79,271}. Il a également été montré que dans des cellules souches embryonnaires de souris et des cellules de levure, la stoechiométrie des PR variait au sein des ribosomes selon les conditions de culture ainsi que selon le nombre de ribosomes en traduction sur un même ARNm. Ainsi, dans des cellules souches embryonnaires de souris, les PR uS3, eS4x (RPS4x), uS4 et uL18 sont enrichies dans les monosomes et les PR uS13, uL15 et eL30 sont enrichies dans les polysomes ; et dans des cellules de levure, la PR uL35b est enrichie dans les monosomes et la PR uL24a est enrichie dans les polysomes²⁶⁶.

La variabilité des protéines ribosomiques peut être à l'origine de spécificités de traduction. Une étude a montré que la présence de la protéine uL16 dans les ribosomes de cellules de mammifères permettait un enrichissement de la traduction des ARNm de certaines fonctions cellulaires (matrice extracellulaire, stress cellulaire par ex), et que la plupart de ces ARNm possédaient des IRES²⁷².

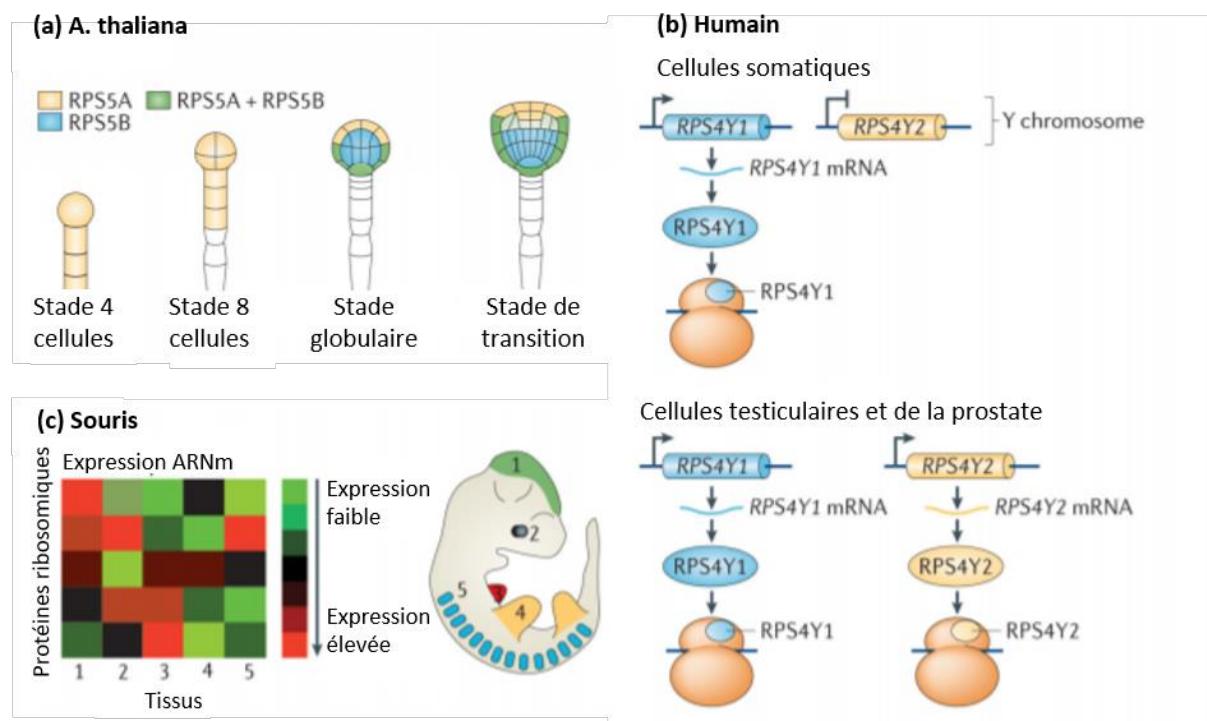


Figure 58 : Hétérogénéité des protéines ribosomiques chez plusieurs espèces

(a) Chez les plantes, les paralogues de protéines ribosomiques présentent différentes fonctions et différents profils d'expression. Chez *A. thaliana*, RPS5A est synthétisée dans les cellules en division rapide au début du développement embryonnaire, tandis que RPS5B est synthétisée dans les cellules en cours de différenciation. (b) Chez l'Homme, seulement quelques paralogues de protéines ribosomiques ont été identifiés. Ici, RPS4Y1 est synthétisée de façon ubiquitaire tandis que RPS4Y2 est synthétisée uniquement dans les testicules et la prostate. (c) Chez la souris, le profil d'expression des ARNm des protéines ribosomiques varie fortement suivant les tissus. Cela peut conduire à une composition en protéines ribosomiques unique des ribosomes dans chaque type cellulaire. *Adapté de Xue, 2012*⁴⁸.

Chez de nombreux procaryotes et eucaryotes, les gènes de RP présentent plusieurs paralogues qui sont synthétisés simultanément ou en réponse à certaines conditions environnementales.

Chez *S. cerevisiae*, 59 des 78 PR présentent deux copies génomiques (dû à une duplication ancienne du génome)²⁷³. De façon remarquable, bien que les paralogues présentent une homologie de séquence élevée, leurs fonctions ne sont pas totalement redondantes et la délétion individuelle de paralogues induit des phénotypes différents⁴⁸. Ainsi, les PR paralogues P1a, P1b, P2a et P2b forment deux hétérodimères qui se lient de façon préférentielle respectivement aux sites A et B de la protéine uL10 (P0). L'étude de mutants dans lesquels un seul des quatre hétérodimères P1/P2 possible est exprimé a permis de montrer que chacun de ces hétérodimères était à l'origine d'un phénotype spécifique. Cela montre qu'ils ne présentent pas des rôles physiologiques identiques mais chacun de ces rôles est important car l'absence d'un seul de ces hétérodimères a un impact négatif sur la croissance cellulaire et réduit la synthèse de la sous-unité 60S²⁷⁴.

Les plantes sont un autre exemple très représentatif de la présence de PR paralogues. Chez *Arabidopsis thaliana*, chaque PR compte deux à sept paralogues²⁷⁵, qui peuvent avoir acquis des spécificités. Par exemple, les deux paralogues uL15C (RPL27aC) et uL15B (RPL27aB) présentent des fonctions redondantes mais chacun d'eux est essentiel au développement de l'ovule²⁷⁶. Ces deux protéines paralogues ont donc suffisamment divergées pour porter des fonctions différentes impliquées dans le développement du gamète. De plus, le parologue uS7A (RPS5A) est fortement exprimé dans les cellules en division tandis que uS7B est exprimé dans les cellules en différenciation, ce qui montre là aussi l'acquisition de fonctions différentes de ces deux paralogues²⁷⁷ (Figure 58a).

Chez *Drosophila melanogaster*, plusieurs exemples de paralogues de PR sont décrits. Ainsi, Rpl22 est transcrit de manière ubiquitaire tandis que Rpl22-like (eL22-like) est enrichi dans les testicules²⁷⁸. De la même façon, les protéines uS7b, eS19a, uL1 (RpL10A) et RPL37b sont aussi enrichis dans les testicules par rapport à leurs paralogues. Une telle hétérogénéité de l'expression des PR dans les gonades suggère que le développement des cellules germinales pourrait nécessiter des variations tissu-spécifiques de la machinerie traductionnelle.

Chez les mammifères, la plupart des gènes codant des PR ne présentent qu'une seule copie, avec quelques exceptions comme eS4 et RPL39-like (eL39-like) qui sont fortement exprimées dans des cellules souches embryonnaires de souris en cours de différenciation^{279,280}. Un enrichissement de RPL39-like a également été observé dans des cellules humaines de tumeurs de carcinomes hépatocellulaires²⁸⁰. La protéine eS4 est codée par trois gènes *RPS4X*, *RPS4Y1* et *RPS4Y2*, localisés sur les chromosomes X et Y. Chez les hommes, *RPS4X* et *RPS4Y1* sont exprimés de

façon ubiquitaire mais l'expression de *RPS4Y2* est restreinte aux testicules et à la prostate, suggérant un rôle masculin-spécifique de cette protéine (Figure 58b). Bien que les séquences des trois protéines *RPS4* soient très similaires, le domaine C-terminal de *RPS4Y2* est distinct et pourrait faciliter des interactions uniques avec des facteurs testicule-spécifiques. Cet exemple conforte l'hypothèse de ribosomes spécifiques²⁷⁹.

Chez la souris, une étude sur le pattern d'expression de toutes les PR et de leurs paralogues a été réalisée dans 22 tissus différents. Cette étude a montré une importante expression tissu-spécifique²⁸⁰ (Figure 58c) et est en accord avec une autre étude réalisée chez la souris qui montre la synthèse spécifique dans les testicules des protéines *RPL10-like* (*eL16-like*) et *RPL39-like* (*eL39-like*), paralogues des PR *RPL10* (*eL16*) et *RPL39* (*eL39*) liées au chromosome X²⁸¹. De plus, chez la souris, durant le développement embryonnaire, la protéine *eL38* est exprimée plus fortement dans les somites en développement et dans un groupe spécifique de neurones moteurs au sein de la moelle épinière. Ces mêmes tissus sont ceux qui se trouvent affectés par une délétion d'*eL38*. Cependant, *eL38* est aussi exprimée dans certains tissus adultes⁷⁹. L'expression de *eL16* dans les embryons de souris est aussi fortement tissu-spécifique car elle est enrichie dans l'épiderme en développement et les bourgeons de membres²⁸² et présente un pattern d'expression vraiment spécifique dans les os fœtaux²⁸³.

Un autre mécanisme qui participe à la variabilité des propriétés de la machinerie de traduction est la modification chimique des protéines ribosomiques incorporées dans le ribosome.

Des méthylations, acétylations et hydroxylations des PR ont été observées chez *S. cerevisiae*²⁸⁴. Par exemple, la diméthylation de la protéine *uS5* joue un rôle dans le clivage et l'export nucléaire des ARNr²⁸⁵ et l'hydroxylation de la protéine *uS12* affecte la fidélité de la terminaison de la traduction, ce qui peut entraîner des défauts de croissance²⁸⁶. De plus, la protéine *uL15* est fortement ubiquitinylée durant la phase S du cycle cellulaire, tandis que l'ubiquitinylation est réduite durant la phase G1. *uL15* est localisée au niveau du PTC du ribosome, et l'ubiquitinylation de cette protéine a un effet stimulateur de la traduction. Une étude a montré que les ribosomes qui comportent *uL15* polyubiquitinylée présentent une traduction plus rapide d'un gène rapporteur que les ribosomes qui comportent *uL15* monoubiquitinylée²⁸⁷.

Chez l'Homme, on sait qu'au moins 11 protéines de la grande sous-unité et la plupart des protéines de la petite sous-unité présentent des modifications post-traductionnelles^{54,55}. Une de ces modifications est l'ajout de O-linked β -d-N-acétylglucosamine (O-GlcNAc). Il a été montré qu'une augmentation de cette modification sur des résidus Sérine et Thréonine par surexpression de la O-GlcNAc transférase entraîne une augmentation des pics de 60S et 80S en profils de polysome²⁸⁸. Cette modification post-traductionnelle contrôle donc le taux de

biogenèse des sous-unités du ribosome. Une autre modification des PR qui a un rôle dans la régulation de la traduction est l'ubiquitylation. Des études ont démontré que l'ubiquitylation des PR est impliquée dans la pause de ribosomes sur la queue poly(A) et dans l'induction des voies de surveillance NGD et NSD. Par exemple, il a été montré que la perte de l'ubiquitylation sur la PR eS10 de la sous-unité 40S entraîne des défauts de pause du ribosome sur la queue poly(A) et donc de mise en place des voies NGD/NSD^{289,290}.

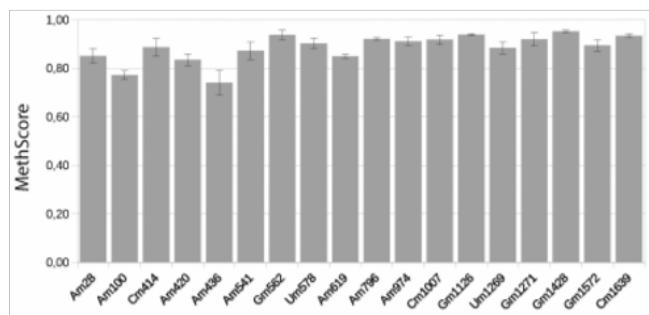
B – Hétérogénéité des séquences des ARN ribosomiques

Il existe, au sein du génome de nombreux organismes dans tous les domaines de la vie, de multiples copies des gènes ADNr. Ces copies peuvent présenter quelques différences mais sont en grande partie identiques et de ce fait sont très difficiles à séquencer.

Chez l'homme, les 400 copies des gènes ADNr ont longtemps été considérés comme identiques et à l'origine d'une population homogène d'ARNr 18S et 28S. Mais des études comparatives ont montré que ces gènes étaient en fait composés de régions conservées et de régions variables. Les régions conservées présentent des tailles, séquences et structures secondaires similaires tandis que les régions variables diffèrent non seulement en taille et en séquence mais aussi en structures secondaires. L'ARNr 28S contient ainsi 11 régions variables. De plus, entre différentes espèces, même proches, les principales différences dans les gènes ADNr se situent dans les régions variables^{291,292}. Une étude plus récente réalisée chez l'Homme et chez la souris a montré que le nombre de copies d'ADNr varie entre les individus et qu'il y avait des variations nucléotidiques dans les séquences ADNr entre les copies au sein d'un même individu et entre individus, pour les quatre ARNr²⁹³. Ces différents variants sont conservés entre la souris et l'Homme. Ils présentent également une variation d'expression selon les tissus et sont intégrés dans des ribosomes qui traduisent activement. Ces résultats confirment une hétérogénéité des ribosomes. Cependant, les aspects physiologiques de cette hétérogénéité sont encore inconnus. Il est possible qu'elle pourrait altérer la liaison de certaines protéines ribosomiques au sein du ribosome, l'association avec des facteurs extraribosomiques ou encore induire des changements de conformation du ribosome, qui affecteraient la synthèse protéique.

Chez les procaryotes, on trouve également une hétérogénéité dans les séquences d'ADNr et plusieurs exemples sont caractérisés. Le génome de *Streptomyces coelicolor* présente six copies différentes de l'ARNr de la grande sous-unité, qui sont à l'origine de molécules d'ARNr différentes au sein d'une même cellule. Ces copies d'ARNr sont différemment transcrites durant le développement²⁹⁴. De même, le génome de *Bacillus subtilis* comporte dix opérons de gènes ADNr. Leur réduction à une seule copie augmente le temps de doublement ainsi que la fréquence de

ARNr 18S



ARNr 25S

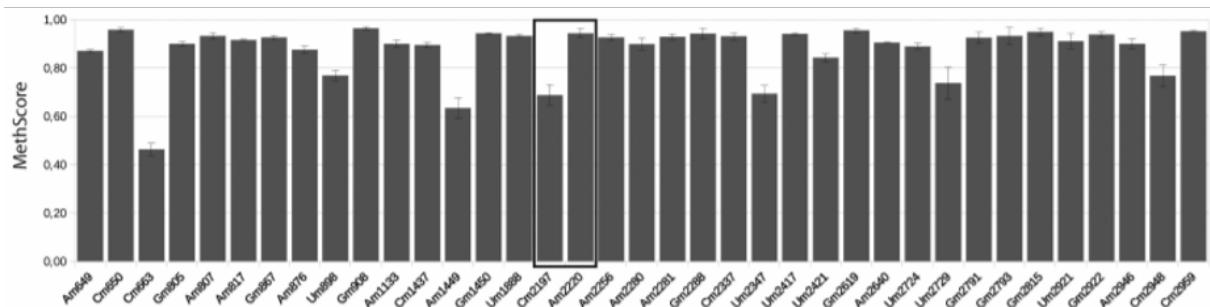


Figure 59 : Variabilité des taux de 2'-O-méthylation des ARNr 18S et 25S chez *S. cerevisiae*

Un MethScore, compris entre 0 et 1 et représentant le taux de méthylation, est calculé pour chaque position méthylée des ARNr 18S et 25S. L'encadré des barres de méthylation des positions Cm2197 et Am2220 dans l'ARNr 25S représente la variabilité des taux de méthylation entre les différentes positions. *Adapté de Marchand, 2016*²⁹⁵.

sporulation et la motilité²⁹⁶. Cela montre une importance fonctionnelle des différentes copies des gènes ADNr. Un autre exemple très intéressant concerne le parasite *Plasmodium berghei* qui présente deux gènes ADNr distincts codant des ARNr de la petite sous-unité. L'expression d'un de ces gènes ADNr (transcrit 1) a été exclusivement retrouvé quand le parasite est dans le moustique, tandis que le transcrit alternatif (transcrit 2) est fortement majoritaire quand le parasite infecte un hôte mammifère. Aucune différence structurale entre les ribosomes contenant l'un ou l'autre des ARNr n'a été détectée. De plus, les parasites qui ne présentent pas le transcrit 1 exclusivement retrouvé chez le moustique, sont capables de se développer à la fois chez les mammifères et chez le moustique. Le profil d'expression de ces deux ARNr ne semble donc pas avoir un rôle fonctionnel très spécifique²⁹⁷.

C – Hétérogénéité des modifications chimiques des ARNr

Les positions des 2'-O-méthylations et pseudouridylations sur les ARNr ont été déterminées récemment par de nouvelles techniques à haut débit^{81,295,298}. Ces études ont confirmé des sites connus depuis longtemps et ont montré l'existence de nouvelles modifications, comme la pseudouridine ψ2345 dans l'ARNr 25S de *S. cerevisiae*²⁹⁸ et deux nouvelles 2'-O-méthylations aux positions 3771 et 3764 de l'ARNr 28S dans les cellules humaines HeLa⁸¹. Certaines de ces études permettent également d'évaluer le taux de modification à chaque position de l'ARNr²⁹⁵ (Figure 59). Il a ainsi été déterminé par exemple que chez *S. cerevisiae*, environ deux tiers des sites méthylés le sont à 100% et un tiers partiellement⁸¹, et que 18 des 112 sites sont méthylés à des taux compris entre 5 et 85%²⁹⁸. La cartographie quantitative des pseudouridylations et des 2'-O-méthylations a montré que le taux de modification à une position donnée de l'ARNr était très similaire pour tous les eucaryotes⁸⁴. Ainsi, la conservation à la fois des positions modifiées et des taux de modification à chaque position montrent qu'ils sont sous pression de sélection.

Une autre étude, réalisée à partir de la structure du ribosome humain 80S obtenue par cryo-EM, a permis de visualiser les modifications chimiques des ARNr⁴⁶. Ainsi 136 modifications ont été visualisées sur le ribosome : 60 2'-O-méthylations dont 50 déjà prédites (sur les 103 décrites) et 10 non prédites, 25 pseudouridylations, toutes prédites (les 71 autres déjà décrites non pas été confirmées ici) et 51 modifications de bases dont seulement 9 étaient déjà prédites. Cette étude a donc mis au jour un nombre important de modifications des ARNr non prédites, et certaines modifications précédemment décrites non pas été confirmées. Ceci suggère que les modifications chimiques des ARNr présentent des variations entre différents types cellulaires.

Une étude réalisée en 2013 dans des cellules humaines HCT-116 a permis de faire varier les taux de 2'-O-méthylation des ARNr et d'évaluer l'impact de ces variations sur la traduction¹⁰¹. Les 2'-O-méthylation sont effectuées par la protéine fibrillarine (au sein des snoRNPs), sous le contrôle de la protéine p53. L'inhibition de p53 dans des cellules épithéliales humaines conduit à une augmentation de l'expression de la fibrillarine et à une modulation du profil de 2'-O-méthylation des ARNr. Cette modification du profil de 2'-O-méthylation s'accompagne de défauts de fidélité de la traduction et d'une augmentation de la traduction IRES-dépendante. Ainsi, l'hétérogénéité des modifications chimiques des ARNs ribosomiques présente un rôle fonctionnel et une altération de cette hétérogénéité entraîne une altération de la traduction¹⁰¹.

Les 2'-O-méthylation et les pseudouridylations sont effectuées par des complexes snoRNPs dont les niveaux d'expression varient selon les tissus²⁹⁹, ce qui pourrait être à l'origine de variations tissus-spécifiques des taux de ces modifications, et potentiellement induire des spécificités de traduction entre différents tissus. Toutefois, l'impact des variations des taux des composants des snoRNPs sur les modifications chimiques des ARNr et sur la traduction n'a pas encore été étudié.

A travers l'hétérogénéité des ribosomes dans leur composition en protéines ribosomiques et en ARNs ribosomiques a émergé la notion de « ribosomes spécialisés ». Les ribosomes spécialisés seraient responsables de la traduction de sous-groupes spécifiques d'ARNm impliqués dans des processus cellulaires et physio-pathologiques, et permettraient ainsi d'introduire un niveau supplémentaire de régulation de l'expression des gènes. Mais aujourd'hui, l'existence de ces ribosomes spécialisés n'est pas vraiment prouvée et les mécanismes ne sont pas encore compris⁸⁴.

Objectifs de thèse

Les différents travaux que j'ai effectués au cours de ma thèse portent sur l'étude du rôle des 2'-O-méthylations des ARNr sur la traduction, sa régulation et sa fidélité. La majorité de mon travail a été réalisé dans des cellules humaines HeLa mais j'ai également fait des analyses dans la levure *Saccharomyces cerevisiae*.

Mon projet s'inscrit au sein d'une collaboration avec Jean-Jacques Diaz, Frédéric Catez et Jenny Erales de l'équipe Domaines nucléaires et pathologies, à Lyon.

Jusqu'à présent, l'étude du rôle des méthylations des riboses des ARNr a été effectuée sur des groupes de méthylation ou sur des événements traductionnels ciblés. L'objectif de ma thèse était de caractériser le rôle des méthylations des ARNr dans la traduction et de dégager un mécanisme moléculaire de l'action de ces méthylations. Pour cela, nous avons choisi d'altérer les taux de méthylation par une extinction de la méthyltransférase fibrillarine, dans des cellules HeLa. La diminution du taux de fibrillarine a ainsi entraîné une baisse globale de la méthylation des ARNr. L'étude de l'impact de cette baisse de la méthylation a été réalisée par une approche globale à l'échelle du génome grâce à la technique de ribosome profiling, et par une approche spécifique d'événements traductionnels particuliers, comme le recodage et l'initiation IRES-dépendante. L'approche globale a permis de réaliser deux analyses. La première analyse a mis en évidence un groupe d'ARNm dont la traduction est altérée par la baisse de la méthylation des ARNr. La deuxième analyse a révélé une liste de sites auxquels l'initiation de la traduction est modifiée quand les ARNr sont hypométhylés. J'ai ensuite débuté une importante analyse sur ces résultats. L'approche spécifique a montré que la fidélité de la traduction et l'initiation IRES-dépendante étaient également affectées par la diminution des 2'-O-méthylations.

Dans une seconde partie de mon projet, j'ai étudié l'impact d'une perte ciblée de méthylations autour d'une région fonctionnelle du ribosome, le tunnel de sortie du peptide. Ces méthylations étant conservées entre la levure et l'Homme, j'ai choisi pour des raisons pratiques de changer d'organisme et j'ai effectué cette étude chez *S. cerevisiae*. L'objectif était de déterminer si la perte de ces méthylations était suffisante pour affecter la fidélité de la traduction. Au cours de cette étude, j'ai montré un effet inédit de l'absence de la protéine Asc1p sur la translecture des codons stop.

Résultats & Discussion

Induction de la baisse du taux des 2'-O-méthylations des ARNr

Résultats et Discussion

Induction de la baisse du taux des 2'-O-méthylations des ARNr

I – Résultats

Ce projet a été initié à la suite de la publication par l'équipe de Jean-Jacques Diaz de la démonstration que la transcription de la fibrillarine, la protéine qui méthyle les riboses des ARNs ribosomiques, est contrôlée par la protéine p53. Ainsi, une extinction de la p53 par siRNA entraîne une augmentation du taux de fibrillarine, ce qui a pour impact une variation des taux de méthylation des ARNr¹⁰¹. Dans ce contexte, nous avons voulu déterminer l'impact d'une variation du taux de méthylation des ARNr sur la fidélité de la traduction en collaboration avec l'équipe de JJD.

1 – Extinction de la fibrillarine par un shRNA

Au cours de ma thèse, je me suis intéressée à la baisse des 2'-O-méthylations des ARNr. Pour cela, une lignée cellulaire dans laquelle l'expression de la protéine fibrillarine est régulée négativement a été construite à partir de cellules HeLa. La fibrillarine étant essentielle, il n'est pas possible de la déléter dans le génome. La stratégie choisie a donc été d'utiliser un shRNA, intégré de façon stable dans le génome, ciblant spécifiquement l'ARNm de la fibrillarine. Les shRNA, ou small hairpin RNA, sont des petits ARNs double brin en forme de tige-boucle qui, après maturation, se fixe à l'ARNm cible par complémentarité de séquence et empêche la traduction et/ou entraîne la dégradation du messager. Nos collaborateurs ont intégré de façon stable dans le génome de cellules HeLa un shRNA dirigé contre la fibrillarine, noté shFBL, à l'aide d'un système lentiviral. Le shFBL est complémentaire de l'exon VIII de la fibrillarine, il est exprimé en fusion transcriptionnelle avec la séquence de la protéine fluorescente RFP (Red Fluorescent Protein) et est sous contrôle d'un promoteur inductible par la doxycycline. Cette stratégie a été choisie car les shRNA sont plus stables que les siRNA et leur expression peut être contrôlée.

Depuis l'expression du shFBL jusqu'à un effet sur la traduction due à l'hypométhylation des ARNr, il faut tenir compte de différentes étapes de biogénèse qui doivent permettre un remplacement de la machinerie de la traduction. En effet, une fois la doxycycline ajoutée au milieu, le shFBL doit être exprimé en quantité suffisante, être mûr puis entraîner une baisse du taux de la protéine fibrillarine. Ensuite, les ARNr normaux doivent être remplacés par des ARNr hypométhylés au sein des ribosomes. Lorsque la population des ribosomes normaux a été renouvelée par des ribosomes aux ARNr hypométhylés, des effets sur la traduction peuvent potentiellement être

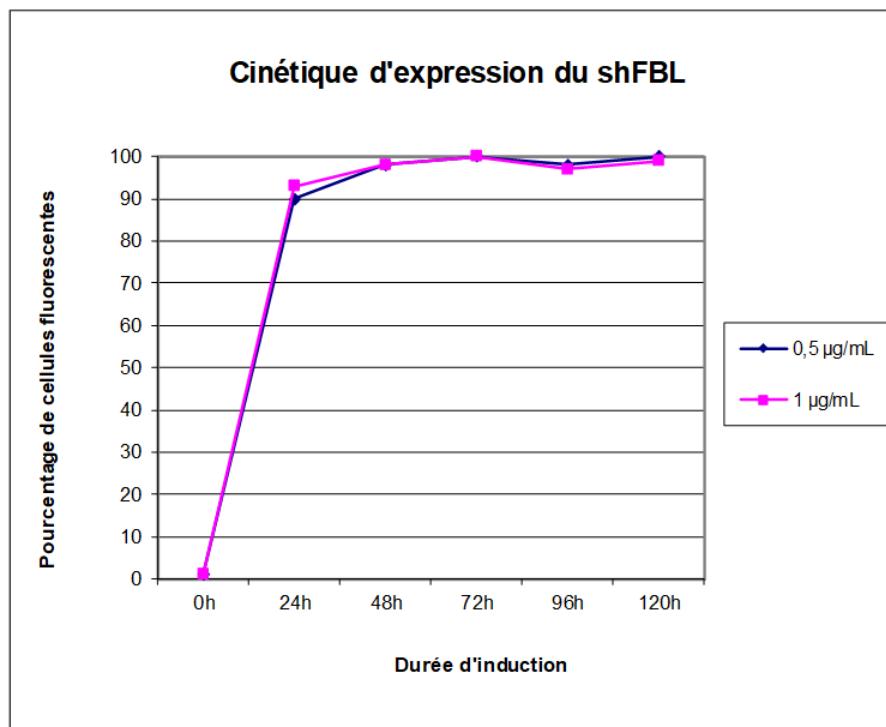


Figure 1 : cinétique d'expression du shFBL, en pourcentage de cellules fluorescentes.

Le pourcentage de cellules HeLa shFBL fluorescentes (RFP) est mesuré après différents temps d'induction à la doxycycline. Deux concentrations de doxycycline sont testées : 0,5 µg/mL et 1 µg/mL.

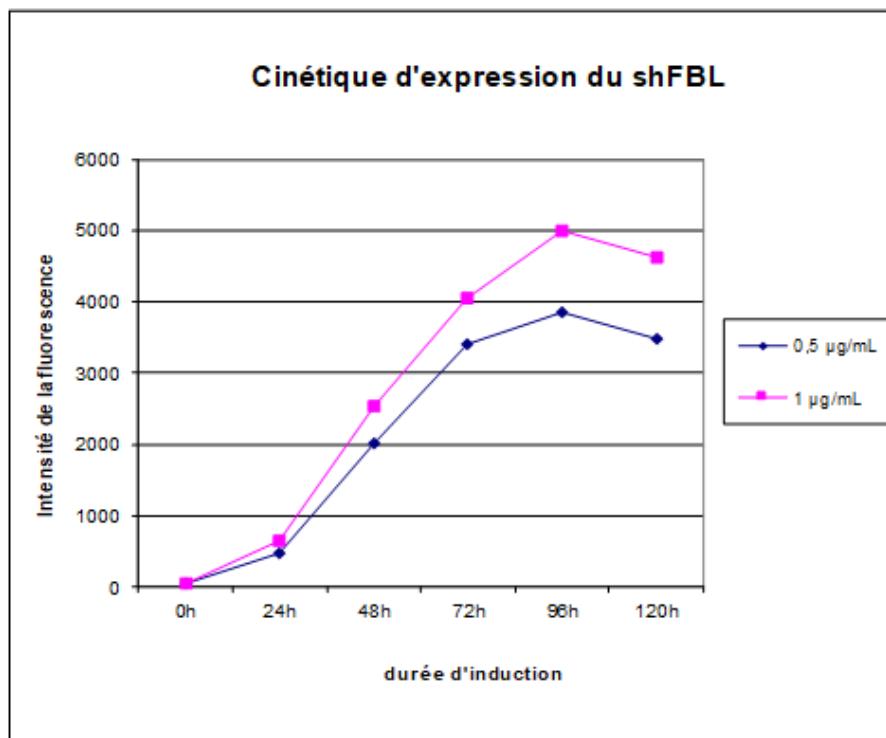


Figure 2 : cinétique d'expression du shFBL, en intensité de fluorescence.

L'intensité de la fluorescence RFP des cellules HeLa shFBL est mesurée après différents temps d'induction à la doxycycline. Deux concentrations de doxycycline sont testées : 0,5 µg/mL et 1 µg/mL.

mesurés. La durée de chaque étape est à prendre en compte, cependant, certaines données ne sont pas complètement connues aujourd’hui, comme la durée de renouvellement des ribosomes que l’on sait varier entre 1 et 5 jours dans des cellules de foie en prolifération³⁰⁰. Il a donc fallu que je mette au point les conditions d’expression du shFBL de façon à obtenir une hypométhylation des ARNr suffisante pour induire des effets sur la traduction sans altérer totalement le fonctionnement de la cellule à cause de la baisse du taux de fibrillarine.

2 – Baisse du taux de fibrillarine et diminution globale de la méthylation des ARNr

La mise au point des conditions d’induction a consisté à déterminer une concentration de doxycycline et une durée d’induction qui permettent l’expression du shFBL, l’extinction de la fibrillarine, la baisse de la 2'-O-méthylation des ARNr et des effets mesurables sur la traduction.

Dans un premier temps, le shFBL étant en fusion transcriptionnelle avec la RFP, j’ai suivi leur expression par cytométrie de flux au cours du temps et à deux concentrations de doxycycline différentes. J’ai mesuré le pourcentage de cellules qui expriment la RFP (et donc le shFBL) et l’intensité de la fluorescence toutes les 24h pendant 120h, pour des concentrations de doxycycline de 0,5 µg/mL et 1 µg/mL. Les résultats montrent que dès 48h d’induction, 100% des cellules expriment la RFP (Figure 1). En ce qui concerne l’intensité de la fluorescence de la RFP, qui reflète le maximum de l’expression du shFBL, elle est maximale à 96h d’induction pour les deux concentrations de doxycycline et plus élevée pour une concentration de 1 µg/mL que 0,5 µg/mL (Figure 2).

J’ai ensuite vérifié par Western Blot l’inhibition de la synthèse de fibrillarine par le shFBL à différents temps d’induction et pour les deux concentrations de doxycycline. Avec une concentration de 0,5 µg/mL, il reste 59% du taux initial de fibrillarine après 96h de traitement, et après 120h de traitement il reste 70% du taux initial (Figure 3). Avec une concentration de 1 µg/mL, le taux de fibrillarine diminue de façon beaucoup plus importante. Dès 48h, il reste seulement 44% du taux initial de fibrillarine et après 120h, il reste 7% du taux initial (Figure 4). Les résultats de ces deux expériences indiquent qu’un traitement de 120h avec 1 µg/mL de doxycycline induit une forte expression du shFBL dans toutes les cellules et une importante diminution du taux de fibrillarine.

Cette lignée cellulaire a été construite dans le but d’induire une baisse du taux de 2'-O-méthylation des ARNr. Il a donc fallu s’assurer que la méthylation baissait. Pour cela nous avons caractérisé l’état des méthylations des ARNr au sein des ribosomes actifs grâce à la technique de RiboMethSequencing. Cette technique, développée par l’équipe de Yuri Motorin à Nancy, permet de

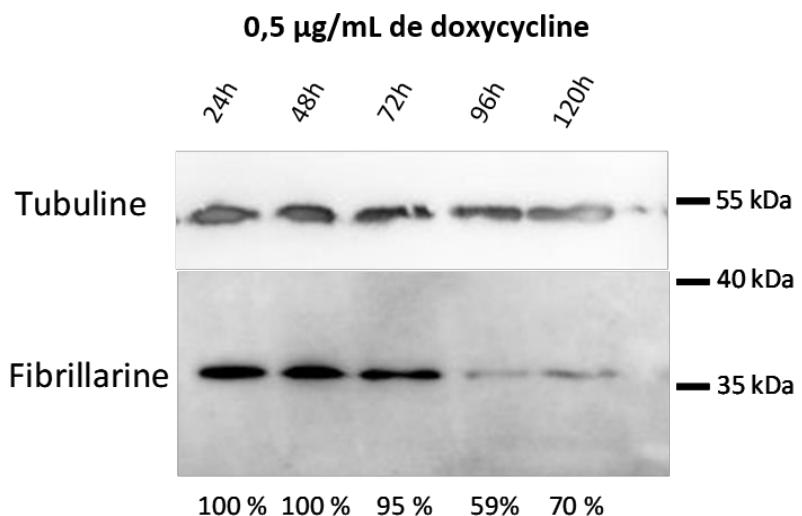


Figure 3 : Western Blot de la fibrillarine à différents temps d'induction avec 0,5 µg/mL de doxycycline.

La fibrillarine et la tubuline sont révélées dans des extraits protéiques de cellules HeLa shFBL traitées avec 0,5 µg/mL de doxycycline pendant différents temps. Le taux de fibrillarine à chaque temps est calculé et rapportée au taux initial au temps t0.

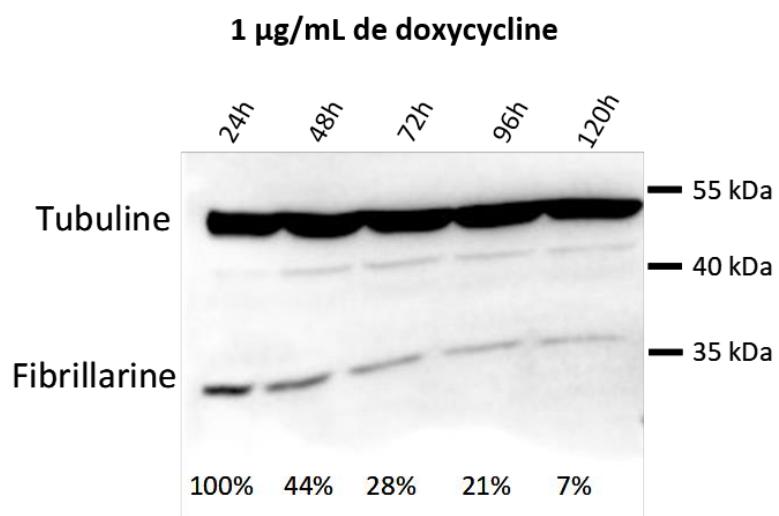


Figure 4 : Western Blot de la fibrillarine à différents temps d'induction avec 1 µg/mL de doxycycline.

La fibrillarine et la tubuline sont révélées dans des extraits protéiques de cellules HeLa shFBL traitées avec 1 µg/mL de doxycycline pendant différents temps. Le taux de fibrillarine à chaque temps est calculé et rapportée au taux initial au temps t0.

cartographier quantitativement les 2'-O-méthylation sur l'ensemble des ARNs. Elle est fondée sur la protection contre l'hydrolyse des groupements méthyles et le séquençage à haut débit des extrémités des fragments générés (Figure 5). J'ai donc extrait les ARNr de cellules HeLa shFBL traitées ou non avec 1 µg/mL de doxycycline pendant 120 heures, et de cellules HeLa témoin ou de cellules HeLa traitées avec un siFBL. Le RibomethSeq a été réalisé en collaboration avec l'équipe de Yuri Motorin à Nancy. Dans ce paragraphe, je me focaliserai uniquement sur les résultats obtenus dans les cellules HeLa shFBL traitées ou non à la doxycycline. Les résultats des cellules HeLa et HeLa siFBL (qui sont identiques à ceux obtenus avec le shFBL) sont détaillés dans l'article « Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes. » dont je suis co-auteur (voir partie correspondante). Nous avons dans un premier temps effectué une carte de référence sur des cellules HeLa shFBL non traitées à la doxycycline. La première observation est que sur l'ensemble des sites méthylés, tous ne le sont pas dans 100% des ribosomes. Cependant, la majorité des sites sont méthylés à plus de 80% (Figure 6). Parmi ces sites efficacement méthylés on retrouve tous les sites conservés entre la levure et l'homme. Ainsi cette carte de référence montre une hétérogénéité du profil des méthylations des ARNr au sein des ribosomes humains dans des conditions normales de croissance.

Le RiboMethSeq réalisé à partir de cellules HeLa dont l'expression du shFBL est induite a révélé une baisse globale de la méthylation des ARNr. De façon surprenante, cette baisse de la méthylation varie fortement entre les sites, allant de 0,2% à 57%, ce qui révèle un impact site-spécifique de la baisse de la fibrillarine. Il est intéressant de noter que les sites qui présentent un taux initial de méthylation de moins de 80% montrent une diminution d'au moins 10% et de 44% en moyenne tandis que les sites qui présentent un taux initial supérieur à 80% montrent une diminution d'en moyenne 7% (Figure 7). Ces résultats laissent suggérer que l'ajout de groupements méthyles à ces positions serait moins efficace et plus sensible à la baisse de la fibrillarine, ou que seuls les ribosomes qui présentent des méthylations aux sites qui apparaissent les moins touchés sont formés correctement.

Par la suite, nous avons positionné tous les sites 2'-O-méthylés sur une structure du ribosome 80S de cellules HeLa obtenue récemment par cryo-EM. Des sites affectés sont présents au cœur du ribosome, la partie la plus conservée au sein des trois règnes du vivant, ainsi que dans des régions impliquées dans le processus de traduction, en particulier proche des sites A et P, des ponts entre les deux sous-unités et du tunnel de sortie du peptide naissant. Les méthylations présentes autour du centre peptidyl-transférase et du centre de décodage ne sont en revanche pas affectées. Ces observations montrent que certaines régions fonctionnelles des ribosomes peuvent être hypométhylées tandis que d'autres régions ne sont pas sujettes à la variation du taux de méthylation des ARNr dans les ribosomes matures.

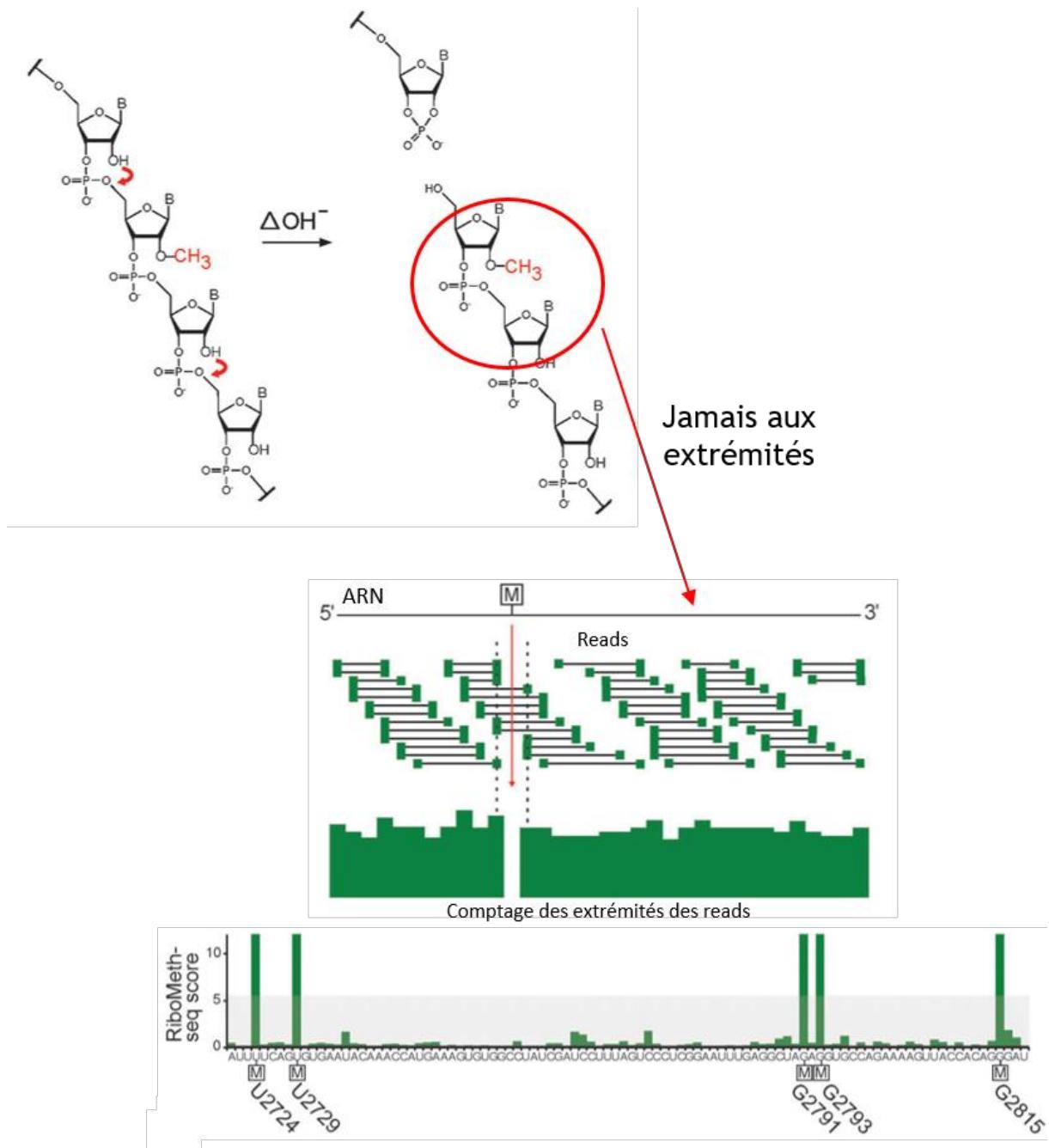


Figure 5 : Evaluation du taux de 2'-O-méthylation par RiboMethSequencing

La technique de RiboMethSequencing permet de cartographier quantitativement les 2'-O-méthylation sur l'ensemble des ARNs. Les groupements méthyles protègent les ARNs contre l'hydrolyse et ne sont donc jamais retrouvés aux extrémités des fragments générés. Le séquençage à haut débit des extrémités de ces fragments donne accès à un RiboMethScore, qui correspond à la fraction de positions méthylées au sein d'une population de ribosomes. *Adapté de Birkedal, 2014*³⁰¹.

3 – Effets de la diminution du taux de fibrillarine et de la baisse des méthylations des ARNr sur la traduction

J'ai ensuite vérifié l'impact de la baisse de la fibrillarine sur la traduction. L'équipe de Jean-Jacques Diaz avait déjà travaillé avec des cellules dans lesquelles l'expression de la fibrillarine était inhibée par siRNA et avait mis en évidence une diminution de l'initiation IRES-dépendante. Les IRES des gènes IGF-1R, FGF1 et c-Myc avaient été testés. Pour les IRES de IGF-1R et de FGF1, l'équipe avait observé une baisse d'environ 20 à 25% de l'initiation tandis qu'elle ne voyait aucun effet sur l'initiation de l'IRES c-Myc. De manière similaire j'ai quantifié l'efficacité d'initiation de chacun des IRES dans la lignée cellulaire exprimant le shFBL grâce à des systèmes double rapporteurs luciférase Renilla - luciférase Firefly (Figure 8). Ce système double rapporteur comporte la séquence codante de la luciférase Renilla sous contrôle de l'initiation canonique suivie de la séquence codante de la luciférase Firefly dont l'initiation de la traduction dépend de l'IRES testé. J'ai réalisé ces tests dans des cellules HeLa shFBL, traitées ou non avec 1 µg/mL de doxycycline pendant 120h. Les données sont ensuite traitées statistiquement par le test de Wilcoxon-Mann-Whitney. J'observe une diminution significative d'environ 22% de l'initiation à l'IRES IGF-1R, une diminution significative d'environ 25% de l'initiation de l'IRES FGF1 et aucune variation significative de l'initiation de l'IRES c-Myc (Figure 9). Les résultats que j'ai obtenus sur les trois IRES IGF-1R, FGF et c-Myc confirment ceux obtenus par l'équipe de Jean-Jacques Diaz avec des cellules traitées avec un siFBL. On peut donc conclure que l'effet du shFBL inducible par la doxycycline est équivalent à celui du siFBL.

Ces différents tests ont permis de déterminer que la durée d'induction la plus courte à laquelle on observe des effets de la baisse de la méthylation des ARNr est de 120h avec 1 µg/mL de doxycycline. Dans l'objectif de réaliser par la suite une étude à l'échelle du génome de l'impact de la baisse de la méthylation sur la traduction, nous souhaitons révéler les effets primaires de cette baisse. C'est pourquoi il est important de réaliser les expériences au premier temps auquel on obtient des effets sur la traduction.

Il a également été important de vérifier qu'avec ces conditions de traitement, la baisse du taux de fibrillarine permet toujours une traduction active. Pour cela, j'ai réalisé des profils de polysomes à partir d'extraits de polysomes de cellules HeLa shFBL traitées ou non avec 1 µg/mL pendant 120h (Figures 10 et 11). Ces deux profils montrent des pics correspondants aux sous-unités 40S et 60S qui révèlent une biogénèse correcte des sous-unités. Ces pics sont suivis par un grand pic de monosomes 80S et des polysomes montrant jusqu'à 7 ribosomes par ARNm, ce qui révèle une

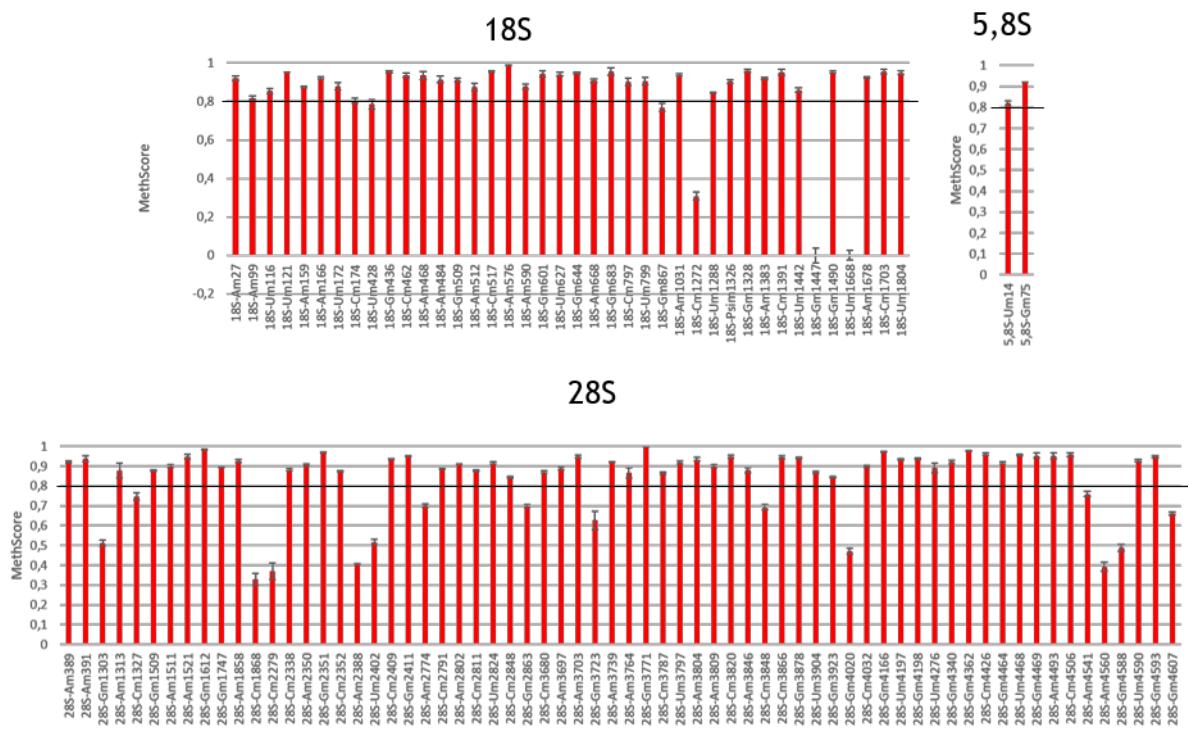


Figure 6 : RiboMethSeq des cellules HeLa shFBL non traitées à la doxycycline

Le RiboMethScore pour chaque position méthylée des ARNr de cellules HeLa shFBL non traitées à la doxycycline est représenté. Les barres noires horizontales indiquent un RiboMethScore de 0,8, soit 80% de ribosomes qui sont méthylés à une position donnée.

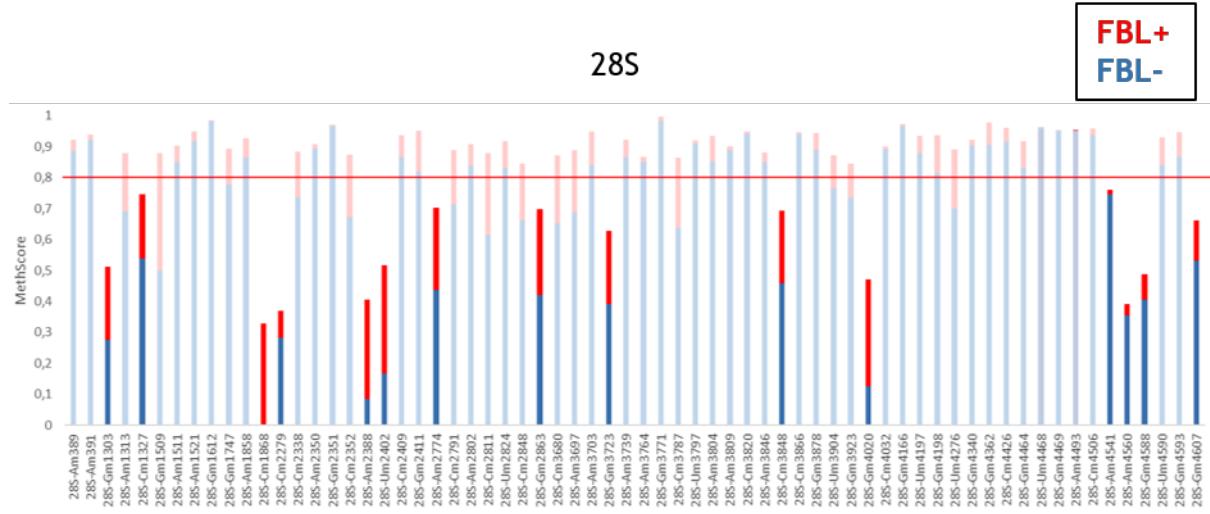


Figure 7 : Superposition des RiboMethScores des cellules HeLa shFBL traitées ou non à la doxycycline

Les RiboMethScores des cellules HeLa shFBL non traitées à la doxycycline (FBL+, en rouge) et des cellules HeLa shFBL traitées pendant 120h avec 1 μ g/mL de doxycycline (FBL-, en bleu) sont superposés. La barre rouge horizontale indique un RiboMethScore de 0,8. Je n'ai représenté que l'ARN 28S car les ARNs 18S et 5,8S présentent des variations très faibles du taux de 2'-O-méthylation.

traduction active, à la fois dans les cellules qui n'expriment pas le shFBL et dans celles qui l'expriment. Ces conditions induisent donc des effets sur la traduction à travers une baisse du taux de fibrillarine et de méthylation des ARNr, tout en permettant toujours une traduction active.

II – Discussion

Des résultats précédemment obtenus par nos collaborateurs de l'équipe de Jean-Jacques Diaz ont montré que des variations du taux de la méthyltransférase fibrillarine entraînent des variations du taux de 2'-O-méthylation des ARNr, qui induisent des changements traductionnels. L'objectif principal de ma thèse a été de déterminer le rôle des 2'-O-méthylation dans la fidélité de la traduction. Nous avons choisi d'induire une diminution du taux de méthylation des ARNr par l'extinction de la fibrillarine. L'injection de siRNA contre l'ARNm de la fibrillarine (siFBL) avait déjà été testée par nos collaborateurs, mais c'est une technique qui peut présenter une certaine variabilité lors de l'étape de transfection du siFBL dans les cellules. Nous avons donc fait le choix d'utiliser un système avec un shRNA (small hairpin RNA) dirigé contre la fibrillarine (shFBL). Le shFBL est intégré de façon stable dans le génome, son expression est inductible et, de par sa structure double brin, il est plus stable qu'un siRNA. Néanmoins, ce système présente des étapes supplémentaires, comme l'expression et la maturation du shRNA, qu'il faut prendre en compte pour la durée d'induction. J'ai réalisé des cinétiques des différentes étapes depuis l'induction de l'expression du shFBL jusqu'à un effet mesurable sur l'initiation de la traduction IRES-dépendante, ce qui m'a permis de déterminer des conditions d'induction de l'expression du shFBL qui induisent des effets sur la traduction tout en conservant une traduction active.

J'ai vérifié la baisse de la méthylation pour chaque position des ARNr par RiboMethSeq. Cela m'a permis de valider une baisse globale de la méthylation. Mais de façon surprenante, cela a également révélé que la diminution de la méthylation varie de manière site-spécifique. Des hypothèses ont été émises pour expliquer cette réponse site-spécifique : (1) puisque ce sont les sites les moins méthylés en condition normale qui sont les plus affectés par la diminution de la fibrillarine, l'ajout d'un groupement méthyle à ces sites serait moins efficace en condition normale. Cette efficacité moins bonne pourrait être due à une hybridation du snoARN guide moins efficace ou à une accessibilité du site plus difficile, et rendrait les sites plus sensibles à la baisse de la fibrillarine ; (2) la formation du ribosome agirait comme un filtre, où seules les sous-unités qui présentent des méthylations aux sites qui apparaissent les moins touchés sont formées correctement. Les ARNr qui ne présentent pas de méthylations à ces sites pourraient ne pas être clivés correctement et être dégradés ou ne pas être incorporés correctement dans les ribosomes, ce qui conduirait également à

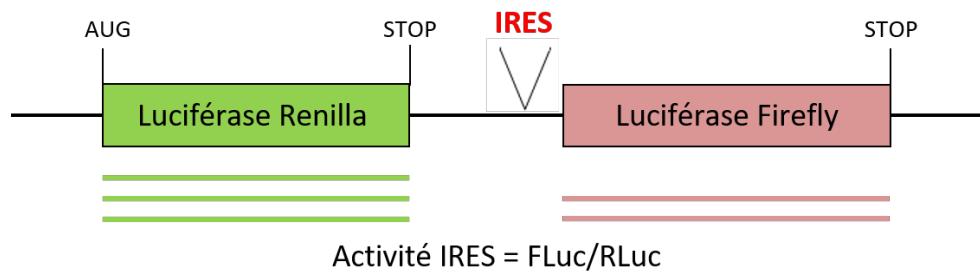


Figure 8 : Schéma du système double rapporteur double luciférase.

Ce système rapporteur est composé de la séquence codante de la luciférase renilla avec ses codons AUG et stop suivie de la séquence codante de la luciférase firefly avec son codon stop. Une séquence IRES est insérée en amont de la séquence codante de la luciférase firefly. La luciférase renilla est synthétisée de façon canonique tandis que la luciférase firefly est synthétisée à partir de la séquence IRES. L'activité de l'IRES est représentée par le rapport de l'activité de la luciférase firefly sur l'activité de la luciférase renilla.

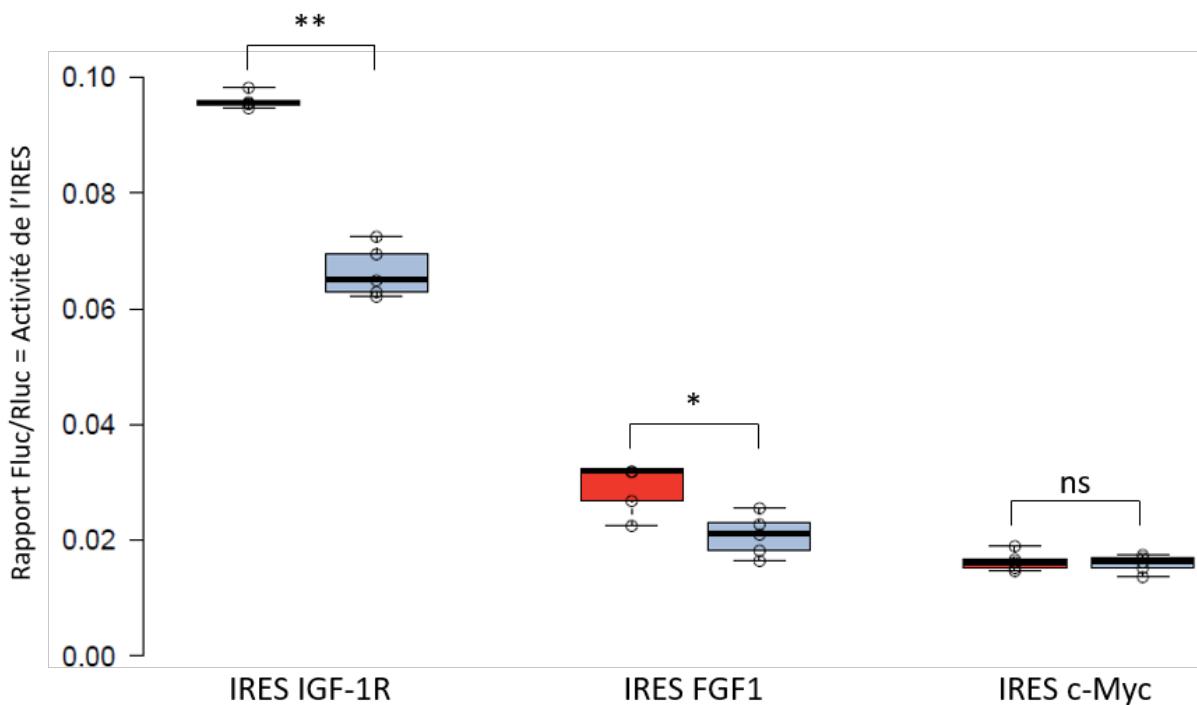


Figure 9 : Activité des IRES IGF-1R, FGF1 et c-Myc dans des cellules HeLa shFBL traitées ou non à la doxycycline

L'activité des IRES IGF-1R, FGF1 et c-Myc est calculée en faisant le rapport de l'activité de la luciférase Firefly sur l'activité de la luciférase Renilla. Elle a été calculée dans des cellules HeLa shFBL traitées (en bleu) ou non (en rouge) pendant 5 jours avec 1 µg/mL de doxycycline. Un test statistique de Wilcoxon-Mann-Whitney a été appliqué sur les données. ns : non significatif ; * p ≤ 0,05 ; ** p ≤ 0,01

leur dégradation ; (3) les niveaux des snoARNs seraient déséquilibrés et certains snoARNs deviendraient fortement limitants. Ce déséquilibre pourrait être causé par une dégradation des snoARNs qui ne sont plus intégrés aux complexes snoRNPs du fait de la forte diminution du taux de fibrillarine. Les différents snoARNs ne présenteraient pas la même sensibilité à la dégradation.

Les positions méthylées ont été localisées au sein du ribosome humain, ce qui a permis de montrer que certaines des méthylations les plus touchées étaient situées dans des régions fonctionnelles, proches des sites A et P, des ponts entre les deux sous-unités et du tunnel de sortie du peptide naissant. Cependant, les méthylations présentes au niveau du centre peptidyl-transférase (PTC) et du centre de décodage (DC) ne sont que très peu affectées. Cela suggère que les ribosomes ne tolèreraient pas une diminution des méthylations dans ces deux régions les plus importantes pour la synthèse protéique, et que ces méthylations auraient un rôle essentiel dans la fonction des ribosomes. Cette hypothèse pourrait être vérifiée en déletant individuellement les snoARNs responsables des 2'-O-méthylations dans le PTC et le DC. Si ces méthylations sont très importantes ou essentielles, leurs pertes devraient entraîner des défauts de traduction très forts voire la mort des cellules.

Pour réaliser une intégration stable dans le génome de cellules humaines, il existe deux méthodes principales. La première consiste à réaliser une intégration contrôlée dans une région ciblée du génome avec le système CRISPR-Cas9. L'utilisation de ce système génère une population clonale et homogène. La deuxième méthode est plus simple et plus rapide et consiste à effectuer une insertion par un lentivirus. C'est cette stratégie que nous avons réalisée. Dans ce cas, le fragment d'ADN s'intègre n'importe où dans le génome des cellules, générant une population non clonale et hétérogène. Il est possible de sélectionner un clone individuel mais ce clone peut présenter un phénotype lié à l'insertion du shFBL dans le génome, c'est pourquoi nous avons choisi de travailler sur une population de cellules HeLa shFBL polyclonale hétérogène. Cette hétérogénéité présente des avantages et des inconvénients. En effet, l'insertion du shFBL par infection lentivirale peut s'effectuer à n'importe quel locus du génome et donc dans des gènes ou des régions importantes, entraînant des effets délétères sur la cellule. Dans une population hétérogène, cet effet est compensé par les cellules dont l'intégration du shFBL n'a aucun effet. Cependant, une telle population hétérogène peut voir sa composition varier au cours des passages cellulaires en s'enrichissant en cellules dont l'intégration du shFBL est peu délétère et s'appauvrir en cellules dont l'intégration du shFBL est plus délétère. C'est ce que l'on appelle la dérive génétique. Concernant les cellules HeLa shFBL, j'ai pu observer ce phénomène de dérive génétique lors de la phase de mise au point, par une diminution du nombre de cellules qui expriment le shFBL et de la quantité de shFBL exprimés au cours des passages. Pour toutes mes expériences il a donc fallu que je m'assure de travailler sur des

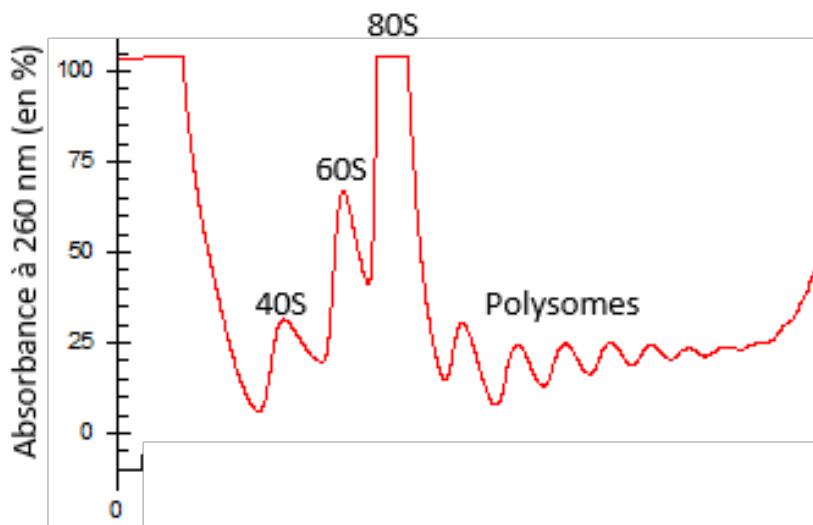


Figure 10 : profil de polysomes de cellules HeLa shFBL non traitées à la doxycycline

Des extraits de polysomes sont récoltés à partir de cellules HeLa shFBL non traitées à la doxycycline, après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucre 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm.

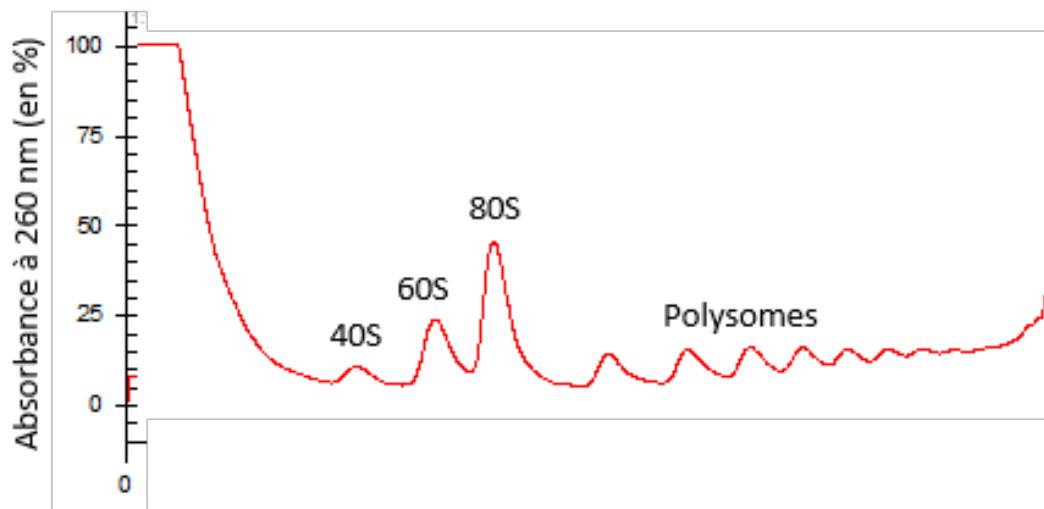


Figure 11 : profil de polysomes de cellules HeLa shFBL traitées avec 1µg/mL de doxycycline pendant 5 jours

Des extraits de polysomes sont récoltés à partir de cellules HeLa shFBL traitées avec 1µg/mL de doxycycline pendant 5 jours, après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucre 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm.

populations de cellules les moins différentes possible. Pour cela j'ai amplifié en très grand nombre des cellules avec un faible nombre de passages et j'ai utilisé pour chaque expérience une nouvelle ampoule issue de cette amplification. A chaque expérience, je me suis assurée que 100% des cellules exprimaient le shFBL et que le taux d'expression était le même par des mesures au cytomètre de flux.

Le RiboMethSeq a montré une baisse globale de la méthylation des ARNr, néanmoins cette baisse est limitée et il serait intéressant d'identifier des conditions dans lesquelles cette baisse est plus importante. De plus, la fibrillarine étant impliquée dans le clivage des ARNr, son extinction entraîne des défauts indépendants de la diminution de la méthylation des ARNr, notamment dans la biogénèse du ribosome³⁰². Chez la levure, un mutant de Nop1 qui présente uniquement des défauts de méthylation, Nop1-3, a été obtenu et étudié⁹⁹. De la même manière nous pourrions introduire une ou plusieurs mutations ponctuelles dans le gène codant la fibrillarine, localisée(s) au niveau du site catalytique ou de la poche de liaison à la molécule SAM, qui affecterai(en)t uniquement l'activité méthyltransférase de la fibrillarine et non son activité dans le clivage des ARNr.

Un autre moyen de bloquer les méthylations des ARNr est d'affecter les snoARNs, soit en les déletant tous, soit en altérant la complémentarité des bases entre les snoARNs et les régions des positions à méthylérer. Nous pourrions alors déterminer si ces ARNr non méthylés ou très peu méthylés peuvent être incorporés dans les sous-unités ribosomiques, elles-mêmes assemblées en ribosomes 80S fonctionnels. Cependant, certains snoARNs guides de la méthylation sont également impliqués dans le clivage des ARNr et l'assemblage des sous-unités du ribosome et leur déletion ou l'altération de leur complémentarité avec l'ARNr pourraient entraîner des défauts supplémentaires qui ne seraient pas liés à la perte des méthylations.

Ces deux méthodes proposent des moyens d'améliorer la diminution de la méthylation des ARNr par rapport à l'extinction de la fibrillarine par shRNA. Cependant, chez les mammifères, la méthylation des ARNr est nécessaire pour un clivage correct du pré-ARNr. Une forte diminution des 2'-O-méthylations du pré-ARNr serait donc très délétère pour les cellules. Le meilleur moyen d'obtenir des ARNr non méthylés ou très peu méthylés serait donc de déméthylérer ces ARNr après le clivage. Aujourd'hui, aucune protéine, aucun complexe ni aucune molécule n'est connue pour enlever les 2'-O-méthylations. Toutefois on sait que d'autres méthylations des ARNs sont retirées par une déméthylase, comme les méthylations m⁶A des ARNm qui sont enlevées par la protéine ALKBH5³⁰³. Il n'est donc pas à exclure qu'un tel facteur existe.

Résultats & Discussion

Analyses globale et ciblée des effets de la baisse de la méthylation des ARNr sur la traduction

Analyses globale et ciblée des effets de la baisse de la méthylation des ARNr sur la traduction

I – Résultats

Jusqu'à présent, les études réalisées sur le rôle de la méthylation des ARNr sur la traduction étaient des études focalisées sur des cibles précises. Le principal objectif de ma thèse a été d'effectuer une étude à l'échelle du génome de l'impact de la baisse globale de la méthylation des ARNr sur la traduction. En parallèle, j'ai déterminé les effets de cette hypométhylation sur la fidélité de la traduction en utilisant des cibles de recodage.

1 – Analyse globale à l'échelle du génome

Afin de réaliser cette étude à l'échelle du génome, j'ai utilisé la technique de ribosome profiling (appelée aussi RiboSeq). Cette technique, mise au point par Ingolia en 2009, est fondée sur le séquençage à haut débit des fragments d'ARNm protégés par le ribosome¹²⁷. Les ARNm avec les ribosomes en traduction sont extraits des cellules, puis digérés avec une RNase. Les ribosomes protègent des fragments d'ARNm d'environ 30 nucléotides, appelés aussi empreintes ou RPFs, qui ne seront pas digérés par la RNase. Le séquençage de ces fragments et leur alignement sur le génome permettent de localiser précisément la position de chaque ribosome. Le comptage des RPF permet en parallèle de déterminer la quantité de ribosome présent sur chaque ARNm. Le RiboSeq permet donc à la fois une analyse quantitative en rendant compte de l'expression des gènes, et une analyse qualitative en donnant accès à la position des ribosomes au nucléotide près.

J'ai effectué la préparation de RPFs à partir de cellules HeLa shFBL. Pendant 5 jours, ces cellules sont mises en culture en milieu normal ou supplémenté avec de la doxycycline pour induire l'expression du shFBL et donc entraîner la baisse du taux de fibrillarine et de méthylation des ARNr. Au bout des 5 jours de culture, la traduction est bloquée par la molécule cycloheximide et les polysomes sont extraits. Les polysomes sont ensuite traités à la RNase I afin d'éliminer l'ARNm qui n'est pas protégé par les ribosomes. Puis les ARNs, une fois débarrassés des ribosomes, sont sélectionnés suivant la taille. Seuls les fragments compris entre 28 et 34 nucléotides sont gardés, ce qui correspond à la taille des fragments protégés par les ribosomes. Parmi ces ARNs, une grande quantité est de l'ARN ribosomique qui résulte de la digestion partielle des ribosomes par la RNase I. Ces contaminants ribosomiques sont éliminés avant de procéder à la création des banques et au séquençage. L'expérience a été réalisée en triplicat.

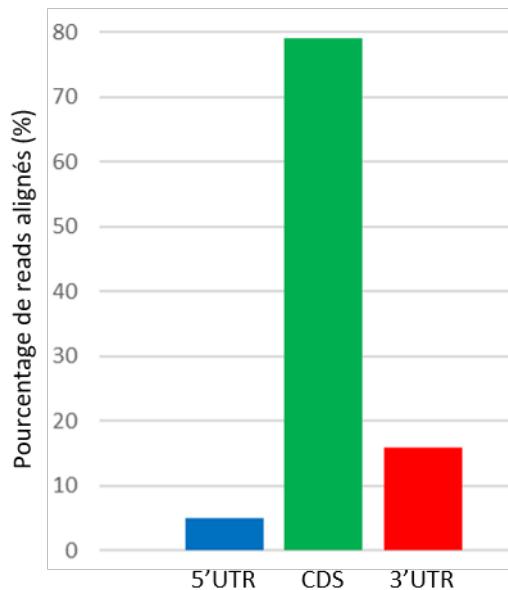


Figure 12 : Répartition des reads dans chacune des régions des ARNs messagers, pour la lignée HeLa shFBL traitée à la doxycycline

Les reads sont comptés suivant la région de l'ARNm où ils s'alignent.

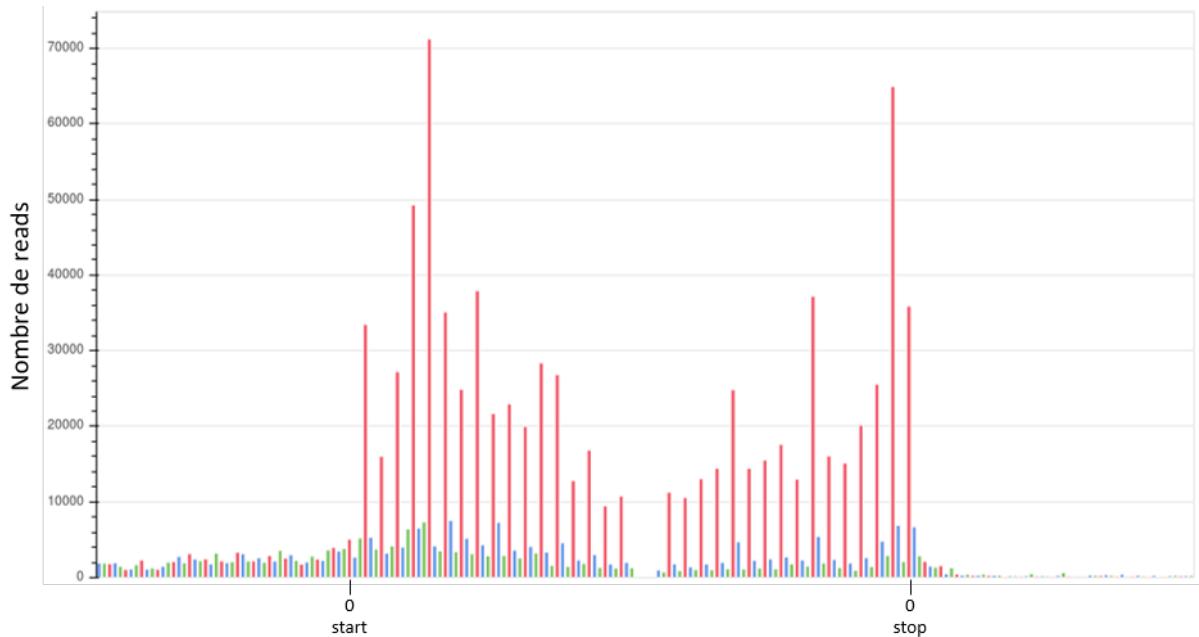


Figure 13 : Phasage des reads dans des cellules HeLa

Les positions du site P des ribosomes sont déterminées sur les reads de tous les gènes, puis sont poolées et comptées autour du start et du stop. Les trois phases de lecture par le ribosome sont représentées par trois couleurs : en rouge la phase 0, en bleu la phase 1 et en vert la phase 2.

En parallèle j'ai extrait les ARNs totaux des cellules HeLa shFBL traitées ou non à la doxycycline pendant 5 jours afin de réaliser des transcriptomes. L'étude des transcriptomes en parallèle des RiboSeq permet de déterminer si les variations d'expression observées avec le RiboSeq sont dues uniquement à une régulation de la traduction ou à une régulation de la transcription. Les transcriptomes sont aussi effectués en triplicat.

2 – Analyse bioinformatique des données de séquençage

Les données de séquençage des transcriptomes et des RiboSeq ont été analysées par Pierre Bertin, le bioinformaticien de notre équipe. Pour les deux expériences, il a procédé à plusieurs étapes avant l'alignement des séquences, ou reads, sur le génome. A la sortie du séquençage, les séquences présentent aux deux extrémités des adaptateurs qui sont éliminés. Puis, pour les RiboSeq, les reads trop courts et trop longs sont éliminés, ainsi seuls ceux dont la taille est comprise entre 25 et 34 nucléotides sont conservés. Ensuite, les reads sont alignés sur les séquences ribosomiques afin d'éliminer les séquences contaminantes restantes. Les séquences qui n'ont pas été éliminées au cours de ces étapes sont alors alignées sur le génome de référence humain (hg38), avec une autorisation de deux mésappariements et en ne conservant que les reads qui s'alignent de manière unique sur le génome. Les reads alignés sont ensuite comptés.

Le déroulement de ces différentes étapes est rythmé par différents contrôles qualité des séquences, et donc de l'expérience.

Pour les RiboSeq, pour chaque étape d'élimination des adaptateurs, des reads trop courts et trop longs et des séquences ribosomiques, le pourcentage de reads restants est calculé. Chaque étape faisant diminuer le nombre de reads, il ne reste à la fin qu'entre 20 et 30% du nombre de reads total en sortie de séquençage. Ces calculs permettent de mettre en évidence des problèmes éventuels au cours de la préparation des RPFs, lors de la digestion par la RNase I (reads trop courts ou trop longs) ou de l'élimination des séquences ribosomiques contaminantes (alignement important des reads sur les séquences ribosomiques) par exemple ; ou lors de la création de la banque, si trop de reads sont éliminés à l'étape d'élimination des adaptateurs.

Les comptages des reads sont compilés pour tous les gènes sous forme de métagène afin de caractériser la répartition des reads le long des gènes. Pour les transcriptomes, la répartition des reads doit être homogène entre les régions 5' et 3'UTR et la séquence codante (CDS) tandis que pour les RiboSeq, comme on ne récupère que les fragments d'ARNm sur lesquels se trouvent des ribosomes, il doit y avoir peu de reads sur les régions UTRs et un enrichissement sur la CDS. Les

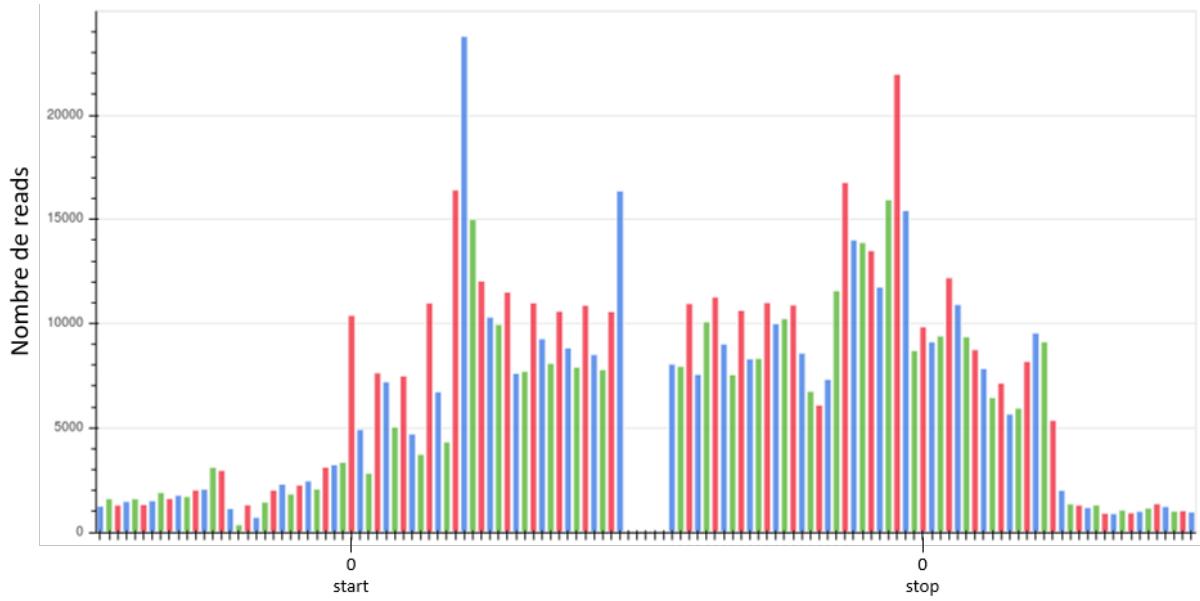


Figure 14 : Phasage des reads dans des cellules HeLa shFBL

Les positions du site A des ribosomes sont déterminées sur les reads de tous les gènes, puis sont poolées et comptées autour du start et du stop. Les trois phases de lecture par le ribosome sont représentées par trois couleurs : en rouge la phase 0, en bleu la phase 1 et en vert la phase 2.

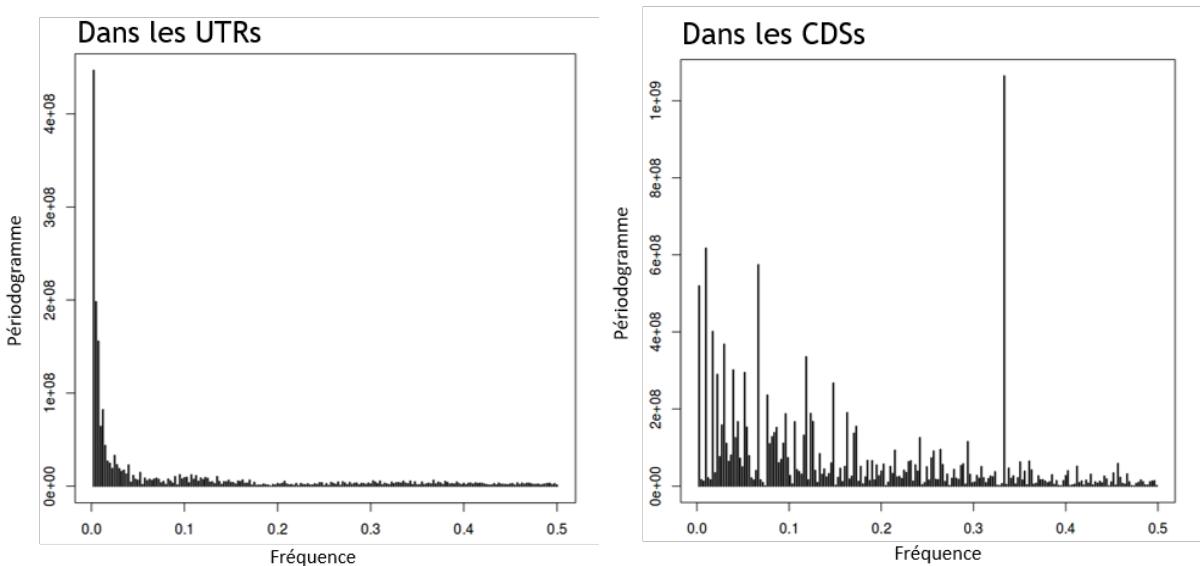


Figure 15 : Transformation de Fourier appliquée aux métagènes de ribosome profiling

La transformation de Fourier permet de dégager la périodicité d'un motif répété. Il n'y a aucun motif répété dans les régions UTRs et un motif répété à une fréquence de 0,33 dans les CDSs. Cette fréquence correspond à 3 nucléotides soit 1 codon.

transcriptomes et RiboSeq que j'ai réalisés présentent bien ces types de profil (non montré pour les transcriptomes, Figure 12 pour les RiboSeq).

De plus, comme la technique de ribosome profiling permet de déterminer la position du ribosome au nucléotide près et que le ribosome se déplace de trois nucléotides (= un codon) en trois nucléotides, l'alignement des reads doit pouvoir permettre de visualiser la phase dans laquelle le ribosome traduit comme cela est visible sur un jeu de données issue d'un RiboSeq réalisé précédemment dans l'équipe (Figure 13). En ce qui concerne mes expériences de RiboSeq, ce phasage n'est pas visible (Figure 14). Cependant, le profil des métagènes des UTR et de la CDS ont subi une transformation de Fourier, qui revient à décomposer le spectre des signaux en fréquences élémentaires. Ainsi, la transformation de Fourier permet de détecter la ou les périodicité(s) d'un ou plusieurs motif(s) répété(s). L'application de la transformation de Fourier n'a révélé aucune répétition de motif dans les UTRs, ce qui est attendu. Au niveau des CDS, elle a révélé la répétition d'un motif de 3 nucléotides, soit un codon (Figure 15). Cela permet de montrer que les RPFs que j'ai préparés proviennent bien de ribosomes en traduction sur les CDS.

Le comptage des reads permet d'évaluer l'expression différentielle de chaque gène entre deux conditions grâce à l'utilisation de l'outil DESeq2. DESeq2 a été utilisé sur les données des transcriptomes et des RiboSeq (définis aussi comme des traductomes), permettant de déterminer des listes de gènes différentiellement exprimés. Les paramètres définis pour déterminer qu'un gène est différentiellement exprimé sont, pour les transcriptomes, un $\log_2(\text{FC}) > 1$ ou < -1 (FC étant le "fold change", ou facteur de variation entre la condition d'hypométhylation des ARNr et la condition témoin) et une $p\text{value} < 0,05$; pour les traductomes, un $\log_2(\text{FC}) > 0,7$ ou $< -0,7$ et une $p\text{value} < 0,05$.

Les résultats des transcriptomes et des RiboSeq sont représentés sous la forme d'un graphique en croix, où la position de chaque gène est déterminée en abscisse par le fold change en transcriptomique et en ordonnée par le fold change en traductomique. En rouge sont représentés les gènes qui varient seulement en RNAseq, en vert les gènes qui varient seulement en RiboSeq et en orange ceux qui varient dans les deux. Le gène de la fibrillarine est le gène le plus touché négativement à la fois en RNAseq et en RiboSeq, ce qui est attendu car il est inhibé par le shFBL (Figure 16).

L'analyse par DESeq2 des données des transcriptomes a révélé 8 gènes dont les taux d'ARNm varient entre la condition témoin et la condition d'hypométhylation des ARNr : 5 sont enrichis dans la condition testée et 3 sont appauvris. Ce résultat montre que la baisse du taux de fibrillarine et la baisse de la méthylation des ARNr a un impact limité sur les taux des ARNm.

L'analyse des données des RiboSeq a révélé une liste de 46 gènes différentiellement exprimés. Parmi ces gènes, 26 présentent une régulation positive et 20 une régulation négative. De

FBL- vs FBL+

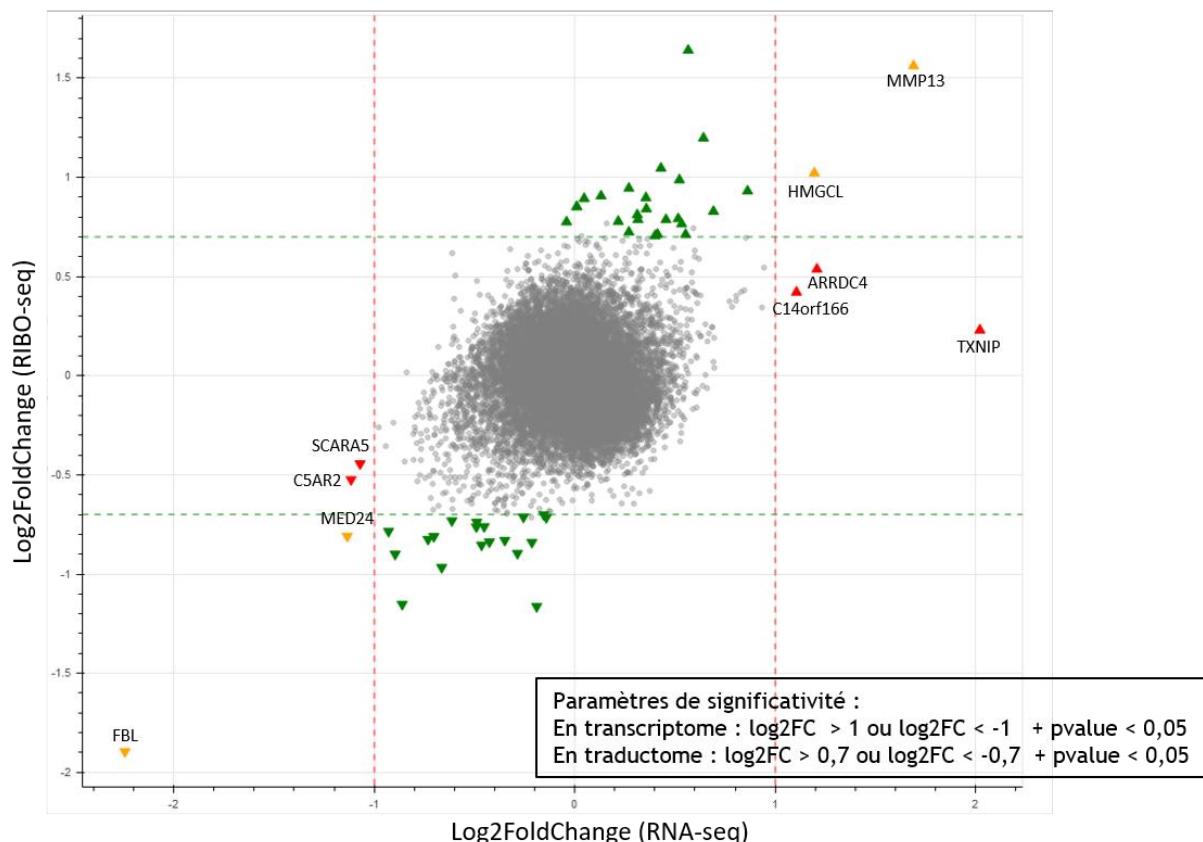


Figure 16 : Graphique en croix FBL- vs FBL+

La position de chaque gène est déterminée en abscisse par le fold change (FC) en transcriptomique et en ordonnée par le fold change en traductomique. En rouge sont représentés les gènes qui varient seulement en transcriptomique, en vert les gènes qui varient seulement en traductomique et en orange ceux qui varient à la fois en transcriptomique et en traductomique.

Les gènes FBL, MED24, MMP13 et TXNIP ont été testés en RT-qPCR.

plus, les gènes MMP13, HMGCL et MED24 présentent à la fois des variations en RNAseq et en RiboSeq (Figure 16). Ces variations d'expression sont probablement une conséquence d'une variation des taux d'ARNm. Les gènes TXNIP, C14orf166, ARRDC4, SCARA5 et C5AR2 montrent une variation du taux de leurs ARNm mais aucune variation au niveau traductionnel, ce qui indique que les régulations transcriptionnelles et post-transcriptionnelles sont compensées au niveau traductionnel (Figure 16). L'extinction de la fibrillarine et la baisse de la méthylation des ARNr entraînent donc une variation positive ou négative de l'efficacité de la traduction sur une liste réduite de 43 gènes (Figure 17).

Tout d'abord, afin de valider les résultats de RNAseq, j'ai utilisé une autre méthode indépendante, la RT-qPCR. J'ai choisi de tester les gènes FBL et MED24, régulés négativement, et MMP13 et TXNIP, régulés positivement. J'ai tout d'abord extrait les ARNs totaux de cellules HeLa shFBL traitées ou non à la doxycycline pendant 5 jours, puis j'ai effectué une rétro-transcription de ces ARNs pour obtenir des ADNc. Enfin j'ai procédé à la qPCR (Figure 18). Les résultats montrent une diminution significative du taux des ARNm de FBL et MED24 en condition d'hypométhylation des ARNr avec respectivement un rapport de 0,259 (0,215 en RNAseq) et 0,580 (0,477 en RNAseq), ainsi qu'une augmentation significative du taux de l'ARNm de TXNIP avec un rapport de 10,627 (4,219 en RNAseq). Concernant le gène MMP13, j'observe une augmentation du taux de son ARNm qui est statistiquement non significative (rapport de 6,716 en RT-qPCR et 4,130 en RNAseq), probablement dû à une variabilité importante des valeurs mesurées. Ces résultats confirment donc les résultats obtenus par l'analyse des données de RNAseq, et montrent que même si les variations observées sont faibles, elles sont significatives.

3 – Recherche d'éléments fonctionnels et/ou moléculaires communs aux gènes affectés

L'analyse de l'expression différentielle des données des RiboSeq a permis de dresser une liste de 43 gènes candidats dont la traduction est affectée quand les ribosomes sont hypométhylés. Afin de dégager un modèle du rôle des 2'-O-méthylation dans la régulation de la traduction, j'ai recherché s'il existait des éléments fonctionnels et/ou moléculaires communs à tout ou partie des gènes de cette liste.

❖ Analyse « Gene ontology » :

Dans un premier temps, j'ai réalisé une analyse « Gene Ontology » avec l'outil PANTHER, qui permet de classer les gènes suivant leurs fonctions moléculaires, les processus biologiques dans lesquels ils sont impliqués et leur localisation au sein de la cellule. J'ai cherché à démontrer si une

Gènes régulés positivement

Nom du gène	Log2(FoldChange)
RTN4RL2	0,706
FOXP2	0,712
IKZF3	0,713
WSCD1	0,724
RCOR2	0,765
ZNF467	0,775
MUC16	0,778
RP11-152F13,10	0,787
RP11-697E2,12	0,787
LOXL4	0,791
CNGB1	0,811
SLITRK5	0,828
UNC13A	0,841
GDF1	0,852
CERS1	0,852
ZFHX3	0,893
C4A	0,897
DDN	0,906
TNC	0,931
ELFN1	0,945
SETBP1	0,987
RUNX2	1,046
SLC22A23	1,198
COL6A3	1,641

Gènes régulés négativement

Nom du gène	Log2(FoldChange)
NKX3-2	-1,164
FGFBP1	-1,153
INSL4	-0,967
TRIML2	-0,9
AMTN	-0,896
MUC13	-0,854
GPR1	-0,84
HMGCS1	-0,838
MSMO1	-0,83
AKR1C2	-0,825
IL7R	-0,811
AKR1C1	-0,785
AOX1	-0,762
TSPAN8	-0,761
UNC5D	-0,739
NKD2	-0,731
RPL22L1	-0,715
CPM	-0,713
CGA	-0,701

Figure 17 : liste des gènes candidats de ribosome profiling

Ces listes regroupent les gènes dont l'expression traductionnelle varie entre la condition normale et la condition d'hypométhylation des ARNr. Le gène FBL et les gènes qui présentent une variation transcriptionnelle n'y figurent pas.

fonction moléculaire, un processus biologique ou une localisation cellulaire était enrichi dans les gènes candidats. Malheureusement, aucun enrichissement n'a été mis en évidence (Figure 19).

❖ Prédiction d'IRES :

On a vu que la baisse du taux de fibrillarine et de méthylation des ARNr entraînait une diminution de l'initiation IRES-dépendante. On peut donc faire l'hypothèse que parmi les gènes qui sont différemment traduits en condition d'hypométhylation des ARNr, on pourrait retrouver des gènes dont l'initiation s'effectue au niveau d'un IRES. J'ai comparé la liste de ces gènes avec une liste des gènes connus pour présenter des IRES dans leurs ARNm, fournie par nos collaborateurs Virginie Marcel et Frédéric Catez. Malheureusement, je n'ai retrouvé aucun gène contenant un IRES parmi les candidats. Je me suis alors orientée vers un outil de prédiction d'IRES fondé sur le calcul de la stabilité de structures secondaires de l'ARNm, IRESPred. J'ai choisi de m'intéresser principalement aux régions 5'UTR pour rechercher des IRES car la plupart des IRES sont situés dans cette région. De plus, la longueur importante des CDS complique la recherche des IRES, en diminuant la capacité de prédiction du logiciel utilisé. J'ai téléchargé les séquences des 5'UTR des gènes candidats pour les soumettre à IRESPred.

L'outil donne une réponse binaire : soit la séquence contient un IRES potentiel, soit elle ne contient aucun IRES prédict. Après avoir soumis toutes les séquences des 5'UTR (certains gènes présentent plusieurs 5'UTR), j'ai dressé la liste des gènes qui contiendraient potentiellement un IRES. Puis, afin de vérifier la robustesse des prédictions de IRESPred, je lui ai soumis des séquences de régions 5'UTR dont je sais qu'elles contiennent ou non des IRES. J'ai choisi les gènes de ménage ACTB, GAPDH, HPRT1 et HMBS, qui ne contiennent pas d'IRES, et les gènes IGF-1R, FGF1 et c-Myc, que je sais contenir des IRES et que j'ai déjà utilisé précédemment. Malheureusement, les tests effectués sur ces gènes ne donnent pas de résultats concluants, la moitié des résultats étant erronés. IRESPred n'est donc pas un outil de prédiction d'IRES assez fiable pour que je me repose sur les résultats qu'il fournit.

De plus, l'initiation de la traduction peut s'effectuer indépendamment de la coiffe sans faire intervenir un IRES. Une étude récente sur la recherche systématique de séquences d'initiation de la traduction indépendante de la coiffe suggère que quatre ARNm candidats seraient traduits par un tel mécanisme¹³³. Ces données sont cependant à confirmer et le mécanisme de recrutement des ribosomes reste à déterminer.

❖ Recherche de motifs nucléotidiques communs :

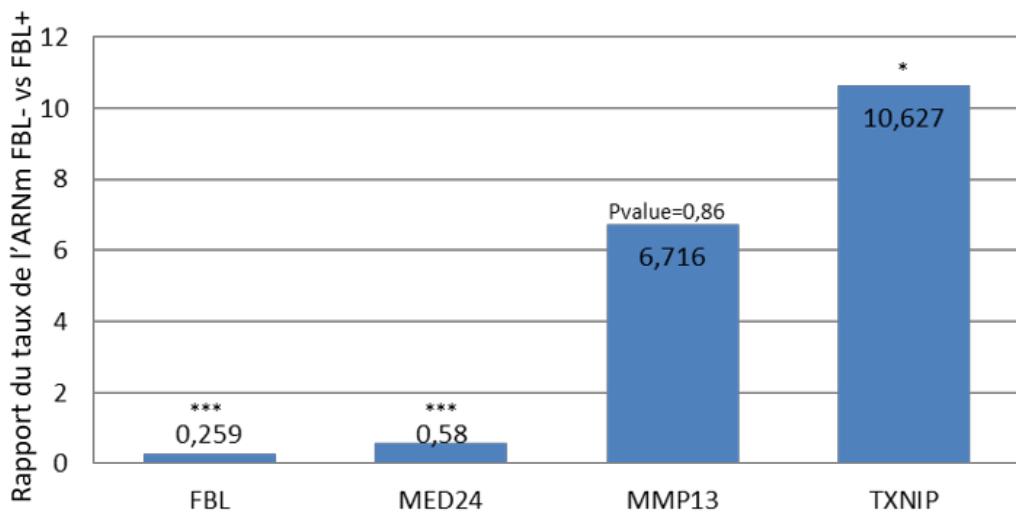


Figure 18 : RT-qPCR des transcrits des gènes FBL, MED24, MMP13 et TXNIP en condition normale et en condition d'hypométhylation des ARNr

Des RT-qPCR sont effectuées à partir d'ARNs totaux extraits de cellules HeLa shFBL traitées ou non avec 1 µg/mL de doxycycline pendant 5 jours, sur les transcrits de FBL, MED24, MMP13 et TXNIP.

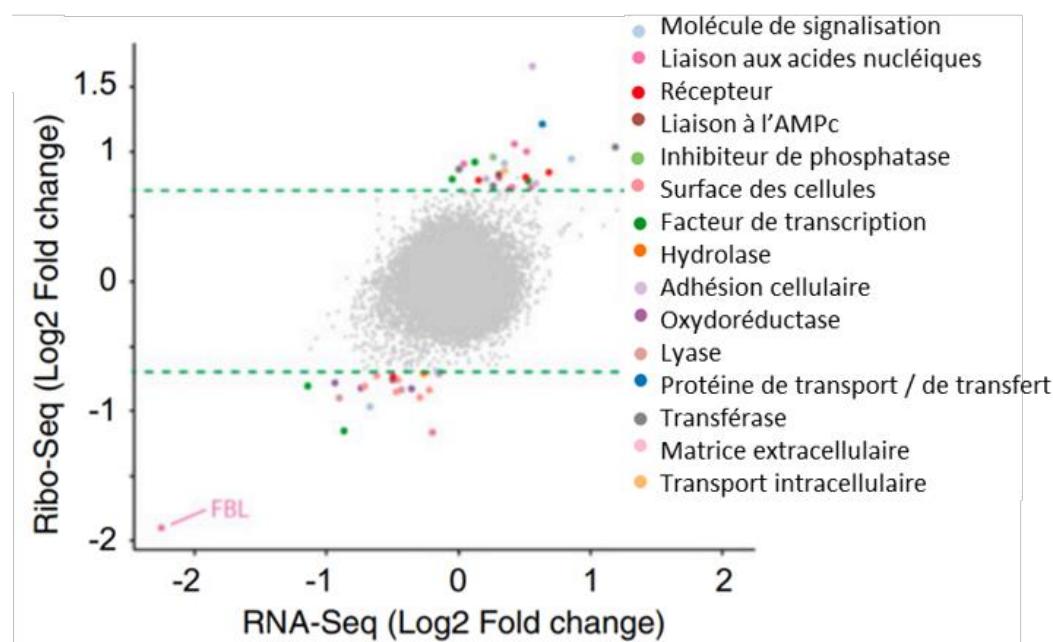


Figure 19 : Fonctions moléculaires des gènes candidats du ribosome profiling

Les fonctions moléculaires des gènes candidats de ribosome profiling, obtenues par une analyse « Gene Ontology », sont indiquées sur le graphique en croix présenté en figure 17.

Les ARNs ribosomiques peuvent interagir avec des séquences nucléotidiques des ARNm, régulant potentiellement ainsi l'initiation de la traduction^{133,304}. On peut émettre l'hypothèse qu'en cas d'hypométhylation des ARNr, certaines de ces interactions se trouvent altérées et la traduction des ARNm correspondants est affectée. Je recherche donc ensuite des motifs nucléotidiques dans les régions 5'UTR qui seraient partagés par plusieurs gènes candidats, avec l'outil MEME. Dans un premier temps, je teste l'ensemble des gènes d'intérêt. Deux motifs, d'une taille de 50 et 49 nucléotides, sont donnés avec une E-value significative. Cependant, lorsque l'on analyse ces motifs, on s'aperçoit que bien qu'ils soient significatifs, ils sont très longs, ne sont partagés que par 6 transcrits sur les 43 gènes et ne sont pas composés de nucléotides fortement représentés au sein des 6 gènes. Biologiquement, on ne peut pas considérer que ce sont des motifs pertinents. Je choisis par la suite de traiter séparément les groupes de gènes régulés positivement et négativement.

Les 5'UTR des ARNm des gènes régulés positivement présentent un enrichissement significatif pour 6 motifs, compris entre 46 et 50 nucléotides. Cependant, 5 de ces 6 motifs ne présentent pas de nucléotides fortement enrichis et ne sont donc pas pertinents. De façon surprenante, le 6ème motif est strictement identique dans le 5'UTR de deux gènes distincts. Ces deux gènes, GDF1 et CERS1, sont en fait présents sur le même transcrit, sous forme bicistronique. Le transcrit, qui présente un codon de terminaison prématûr - le codon stop de la première séquence codante de CERS1 - échappe à la voie de dégradation NMD car l'AUG de la deuxième séquence codante de GDF1 est présente sur le même exon que le codon stop de CERS1. Ce cas particulier n'apporte cependant aucune réponse à la question de savoir quel(s) effet(s) la baisse de la méthylation des ARNr a sur la traduction. De plus, la présence de deux séquences de 5'UTR identiques apporte un biais dans la recherche de motifs nucléotidiques enrichis. Je retire donc une des deux séquences et relance une recherche par l'outil MEME. Cette fois-ci, seuls deux motifs sont enrichis significativement mais ne sont partagés au maximum que par 6 transcrits et ne comportent pas de nucléotides présents dans 100% des transcrits de ces gènes.

De la même façon, les 5'UTR des ARNm des gènes régulés négativement présentent un enrichissement significatif pour 3 motifs, deux de 50 nucléotides et un de 36 nucléotides. Le motif de 36 nucléotides est riche en nucléotides U et C, qui sont fortement représentés dans les 4 transcrits qui partagent ce motif. Malheureusement, le nombre de transcrits qui présentent ce motif n'est pas suffisamment élevé pour dégager un lien entre ce motif et les 2'-O-méthylation des ARNr.

❖ Recherche de protéines de liaison à l'ARN (RBPs) :

Les gènes différentiellement traduits et leurs transcrits ne présentant jusqu'à présent aucune caractéristique commune, je cherche à trouver des mécanismes de régulation de la traduction qui

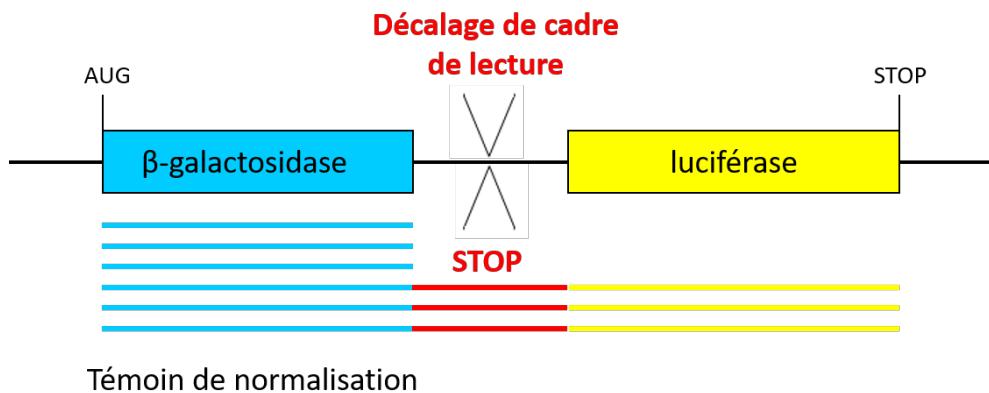


Figure 20 : Schéma du système double rapporteur β -galactosidase / luciférase

Ce système rapporteur est composé de la séquence codante de la β -galactosidase avec son codon d'initiation AUG mais sans codon stop, suivie de la séquence codante de la luciférase firefly sans codon d'initiation mais avec son codon stop. L'élément cible à tester, un codon stop ou un élément de décalage de cadre de lecture par exemple, est inséré entre les deux séquences codantes. Ainsi, tous les ribosomes synthétisent la β -galactosidase tandis que seuls ceux qui franchissent la cible traduisent la luciférase. La β -galactosidase est un témoin interne de normalisation et le rapport luciférase/ β -galactosidase permet une quantification du recodage au niveau de la cible.

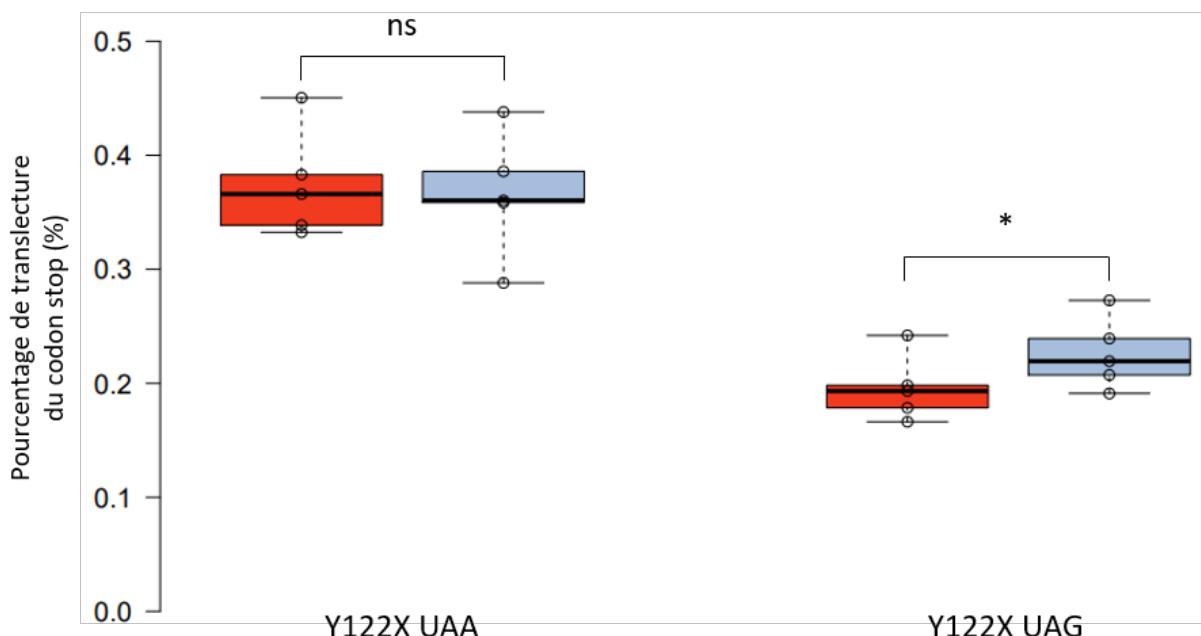


Figure 21 : Translecture des codons stop Y122X UAA et UAG dans des cellules HeLa shFBL traitées ou non à la doxycycline

Les taux de translecture des codons stop Y122X UAA et UAG sont calculés en rapportant l'activité de la luciférase à l'activité de la β -galactosidase, dans des cellules HeLa shFBL traitées (en bleu) ou non (en rouge) pendant 5 jours avec 1 μ g/mL de doxycycline. Un test statistique de Wilcoxon-Mann-Whitney a été appliqué sur les données. ns : non significatif ; * $p \leq 0,05$

pourraient être communs. La présence de protéines de liaison à l'ARN est un niveau de régulation intéressant à explorer. Pour cela, j'utilise l'outil RBPmap, qui recherche des motifs de liaison aux RBPs dans les séquences qui lui sont données. Je lui soumets alors toutes les séquences des 5'UTR des gènes candidats, et il me transmet une liste de RBPs qui sont susceptibles de se lier à chaque séquence. Je répertorie toutes les RBPs pour chaque ARNm des gènes d'intérêt. Afin d'éliminer les RBPs qui se lient à une quantité importante d'ARNm de façon non spécifique, j'effectue la même recherche sur des ARNm témoins, des gènes ACTB, GAPDH, HPRT1 et HMBS. Puis je soustrais les RBPs qui y sont associés des listes de RBPs des gènes d'intérêt. Il ne reste alors plus que quelques RBPs, qui ne sont associés qu'à quelques transcrits candidats, ce qui n'est pas suffisant pour établir un mécanisme général. La régulation de la traduction des transcrits des gènes candidats par les 2'-O-méthylation des ARNr ne semble donc pas faire intervenir des protéines de liaison aux ARNs RBPs.

4 – Impact sur la fidélité de la traduction

En parallèle, j'ai réalisé une étude plus spécifique des effets sur la fidélité de la traduction. La fidélité de la traduction peut se mesurer par le taux des événements de recodage que sont la translecture du codon stop et le décalage de cadre de lecture en -1 et en +1. Au laboratoire, un système double rapporteur qui permet de quantifier précisément ces événements a été développé et est utilisé en routine (Figure 20). Ce système comporte la séquence codante de la β -galactosidase dont l'initiation de la traduction a lieu au niveau du codon AUG canonique et qui ne présente pas de codon stop, ainsi que de la séquence codante de la luciférase, qui ne présente pas de codon d'initiation mais se termine par un codon stop. L'élément cible à tester, qui peut être un codon stop avec son contexte nucléotidique ou une région où se produit un décalage du cadre de lecture, est inséré entre les deux séquences codantes. Ainsi, tous les ribosomes synthétisent la β -galactosidase tandis que seuls ceux qui franchissent la cible traduisent la luciférase. La β -galactosidase est un témoin interne et le rapport luciférase/ β -galactosidase permet une quantification du recodage au niveau de la cible. J'ai donc utilisé ce système pour tester l'impact de l'hypométhylation des ARNr sur la translecture des trois codons stop UAA, UAG et UGA dans un contexte naturel, au sein d'une séquence codante (codon stop prématûre à l'origine d'une pathologie), le décalage de cadre de lecture -1 HIV (gène Gag-Pol du virus HIV) et le décalage de cadre de lecture présent dans le gène humain OAZ1. J'ai effectué des mesures dans des cellules HeLa shFBL traitées ou non à la doxycycline et réalisé le test statistique de Wilcoxon-Mann-Whitney sur ces mesures. Les résultats ne montrent aucun effet significatif sur la translecture du codon UAA (0,37% de translecture avec ou sans doxycycline) (Figure 21) et sur le décalage de cadre de lecture -1 HIV (3,12% sans doxycycline, 3,35% avec doxycycline) (Figure 23). En revanche, ils montrent une augmentation légère mais significative

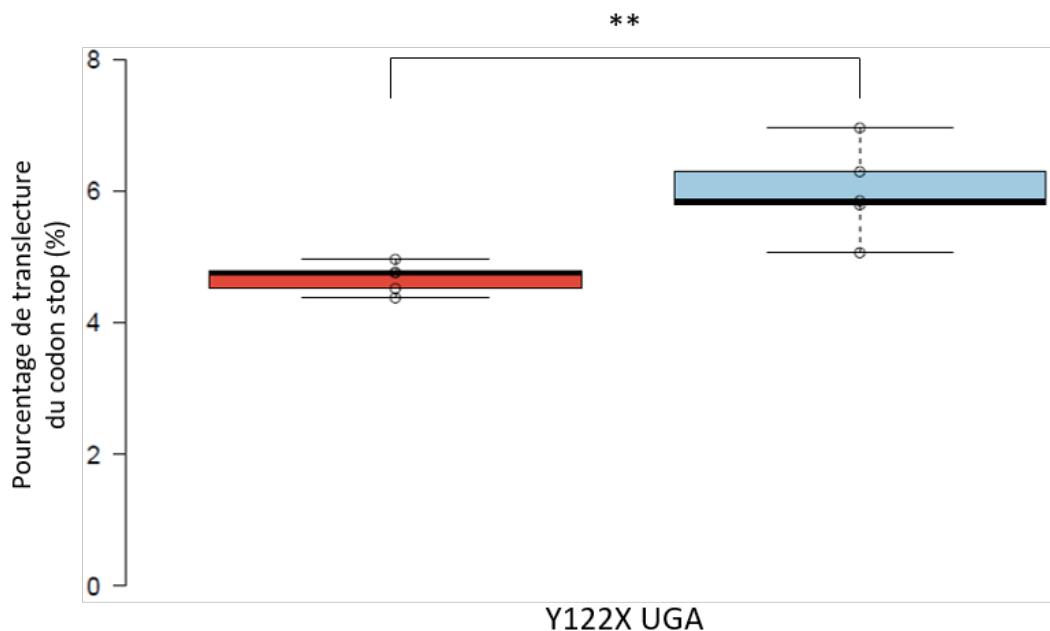


Figure 22 : Translecture du codon stop Y122X UGA dans des cellules HeLa shFBL traitées ou non à la doxycycline

Le taux de translecture du codon stop Y122X UGA est calculé en rapportant l'activité de la luciférase à l'activité de la β -galactosidase, dans des cellules HeLa shFBL traitées (en bleu) ou non (en rouge) pendant 5 jours avec 1 μ g/mL de doxycycline. Un test statistique de Wilcoxon-Mann-Whitney a été appliqué sur les données. ** $p \leq 0,01$

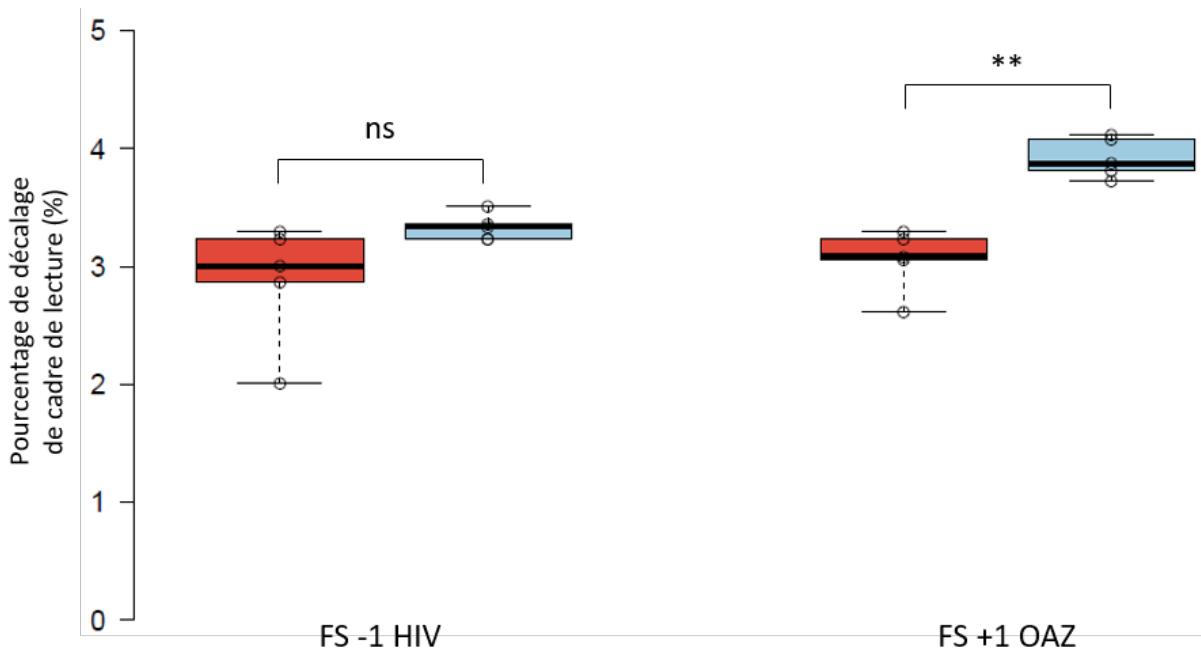


Figure 23 : Décalages de cadre de lecture -1 HIV et +1 OAZ dans des cellules HeLa shFBL traitées ou non à la doxycycline

Le taux des décalages de cadre de lecture -1HIV et +1OAZ sont calculés en rapportant l'activité de la luciférase à l'activité de la β -galactosidase, dans des cellules HeLa shFBL traitées (en bleu) ou non (en rouge) pendant 5 jours avec 1 μ g/mL de doxycycline. Un test statistique de Wilcoxon-Mann-Whitney a été appliqué sur les données. ns : non significatif ; ** $p \leq 0,01$

de la translecture des codons UAG (0,19% sans doxycycline et 0,23% avec doxycycline) (Figure 21) et UGA (4,77% sans doxycycline, 6,08% avec doxycycline) (Figure 22) ainsi que du décalage de cadre de lecture OAZ (3,07% sans doxycycline et 3,98% avec doxycycline) (Figure 23). Bien que faible, l'augmentation du taux de recodage est observée au niveau de plusieurs cibles, ce qui indique une diminution globale de la fidélité de la traduction par des ribosomes présentant des ARN hypométhylés. Cependant, l'augmentation du recodage n'a pas lieu au niveau de tous les codons stops ni de toutes les structures de décalage de cadre de lecture : elle est donc spécifique.

II – Discussion

La traduction peut être étudiée depuis quelques années de manière globale, à l'échelle du génome, grâce à la technique de ribosome profiling. Cette technique permet d'accéder à la fois à des données quantitatives en rendant compte du nombre de ribosomes en cours de traduction sur chaque transcrit, et à des données qualitatives par la position des ribosomes au nucléotide près. J'ai utilisé le ribosome profiling ainsi que le RNA-Seq en condition témoin (cellules HeLa shFBL non traitées à la doxycycline) ou en condition d'hypométhylation des ARNr (cellules HeLa shFBL traitées à la doxycycline pendant 5 jours) afin de déterminer si certains transcrits sont différentiellement traduits et si la fidélité de la traduction, à travers les événements de translecture et de décalage de cadre de lecture, est affectée.

La réalisation des RiboSeq et RNA-Seq a mis en évidence que la baisse des méthylations des ARNr a peu d'impact sur les taux d'ARN messagers – seuls 8 transcrits sont touchés – et entraîne une variation de l'expression traductionnelle de 43 gènes, dont 24 sont régulés positivement et 19 négativement. Ce résultat montre que seul un groupe restreint de gènes répond à la baisse de la méthylation des ARNr. Le mécanisme d'action des 2'-O-méthylation sur la traduction est donc spécifique à certains ARNs messagers.

A partir de cette liste de 43 gènes candidats, j'ai recherché des éléments fonctionnels et/ou moléculaires communs à plusieurs candidats.

L'analyse « Gene Ontology » réalisée sur ces gènes n'a permis de montrer aucun enrichissement fonctionnel. Néanmoins, une analyse manuelle réalisée par nos collaborateurs Virginie Marcel et Frédéric Catez suggère que les protéines synthétisées à partir des gènes candidats seraient en grande partie associées au réticulum endoplasmique et aux membranes. Cette piste reste aujourd'hui à explorer.

J'ai ensuite essayé de trouver des éléments moléculaires qui seraient enrichis dans les transcrits des gènes candidats. J'ai cherché la présence d'éventuels IRES, sans succès, l'outil de prédiction que j'ai utilisé n'étant pas suffisamment robuste. J'ai ensuite cherché si un ou plusieurs motifs nucléotidiques étaient enrichis dans les régions 5'UTR des transcrits des gènes candidats, mais je n'ai trouvé aucun motif. Bien que l'importance d'une interaction entre l'ARNr et des séquences des ARNm dans l'initiation de la traduction semble limitée aux procaryotes, il semblerait qu'il existe également de telles interactions chez les eucaryotes^{133,304}. Il pourrait donc être intéressant d'effectuer une recherche de séquences complémentaires à l'ARNr 18S au sein des transcrits des gènes candidats pour vérifier si ce type d'interaction pourrait faire intervenir les 2'-O-méthylation des ARNr.

Par la suite j'ai cherché si des protéines de liaison à l'ARN (RNA Binding Proteins ou RBPs) étaient retrouvées plus fréquemment associées aux transcrits des gènes candidats. J'ai donc listé, grâce à l'outil RBPmap, toutes les RBPs associées à tous les transcrits des gènes candidats, puis j'ai éliminé les RBPs qui s'associent de façon non spécifique à la majorité des transcrits de la cellule. Il ne reste alors seulement que quelques RBPs associées à un ou quelques transcrits, ce qui n'est pas suffisant pour permettre de dégager un mécanisme global d'action des 2'-O-méthylation des ARNr sur la traduction. Néanmoins, l'analyse spécifique et plus approfondie de ces RBPs pourrait potentiellement éclaircir une part du rôle des méthylations des ARNr dans la traduction.

Le ribosome profiling est une technique qui permet habituellement de réaliser une analyse qualitative des données en donnant accès aux positions des ribosomes au nucléotide près. Dans le cas de mon expérience, nous avons vu que le phasage, c'est-à-dire le fait que le ribosome traduit trois nucléotides par trois nucléotides suivant une phase de lecture déterminée, n'était pas directement observable. L'hypothèse la plus probable pour expliquer l'absence de phasage est une sous-digestion de l'ARNm par la RNase I, ce qui est à l'origine de RPFs plus longs d'un ou plusieurs nucléotides en 5' ou en 3'. Ce défaut de phasage est problématique car il empêche de déterminer la position au nucléotide près des ribosomes. Nous n'avons donc pas pu réaliser l'analyse qualitative que nous voulions faire, sur la traduction éventuelle de régions non annotées comme codantes (traduction dans la région 3'UTR après translecture du codon stop par exemple) au sein du génome entier. Cependant, j'ai réalisé en parallèle une étude sur quelques cibles de translecture et de décalage de cadre de lecture à l'aide de systèmes double rapporteurs afin d'évaluer l'impact de la baisse du taux de méthylation des ARNr sur la fidélité de la traduction. Dans une étude précédente, nos collaborateurs ont montré qu'une augmentation de la méthylation entraînait une augmentation de la translecture du codon stop, c'est-à-dire une diminution de la fidélité de la traduction¹⁰¹. En induisant une hypométhylation des ARNr, nous nous attendions également

à une altération de la fidélité de la traduction à travers une augmentation de la translecture du codon stop et/ou du décalage de cadre de lecture. Par les mesures que j'ai effectuées sur 3 codons stop prématurés et les éléments responsables des décalages de cadre de lecture -1 HIV et +1 OAZ, j'ai montré que les taux de translecture des codons UAG et UGA et de décalage de cadre de lecture +1 augmentaient tandis que la translecture du codon UAA et le décalage de cadre de lecture -1 ne sont pas affectés. Ces résultats révèlent que les ribosomes dont les ARNr sont hypométhylés sont moins fidèles que les ribosomes normaux, et que cette baisse de la fidélité est spécifique à certains événements.

Lorsque j'ai mesuré les taux de translecture et de décalage de cadre de lecture ainsi que l'initiation IRES-dépendante, j'ai effectué également des mesures sur des cellules HeLa traitées ou non à la doxycycline afin de m'assurer que cette molécule n'a pas d'effet sur la traduction. Les résultats préliminaires indiquent que dans certains cas, la doxycycline aurait un léger effet sur la traduction. Néanmoins ces effets sont à confirmer et s'ils s'avèrent importants, nous envisageons de refaire des mesures en parallèle sur des cellules HeLa et des cellules HeLa shFBL traitées à la doxycycline.

Afin d'élargir les spécificités d'impact de la baisse de la méthylation des ARNr sur la fidélité de la traduction, il pourrait être intéressant de tester la translecture d'autres codons stop sujets à la translecture programmée. L'exemple le plus connu et utilisé de translecture programmée est le stop TMV issu du virus de la mosaïque du tabac. Chez les mammifères et plus précisément chez l'homme, seuls quelques exemples sont connus aujourd'hui, comme les stops des transcrits des gènes VEGFA, MDH1 et LDHB²¹⁴. Il serait aussi intéressant de tester d'autres décalages de cadre de lecture, comme ceux qui ont lieu pour les transcrits des gènes PEG10 (FS-1)³⁰⁵ et IL-10 (FS+1)³⁰⁶.

Ces études globale et ciblée ont permis de révéler que la baisse du taux de méthylation des ARNr a un effet spécifique sur la traduction en affectant l'expression traductionnelle d'une liste restreinte de gènes et en augmentant la translecture de certains codons stop et le décalage de cadre de lecture +1. Les 2'-O-méthylations des ARNr semblent donc être impliquées dans la traduction spécifique d'un groupe d'ARNm et dans la fidélité de la traduction.

Bien que ces premiers résultats soient encourageants, nous n'avons pas pu dégager le mécanisme par lequel les 2'-O-méthylations des ARNr agissent sur la traduction. Je n'ai pas pu déterminer d'élément fonctionnel ou moléculaire commun aux gènes révélés par ribosome profiling et dont l'expression traductionnelle varie.

Une première hypothèse est qu'il est possible que je n'aie pas révélé uniquement les effets primaires causés par la baisse de la 2'-O-méthylation mais aussi des effets secondaires induits par les effets primaires. En effet, il semblerait que parmi les ARNm candidats de RiboSeq, quatre présentent une initiation indépendante de la coiffe^{133,302}. La traduction de ce sous-groupe d'ARNm pourrait être affectée de façon primaire et les protéines qui en résultent pourraient altérer à leur tour la traduction des autres ARNm de façon secondaire. Cependant, une telle analyse est compliquée et se heurte aux difficultés d'interroger les bases de données disponibles. De manière générale, après avoir testé des éléments assez évidents, comme des éventuels enrichissements fonctionnel et nucléotidique, il m'a été difficile de déterminer de nouveaux éléments à tester et d'interroger les bases de données.

Une deuxième hypothèse est que la stratégie utilisée pour induire la baisse de la méthylation des ARNr peut présenter des limites. En effet, l'extinction de la fibrillarine par un shRNA induit une baisse partielle du taux de 2'-O-méthylation. Il se peut que cette baisse ne soit pas suffisante pour induire des effets mesurables par les techniques que j'ai utilisées. De plus, il est très probable que les effets sur la traduction d'une baisse de la méthylation des ARNr soient faibles en conditions normales mais beaucoup plus importants en conditions de stress cellulaires, lorsque les cellules sont soumises à un agent pathogène ou au cours de la progression tumorale.

D'une part, les cellules au sein d'un organisme sont fréquemment soumises à un stress oxydatif dû à des espèces réactives de l'oxygène, les ROS. On sait que la production de ROS est impliquée dans la prolifération des tumeurs en induisant des lésions de l'ADN, en favorisant le processus de carcinogenèse et en activant directement des signaux intracellulaires impliqués dans le contrôle de la prolifération³⁰⁷. Il pourrait donc être intéressant de soumettre les cellules HeLa shFBL à un stress oxydatif en ajoutant des ROS dans le milieu de culture, et de chercher si l'hypométhylation des ARNr a un effet sur la réponse au stress oxydatif.

D'autre part, l'hypoxie correspond aussi à un stress cellulaire impliqué dans la tumorigénèse. L'hypoxie est la baisse de la quantité de dioxygène apporté aux tissus par le sang. La plupart des tumeurs solides présentent des zones hypoxiques. Les cellules tumorales s'adaptent à ces conditions en stabilisant les facteurs de transcription HIF (Hypoxia Inducible Factor), qui activent l'expression de nombreux gènes de survie et de maintien des fonctions cellulaires³⁰⁸. Ainsi, l'hypoxie tumorale induit une diminution de la sensibilité aux agents cytotoxiques, dont les agents utilisés en radiothérapie et en chimiothérapie. Il est donc très important de trouver un moyen de réduire la réponse des cellules cancéreuses en état d'hypoxie afin d'améliorer l'efficacité des traitements anticancéreux. Or il a été montré que le taux de fibrillarine et la méthylation des ARNr sont altérés dans certains cancers. Il pourrait donc être intéressant de regarder si l'hypométhylation des ARNr a un rôle dans la réponse des cellules à l'hypoxie.

Dans les cellules cancéreuses, l'activité de la protéine supresseur de tumeur p53 est très souvent altérée. En dehors de cette fonction très connue de p53, la protéine contribue également à la réponse immunitaire de l'hôte contre les infections par les virus de la stomatite vésiculaire (VSV), de la maladie de Newcastle (NDV) et de l'hépatite C (HCV). De plus, il a été montré que l'infection par le virus Influenza A (IAV) augmentait la transcription de p53, qui joue un rôle dans l'inhibition de la réplication du virus et régule l'apoptose des cellules infectées³⁰⁹. Or on sait que la protéine p53 régule négativement l'expression de la fibrillarine. Quand p53 est surexprimée, l'expression de la fibrillarine est donc réprimée, ce qui entraîne une diminution de la méthylation des ARNr. J'ai montré que la baisse de la méthylation des ARNr entraînait une diminution de la traduction IRES-dépendante et une augmentation de la translecture et du décalage de cadre de lecture. Les IRES et les décalages de cadre de lecture sont des éléments fréquemment présents chez les virus. Par exemple, le virus HCV présente un IRES³¹⁰ et les virus HCV, NDV et IAV présentent un décalage de cadre de lecture³¹¹⁻³¹³. Comme il a été montré qu'une infection par ces virus entraîne une réponse immunitaire qui fait intervenir p53, nous pouvons également émettre l'hypothèse que le taux de méthylation des ARNr est altéré par l'intermédiaire de p53. Il serait donc intéressant de chercher si l'infection par HCV, NDV ou IAV entraîne une hypométhylation des ARNr et si cette hypométhylation a un effet sur la traduction dépendante de l'IRES du virus et le décalage de cadre de lecture.

La réplication par des virus ADN nécessite une localisation nucléaire. Il a été montré que cette localisation nucléaire fait intervenir des interactions entre des protéines virales et des protéines nucléolaires, comme la fibrillarine. Ainsi, la fibrillarine interagit avec la protéine ORF3 du virus de la rosette de l'arachide GRV, avec la protéine NS1¹¹⁶ du virus Influenza A ou encore la protéine Tat du virus HIV⁸⁶. Par ces interactions, la protéine fibrillarine est en partie séquestrée et ne participe plus aux complexes snoRNPs pour la maturation des ARNr. Si cette séquestration est suffisamment importante, elle peut entraîner une diminution de la méthylation des ARNr, qui peut affecter la réplication du virus. Il pourrait être intéressant de vérifier si un tel mécanisme est mis en place dans les cellules infectées par un virus à ADN.

Résultats & Discussion

Article « Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes »

Résumé de l'article « Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes »

Des études récentes ont révélé que le ribosome lui-même pouvait être un niveau de régulation de la traduction à travers sa composition. En effet, la composition protéique des ribosomes peut varier ainsi que les profils de modifications chimiques des ARNr, ce qui serait à l'origine de spécificités de traduction pour certains ARNm. C'est le concept de « ribosome spécialisé », qui reste aujourd'hui à valider. Dans le cadre de ce contexte, mon équipe et nos collaborateurs avons cherché à caractériser la variabilité du taux de 2'-O-méthylation des ARNr et l'implication de ces modifications dans les capacités traductionnelles des ribosomes.

Dans le but d'affecter les 2'-O-méthylation des ARNr, l'expression de la fibrillarine a été inhibée dans des cellules HeLa par l'utilisation d'un siRNA, noté siFBL. Dans ces cellules, le clivage des ARNr présente des défauts, affectant la biogenèse des ribosomes tandis que l'incorporation des protéines ribosomiques n'est pas touchée. Environ 80% de la production normale des ribosomes est assurée.

Nous avons ensuite cherché quel est l'impact de la baisse de la fibrillarine sur les taux de 2'-O-méthylation. Tout d'abord, l'ajout des groupements méthyles ayant lieu sur le transcript primaire pré-ARNr, une expérience de marquage à la H3-méthyl-méthionine a été effectuée sur ce transcript. Elle a révélé une diminution globale de la méthylation du pré-ARNr. Puis nous avons caractérisé l'état des méthylation des ARNr au sein des ribosomes actifs grâce à la technique de RiboMethSequencing. Ces résultats sont déjà détaillés et discutés dans la partie « Induction de la baisse du taux des 2'-O-méthylation des ARNr ».

Après avoir validé que l'extinction de la fibrillarine par le siFBL induisait bien une baisse de la méthylation des ARNr, nous avons recherché les effets sur la traduction de la baisse des méthylation. Tout d'abord, nous avons réalisé une expérience de marquage des protéines au soufre S³⁵ qui a permis de montrer une baisse globale de la synthèse protéique. Cette baisse est causée par la diminution du nombre de ribosomes actifs due aux défauts de biogenèse des ribosomes induits par l'extinction de la fibrillarine.

Puis, j'ai effectué une analyse globale, à l'échelle du génome, de la traduction par ribosome profiling. Cette technique est fondée sur le séquençage à haut débit des fragments d'ARNm protégés par les ribosomes (Figure 24). Afin de s'affranchir de la variabilité de l'efficacité de la transfection des cellules HeLa par le siFBL, j'ai travaillé à partir des cellules HeLa qui expriment le shRNA dirigé contre l'ARNm de la fibrillarine. Le ribosome profiling a été effectué sur des cellules dans lesquelles le taux

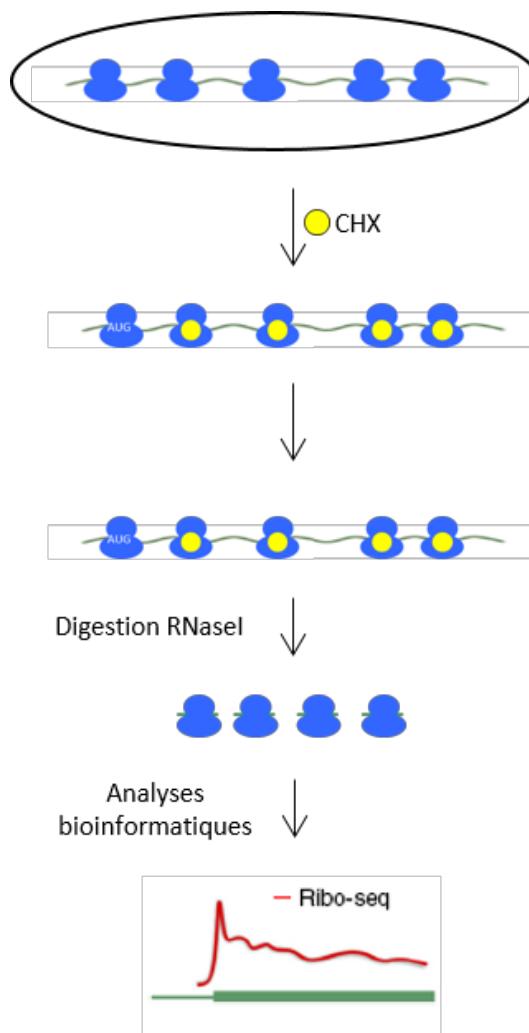


Figure 24 : La technique de ribosome profiling

La technique de ribosome profiling repose sur le séquençage à haut débit de fragments d'ARNm protégés par les ribosomes. Les ribosomes sont bloqués sur les ARNm par l'inhibiteur de traduction cycloheximide. Puis les ARNm sont extraits des cellules et digérés à la RNase I. Ensuite les fragments d'ARNm protégés par les ribosomes sont purifiés et séquencés. Le séquençage est suivi par des analyses bioinformatiques.

de fibrillarine est normal ou diminué. En parallèle, j'ai réalisé un transcriptome de ces cellules. Les résultats que j'ai obtenus montrent que la baisse du taux de fibrillarine et de méthylation des ARNr n'a que peu d'impact sur la quantité d'ARNm mais entraîne une variation de l'expression traductionnelle de 43 gènes. Afin de déterminer si ces gènes partagent une fonction commune, j'ai fait une analyse de « Gene Ontology » à partir de cette liste mais je n'ai pu mettre aucun enrichissement en évidence.

Ensuite, nous avons montré par des mesures de l'initiation IRES-dépendante avec des systèmes double rapporteurs une baisse de l'initiation au niveau des IRES cellulaires FGF1, IGF-1R et de l'IRES viral EMCV mais pas au niveau de l'IRES cellulaire VEGFA en cas de diminution du taux de fibrillarine. La baisse des taux de fibrillarine et de méthylation des ARNr entraîne donc une diminution de l'initiation IRES-dépendante spécifiquement au niveau de certains IRES cellulaires et viraux. Afin de vérifier que c'est bien uniquement la baisse de la méthylation des ARNr qui est à l'origine de cet effet et non d'autres facteurs, nous avons effectué de nouvelles mesures sur l'IRES du CrPV, qui ne nécessite la présence d'aucun facteur d'initiation de la traduction, et confirmé la baisse de l'initiation à cet IRES. Un autre moyen de vérification a consisté à récupérer les ribosomes hypométhylés de cellules transfectées par le siFBL pour les introduire dans un système de traduction *in vitro* afin de mesurer l'initiation des IRES viraux CrPV, DCV et EMCV et l'initiation canonique des transcrits de GAPDH et de la globine. Aucune variation de l'initiation de la traduction de GAPDH et de la globine n'est révélée tandis que l'initiation au niveau des IRES diminue. Ces résultats confirment que c'est bien uniquement la diminution de la méthylation des ARNr qui entraîne la diminution de l'initiation à certains IRES.

En utilisant une lignée cellulaire dans laquelle la synthèse de la protéine fibrillarine est inhibée par un siRNA ou un shRNA, nous avons montré que le taux de 2'-O-méthylation des ARNr pouvait être sujet à variation et que la cellule tolérait la production de ribosomes hypométhylés. Les 2'-O-méthylation sont donc une source d'hétérogénéité des ribosomes. Dans le cas d'une extinction de la fibrillarine, l'altération de la méthylation des ARNr est, de façon inattendue, site-spécifique. Les mécanismes impliqués dans cette spécificité restent à clarifier.

D'un point de vue fonctionnel, nous avons montré que la diminution des 2'-O-méthylation des ARNr a un impact sur l'initiation de certains IRES et entraîne une altération de l'expression traductionnelle d'une liste restreinte de gènes. Cet effet sur l'elongation de la traduction est à approfondir et les mécanismes moléculaires impliqués sont à élucider.



Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes

Jenny Eralas^{a,b}, Virginie Marchand^c, Baptiste Panthu^{b,d,e}, Sandra Gillot^f, Stéphane Belin^{a,b,1}, Sandra E. Ghayad^{a,b,2}, Maxime Garcia^{a,b}, Florian Laforêts^{a,b}, Virginie Marcel^{a,b}, Agnès Baudin-Baillieu^f, Pierre Bertin^f, Yohann Couté^g, Annie Adrait^g, Mélanie Meyer^h, Gabriel Therizols^{a,b,3}, Marat Yusupovⁱ, Olivier Namy^f, Théophile Ohlmann^{b,d,e}, Yuri Motorinⁱ, Frédéric Catez^{a,b,4,5}, and Jean-Jacques Diaz^{a,b,4,5}

^aCentre de Recherche en Cancérologie de Lyon, UMR, INSERM 1052, CNRS 5286, Centre Léon Bérard, F-69373 Lyon, France; ^bInstitut des Sciences Pharmaceutiques et Biologiques, Université de Lyon, F-69003 Lyon, France; ^cNext-Generation Sequencing Core Facility, FR3209 Bioingénierie Moléculaire Cellulaire et Thérapeutique, CNRS, Lorraine University, F-54505 Vandoeuvre-les-Nancy, France; ^dCentre International de Recherche en Infectiologie, Université de Lyon, 69364 Lyon, France; ^eÉcole Normale Supérieure de Lyon, 69324 Lyon, France; ^fInstitute for Integrative Biology of the Cell, Commissariat à l'Energie Atomique, CNRS, Université Paris-Sud, Université Paris-Saclay, 91190 Gif-sur-Yvette, France; ^gLaboratoire Biologie à Grande Échelle, Institut de Biosciences et Biotechnologies de Grenoble, Université Grenoble Alpes, Commissariat à l'énergie Atomique et aux Énergies Alternatives, INSERM, 38000 Grenoble, France; ^hInstitut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS UMR 7104 INSERM U 964, F-67404 Illkirch, France; and ⁱIngénierie Moléculaire et Physiopathologie Articulaire, UMR7365, CNRS, Université de Lorraine, F-54505 Vandoeuvre-les-Nancy, France

Edited by Ada Yonath, Weizmann Institute of Science, Rehovot, Israel, and approved October 31, 2017 (received for review May 9, 2017)

Ribosomal RNAs (rRNAs) are main effectors of messenger RNA (mRNA) decoding, peptide-bond formation, and ribosome dynamics during translation. Ribose 2'-O-methylation (2'-O-Me) is the most abundant rRNA chemical modification, and displays a complex pattern in rRNA. 2'-O-Me was shown to be essential for accurate and efficient protein synthesis in eukaryotic cells. However, whether rRNA 2'-O-Me is an adjustable feature of the human ribosome and a means of regulating ribosome function remains to be determined. Here we challenged rRNA 2'-O-Me globally by inhibiting the rRNA methyl-transferase fibrillarin in human cells. Using RiboMethSeq, a nonbiased quantitative mapping of 2'-O-Me, we identified a repertoire of 2'-O-Me sites subjected to variation and demonstrate that functional domains of ribosomes are targets of 2'-O-Me plasticity. Using the cricket paralysis virus internal ribosome entry site element, coupled to in vitro translation, we show that the intrinsic capability of ribosomes to translate mRNAs is modulated through a 2'-O-Me pattern and not by nonribosomal actors of the translational machinery. Our data establish rRNA 2'-O-Me plasticity as a mechanism providing functional specificity to human ribosomes.

2'-O-methylation | fibrillarin | ribosomal RNA | translational control | RNA epigenetics

Translational control is one of the most important regulators of gene expression (1). Translation is regulated through different mechanisms and coordinated with cell signaling. The best-described translational regulation pathways operate through non-ribosomal elements, such as the messenger RNA (mRNA) sequence and modification, canonical translation factors, transfer RNAs (tRNAs), micro RNAs (miRNAs), and RNA binding proteins (2, 3). Recently, several studies have provided compelling evidence that regulation of ribosomal proteins or ribosome biogenesis factors was associated with selective regulation of mRNA subsets (4–7). These observations led to the hypothesis of a ribosome-mediated translational control through functionally “specialized ribosomes.” However, direct molecular evidence that ribosomes displaying a different ribosomal RNA (rRNA) or protein composition carry different translational capabilities remains to be provided to validate the concept of specialized ribosomes.

In eukaryotes, rRNAs undergo 12 different types of chemical modification, on at least 112 (of 5,475 nt) and 212 (of 7,184 nt) nucleotides in yeast and human, respectively (8). However, despite being one of the best-characterized, the role of the rRNA epitranscriptome remains largely unknown. Among the different types of chemical modifications, 2'-O-methylation (2'-O-Me) is the most abundant modification of eukaryotic rRNA, with

55 and 106 sites mapped in yeast and in human rRNA, respectively (9, 10). In human rRNA, 2'-O-Me is carried out by the methyl transferase fibrillarin (FBL) associated with the

Significance

Translational control is a cornerstone of gene-expression regulation in physiological and pathological contexts. The contribution of nonribosomal factors, including messenger RNAs (mRNAs) and mRNA-bound factors, to translational control have been extensively studied. Recently, the hypothesis of a ribosome-mediated regulation emerged, which proposes that cells produce ribosomes of different composition and displaying different translational properties. This work reveals that ribosomal RNA 2'-O-methylation can be modulated in human ribosomes, including at key functional sites for translation, and that changes in the 2'-O-methylation pattern control the intrinsic capabilities of ribosomes to translate mRNAs. This work directly demonstrates the existence of composition-modified ribosomes and their associated change in translational activity as conceptualized by the specialized ribosome concept.

Author contributions: M.Y., O.N., T.O., Y.M., F.C., and J.-J.D. designed research; J.E., V. Marchand, B.P., S.G., S.B., S.E.G., M.G., F.L., V. Marcel, A.B.-B., Y.C., A.A., and F.C. performed research; M.M., G.T., M.Y., T.O., and Y.M. contributed new reagents/analytic tools; J.E., V. Marchand, B.P., S.B., S.E.G., P.B., Y.C., O.N., Y.M., F.C., and J.-J.D. analyzed data; J.E. and M.M. performed structural analysis; and J.E., V. Marchand, B.P., F.C., and J.-J.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository (accession no. [PXD00724](#)). RNA-Seq data of RiboMethSeq data have been deposited in the European Nucleotide Archive database (accession no. [PRJEB43738](#)). The RNA-Seq data of Ribosome Profiling have been deposited in the Gene Expression Omnibus (GEO) database, <https://www.ncbi.nlm.nih.gov/geo> (accession no. [GSE105248](#)).

¹Present address: Grenoble Institut des Neurosciences, University of Grenoble Alpes, F-38000 Grenoble, France.

²Present address: Department of Biology, Faculty of Science II, Lebanese University, Jdeidet, Lebanon.

³Present address: Department of Urology, University of California, San Francisco, CA 94158.

⁴F.C. and J.-J.D. contributed equally to this work.

⁵To whom correspondence may be addressed. Email: frederic.catez@lyon.unicancer.fr or jean-jacques.diaz@lyon.unicancer.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1707674114/DCSupplemental.

RNA-binding protein 15.5kDa and the core proteins NOP56 and NOP58. Methylation at each site is guided by small nucleolar RNAs (snRNAs) from the C/D box snoRNA family, which carry a complementary sequence to the target rRNA.

A significant amount of data supports an essential role for rRNA 2'-O-Me in ribosomal activity. About 70% of 2'-O-Me sites are conserved from yeast to human, particularly those located within functional regions of rRNAs (11). Studies using snoRNA knockout yeast strains revealed the importance of 2'-O-Me for the molecular functioning of ribosomes and for cell fitness, and their potential impact on rRNA folding. In yeast, inhibition of 2'-O-Me at several positions was required to severely impair translation and cell growth (12, 13). In contrast, inhibition of 2'-O-Me at single sites in zebrafish was sufficient to induce embryonic lethality, indicating that the role of individual 2'-O-Me is dependent on the cellular context (14). Finally, dysregulations in C/D box snoRNA gene expression have been linked to human diseases, including cancer or inherited genetic disorders, such as the Prader-Willy syndrome (15). The mechanisms by which C/D box snoRNAs adversely impact human cell behavior remain to be determined, and a link with their 2'-O-Me guiding activity and ribosomal function needs to be established, since an impact of snoRNAs on other cellular functions cannot be excluded. *FBL* (encoded by *NOPI* in yeast) is essential for rRNA 2'-O-Me in yeast and crucial for proper mouse development (16, 17). In addition, in yeast and mammals, *FBL* participates in pre-rRNA cleavage by association with C/D box snoRNAs, such as U3 or U14 (18), and regulates RNA *Pol I* activity on rDNA gene promoters by methylating a glutamine residue of histone H2A, by an unknown mechanism (19). *FBL* expression was recently shown to be highly modulated in physiological and pathological contexts, such as development (20), stem cell differentiation (21), viral infection (17), and cancer (7, 22). In cellular models of cancer, forced *FBL* up- or down-regulation modulated tumor progression (7). In addition, maintained expression of *FBL* in mouse embryonic stem cells prolonged their pluripotent state (21). In breast cancer cells, changes in *FBL* expression were correlated with alterations in the level of rRNA 2'-O-Me, with alterations in translational accuracy and with efficient translational initiation of mRNAs containing internal ribosome entry site (IRES) elements (7, 22, 23). However, due to the different activities of *FBL*, more data are needed to demonstrate that the effect of *FBL* modulation on translational activity is due to its impact on 2'-O-Me.

While the functional importance of 2'-O-Me is supported by genetic, developmental, cellular, and structural studies, whether the 2'-O-Me pattern represents an adjustable feature of ribosomes and a molecular basis of ribosome regulation is not yet determined. Initial proof supporting that 2'-O-Me could be modulated was provided in cellular models of breast cancer and in thalassemia patients using site-by-site analyses (7, 24, 25). However, a comprehensive view of 2'-O-Me within the four rRNAs, as well as a quantitative evaluation of the level of methylation at each site, is still missing. In the present study, we extensively characterize ribosomes following *FBL* down-regulation in HeLa cells. Using the recently developed RiboMethSeq approach, we show that the rRNA 2'-O-Me pattern can be qualitatively and quantitatively modulated. Mapping of the position of methylated nucleotides and their methylation frequency on the 3D structure of the human ribosome revealed an unsuspected 2'-O-Me plasticity within the critical functional domains of the ribosome, responsible for the ribosome translational activity. Using IRES-containing mRNAs as models coupled to hybrid in vitro translation assays, we demonstrate that the intrinsic capability of ribosomes to translate mRNAs is directly controlled by 2'-O-Me. Taken together, these studies establish rRNA 2'-O-Me and its plasticity as a molecular mechanism to regulate the translational activity of ribosomes.

Results

***FBL* Knockdown Decreases Ribosome Biogenesis and Global rRNA 2'-O-Me in Human Cells.** With the aim of altering global rRNA 2'-O-Me, we inhibited *FBL* expression in HeLa cells using small interfering RNA (siRNA). Transfection conditions were set

up to obtain a 5- to 10-fold *FBL* knockdown over a period of 5 d to enable ribosome turnover (Fig. S1A). The decrease in *FBL* did not induce a widespread disorganization of nucleoli or instability of major nucleolar proteins (Fig. S1B and C). *FBL* knockdown induced a clear, yet incomplete inhibition of the processing of the 5'-ETS region of the pre-rRNA, consequently inhibiting 18S rRNA maturation (Fig. S1D), an observation in agreement with previous studies on yeast *NOPI* and with the association of *FBL* with C/D box snoRNAs involved in pre-rRNA folding and cleavage (18). In contrast, the processing of 5.8S and 28S rRNAs was not affected by *FBL* knockdown. Consistently, ribosome biogenesis was sufficient to maintain ribosome production at ~80% of that of control cells (Fig. S1E). Since *FBL* participates in rRNA processing (Fig. S1D), we speculated that *FBL* knockdown could alter the assembly of ribosomal proteins (RPs). The assembly of newly synthesized ribosomal subunits appeared similar in *FBL* knockdown and control cells as evaluated using 2D-PAGE on ribosomes purified from isotope pulse-labeled cells (Fig. S1F). This observation was strengthened by label-free quantitative proteomics analysis, which showed no significant difference between ribosomes extracted from *FBL* knockdown cells compared with control cells (Fig. 1A and Dataset S1). Taken together, these findings indicate that *FBL* does not control the final stoichiometry of proteins in cytoplasmic ribosomes.

Next, we investigated the impact of a decrease in *FBL* on levels of rRNA 2'-O-Me. Because 2'-O-Me was shown to be an early and primarily cotranscriptional event (26, 27), we first analyzed methylation of the pre-rRNA by pulse labeling (Fig. 1B). *FBL* knockdown induced a 33.8% (± 19.2 , $P = 0.064$) decrease in the level of pre-rRNA methylation. Thus, as could be anticipated, knockdown of the rRNA methyl-transferase fibrillarin induced a global decrease in methylation of the pre-rRNA.

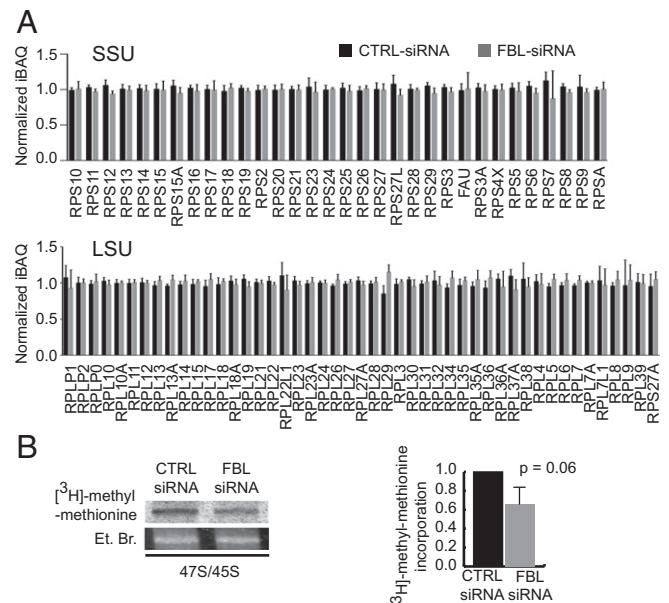


Fig. 1. *FBL* knockdown impacts rRNA 2'-O-Me and not ribosome protein composition in human cells. (A) Label-free quantitative proteomic analysis of 0.5 M KCl-purified cytoplasmic ribosomes from siRNA transfected cells. Normalized Intensity-based absolute quantification (niBAQ) values are shown for RPs of the small subunit (SSU, *Upper*) and the large subunit (LSU, *Lower*). Values are presented as mean \pm SD ($n = 5$) (see Dataset S1 for values). (B) Agarose gel electrophoresis (Left) of nuclear RNA purified from cells pulse labeled with [3 H]-methyl-methionine. The gels show the [3 H]-methyl-methionine incorporation in the 45S/47S pre-RNA (*Upper*), and the corresponding band stained with ethidium bromide as a loading control (*Lower*). The radioactive signal was normalized against the ethidium bromide signal (*Right*). Values are presented as mean \pm SD ($n = 2$). See also Fig. S1 and Dataset S1.

Altogether, these findings revealed that altering *FBL* expression in HeLa cells impacted ribosome biogenesis, notably rRNA maturation. However, although 2'-O-Me had decreased, the cytoplasmic ribosomes presented a normal protein composition.

***FBL* Knockdown Impacts 2'-O-Me of Nucleotides in a Site-Specific Manner, Including Nucleotides at Key Positions Within the Ribosome**

To identify potential site-specific methylation events and to quantify individual variations following *FBL* knockdown, we analyzed the methylation frequency of every nucleotide known to be ribose-methylated in human ribosomes (10). Several RNA-Seq-based 2'-O-Me mapping methods have been developed and used to refine the map of rRNA 2'-O-Me; however, these methods have so far not been applied to studying the dynamics of individual rRNA 2'-O-Me (9, 10, 28). We modified and applied our recently developed high-throughput RiboMethSeq technology (9) to human rRNA. RiboMethSeq is based on the protection of RNA hydrolysis provided by the methyl group, and on high-throughput sequencing to quantify the fraction of methylated nucleotides. The calculated MethScore represents the fraction of methylated rRNA at a given nucleotide in the ribosomal population (see *Materials and Methods* for details). We first established a reference map of rRNA 2'-O-Me in HeLa cells, using three independent biological replicates (Fig. 2). All of the 106 previously validated 2'-O-methylated nucleotides were highly methylated in rRNA of HeLa cells, except the 18S-Gm1447 nucleotide (MethScore = 0.09 ± 0.08). This was likely not due to a technical bias, since high MethScore values for this position were obtained in other cell lines. The majority of 2'-O-Me sites were methylated in over 80% of ribosomes, and only 16 sites (15%) were less-frequently methylated (MethScore ranging from 0.2 to 0.8). Among the sites conserved between yeast and human, all except one belonged to the highly methylated category (MethScore > 0.8), which is consistent with a high frequency of methylation of these nucleotides in yeast rRNA (9, 27). MethScore of individual sites displayed low dispersion among biological replicates, with a mean SD of 2.9%. Of the 106 known sites, 100 sites showed a level of variability below 5%, whereas only two sites in the 18S rRNA (Gm1447 and Cm174) and one in the 28S rRNA (Am4560) showed variability exceeding 10% (Fig. S2 and *Dataset S2*). This indicates that RiboMethSeq provides a robust measurement of 2'-O-Me levels. Altogether, these results demonstrate that the rRNA 2'-O-Me pattern is heterogeneous among human ribosomes.

RiboMethSeq was then applied to analyze the rRNA 2'-O-Me pattern upon *FBL* knockdown. Importantly, the transfection procedure did not introduce any experimental bias (Fig. S3A). Upon *FBL* knockdown, the frequency of 2'-O-Me decreased at almost all sites, although this variation was not statistically significant for all of the positions (Fig. 3A, Fig. S3 B and C, and *Dataset S2*). Surprisingly, the decrease in methylation was very different among sites, ranging from 0.2 to 57% (Fig. S3C), indicating that 2'-O-Me is likely

controlled in a site-specific manner rather than systemically. The level of methylation significantly decreased for 59 sites ($P < 0.05$; 10 in 18S rRNA, 1 in 5.8S rRNA and 48 in 28S rRNA) (*Dataset S2*). Interestingly, each site with an initial methylation level below 80% decreased by at least an additional 10% upon *FBL* knockdown, suggesting that partial methylation might render these sites more sensitive to *FBL* knockdown, or that they are intrinsically prone to variation (Fig. S3D). Of note, the decrease in 2'-O-Me was greater for 28S rRNA than for 18S rRNA, which we attributed to the lower turnover of the 18S rRNA in *FBL* knockdown cells (Fig. S1D).

Because a majority of 2'-O-methylated nucleotides are localized within functional domains of the rRNA, as evidenced by 2D maps of rRNAs (11), we investigated whether the nucleotides displaying an altered 2'-O-Me upon *FBL* knockdown were localized in particular domains within the ribosome structure. Each 2'-O-Me site was mapped on the 3D structure of the HeLa cell 80S ribosome recently obtained by cryo-EM (29), and was assigned a color based on the decrease in methylation in *FBL* knockdown cells, according to four different groups (Fig. 3 B and C and *Dataset S3*). Affected sites (yellow, orange, and red in Fig. S3E) were distributed throughout the ribosome structure, including in the “core” of the ribosome, the most conserved region compared with bacterial ribosomes (30). Strikingly, several affected 2'-O-Me sites were located in regions that are known to be involved in the translational process, in particular close to the A and P-sites, the intersubunit bridges, and the peptide exit tunnel (Fig. 3B and Fig. S3F), demonstrating that these important regions are subjected to variations in methylation. In contrast, 2'-O-Me sites close to the peptidyl transferase center (PTC) were not affected, indicating that this functional region might be protected from variations in methylation (Fig. 3C). The decoding center within 18S rRNA was also devoid of altered sites (Fig. S3G).

In conclusion, these data demonstrate that the down-regulation of *FBL*, a factor of the general ribose methylation machinery, induces site-specific modulation of the 2'-O-Me pattern. While several functional domains of the ribosome are subjected to 2'-O-Me variation, other key domains might be protected.

2'-O-Me Inhibition Selectively Modifies the Intrinsic Capability of Ribosomes to Initiate Translation from Dicistrovirus IRES Elements and Not from the m⁷G-Cap. To evaluate whether *FBL* knockdown impacts protein synthesis at a global level, we performed both a puromycylation assay (31), the signal produced by which represents the number of nascent peptides (Fig. S4 A and B), and an isotope pulse labeling with [³⁵S]-labeled amino acids, to evaluate the rate of amino acid incorporation (Fig. 4A and Fig. S4C). The results show a decrease in global synthesis of proteins, which indicates a reduction in the number of actively translating ribosomes, and is consistent with a decrease in ribosome production (Fig. S1E). Next, we sought whether *FBL* knockdown selectively altered mRNA translation. For this we applied ribosome profiling on HeLa cell lines expressing a *FBL* shRNA or a CTRL shRNA in an

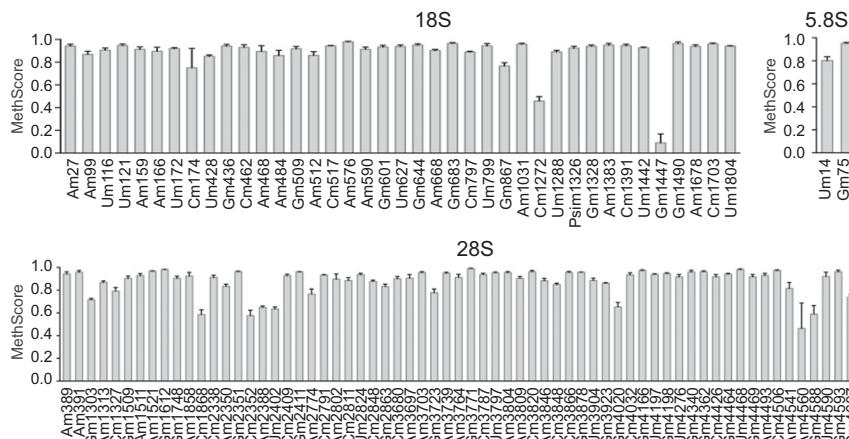


Fig. 2. Quantitative mapping of rRNA 2'-O-Me in human cells. The 2'-O-Me levels at each site of 18S, 28S, and 5.8S rRNA, evaluated by RiboMethSeq on nontreated HeLa cell rRNA. Data are expressed as mean MethScore values ± SD ($n = 3$ independent biological replicates) for each known methylated nucleotide in 18S, 5.8S, and 28S human rRNA. See also Fig. S2.

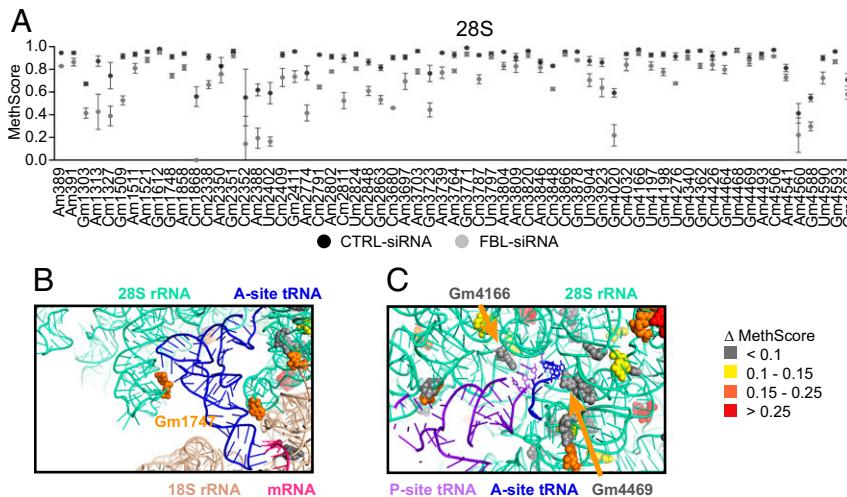


Fig. 3. *FBL* knockdown impacts 2'-O-Me of nucleotides in a site-specific manner, including nucleotides at key positions within the ribosome. (A) Mean MethScore values \pm SD ($n = 3$ independent biological replicates) for each methylated nucleotide in 28S rRNA from HeLa cells transfected with CTRL-siRNA (black circle) or FBL-siRNA (gray circle). (B) View of the A-site in a HeLa cell ribosome 3D structure. Methylation sites are color coded according to the variation in MethScore comparing FBL siRNA cells with CTRL siRNA cells, as indicated on the right. The Gm1747 methylation site (orange, methylation decreased by 16.7%), is oriented with the 2'OH group close to the D-loop of the A-site tRNA (blue). (C) View of the PTC showing the tRNAs in the A-site (blue) and P-site (purple). Methylation frequency of nucleotides Gm4469 and Gm4166 (Gray) was not altered by *FBL* knockdown. See also Fig. S3 and Data-sets S2 and S3.

inducible manner (Fig. S4 D, F, and G). Notably, change in the 2'-O-Me pattern, analyzed by RiboMethSeq, was similar after *FBL* knockdown induced by shRNA compared with the one induced by siRNA (Fig. S4E). Upon *FBL* knockdown, several genes were translationally altered (Fig. 4B). Translation efficiency of altered genes was either higher ($n = 28$) or lower ($n = 22$). This observation further supported that *FBL*, and possibly 2'-O-Me, could selectively regulate the translation efficiency of particular mRNA, although in this cellular model, there was no enrichment in particular molecular or cellular function (Fig. 4B).

Changes in *FBL* expression have been associated with alterations in IRES-dependent translation initiation (7, 22, 23). Within the subset of translationally altered mRNAs, 8% (four mRNAs) were previously identified in a large-scale screen for mRNAs able to drive Cap-independent translation (32). As a readout of changes in ribosome behavior, we analyzed IRES-dependent translational initiation *in cellulo* for a panel of cellular and viral IRESs using bicistronic constructs that code for two luciferases, the translation of which is either driven by the m⁷G-cap (*Renilla* luciferase) or by an IRES structure (firefly luciferase) (Fig. S4H). The firefly/*Renilla* ratio revealed that *FBL* knockdown induced a decrease in translation initiation from cellular IRESs of FGF1, IGF-1R, and from the type II encephalomyocarditis virus (EMCV) IRES, but not from VEGFA IRES (Fig. 4C). Consistently, luciferase activity/mRNA ratios, which reflect translation efficiency, showed a decrease in Cap-dependent translation consistent with the global protein synthesis reduction observed in *FBL*-siRNA cells (Fig. 4A), and a stronger decrease in IRES-dependent translation (Fig. S4I). Thus, *FBL* knockdown alters IRES-dependent translational initiation with a selective impact depending on the nature of the IRES.

To determine whether altering the pattern of rRNA methylation directly contributes to the *FBL*-induced reduction in IRES-dependent translation, we analyzed the impact of *FBL* knockdown on translation initiation using the cricket paralysis virus intergenic region (CrPV-IGR) IRES, which is able to trigger the assembly of an active 80S ribosome in the absence of any cellular translation initiation factor (33). First, we observed that translation initiation from the CrPV-IGR IRES was significantly reduced upon *FBL* knockdown compared with control cells using a bicistronic construct (Fig. 4D). Second, to consolidate these data and further exclude the contribution of other factors involved in translation, such as tRNA, mRNA, or miRNA, we analyzed the translational capability of ribosomes extracted from *FBL* knocked-down cells in a hybrid *in vitro* translation assay, which we developed recently (34) (Fig. 4E). In this assay all of the translation machinery components, except for the ribosomes, are provided as purified products so that the cell-extracted ribosomes are the only variable components (see *Materials and Methods* for details). In this context, translation initiation from a Cap-less mRNA containing the CrPV-IGR IRES was severely

impaired using ribosomes from *FBL* knockdown cells (Fig. 4F). In contrast, m⁷G-cap-driven translation from mRNAs containing the GAPDH or globin 5'UTR, was not significantly affected (Fig. 4F). In addition to the CrPV-IGR IRES element, translation from the *Drosophila* C virus (DCV) IRES, another dicistrovirus type IV IRES, and from the type II IRES EMCV was also strongly impaired (Fig. S4J). Any artifact due to nonspecific binding of mRNA to ribosomes was excluded by reproducing the experiment using a range of mRNA:ribosome ratios (Fig. S4 K and L). In conclusion, these data demonstrate that modulation of the 2'-O-Me pattern alters the intrinsic capability of ribosomes to initiate translation from IRES elements, but not from the m⁷G-cap structure of mRNAs.

Discussion

The most abundant modification in human rRNA, 2'-O-Me, is a highly complex and specific posttranscriptional modification, which is present in functionally important domains of the ribosome, indicating a significant contribution to ribosome functioning. However, existence of distinct 2'-O-Me patterns and the direct contribution of 2'-O-Me on the translational activity of ribosomes remain to be demonstrated. Here we show that rRNA 2'-O-Me patterns can be extensively modulated, although in a site-specific manner, including sites present in known functional regions of the ribosome, demonstrating that 2'-O-Me is a regulated, complex, and plastic feature of human ribosomes, and a molecular mechanism controlling ribosome functioning.

RiboMethSeq represents a unique method to simultaneously map and quantify 2'-O-Me on each site present in human rRNA, and was used here to explore the dynamics of 2'-O-Me. In HeLa cells, addition of 2'-O-Me appeared to be highly efficient since the majority of sites were methylated in almost 100% of the ribosomes. However, in contrast to yeast rRNA (9), a subset of sites was partially methylated, which has several conceptual implications: first, 2'-O-Me is not constitutively added at all sites in each ribosome; second, cells tolerate the production of ribosomes lacking some 2'-O-Me; and third, 2'-O-Me is a source of heterogeneity for the ribosomal population. In addition, a decrease in methylation was observed as a consequence of *FBL* knockdown (Fig. 3), and establishes 2'-O-Me as an adjustable and dynamic process, and a source of ribosome diversity. Subsequently, 2'-O-Me sites unaffected or weakly affected by *FBL* knockdown may represent sites for which methylation is highly efficient, or for which absence of methylation cannot be tolerated during ribosome biogenesis and subsequent quality control of ribosome fitness. The presence of 12 sites with a decrease in methylation exceeding 30% implies that *FBL* knocked-down cells contain ribosomes lacking 2'-O-Me at several sites. Consequently, 2'-O-Me should be considered and studied as a combination of sites, and

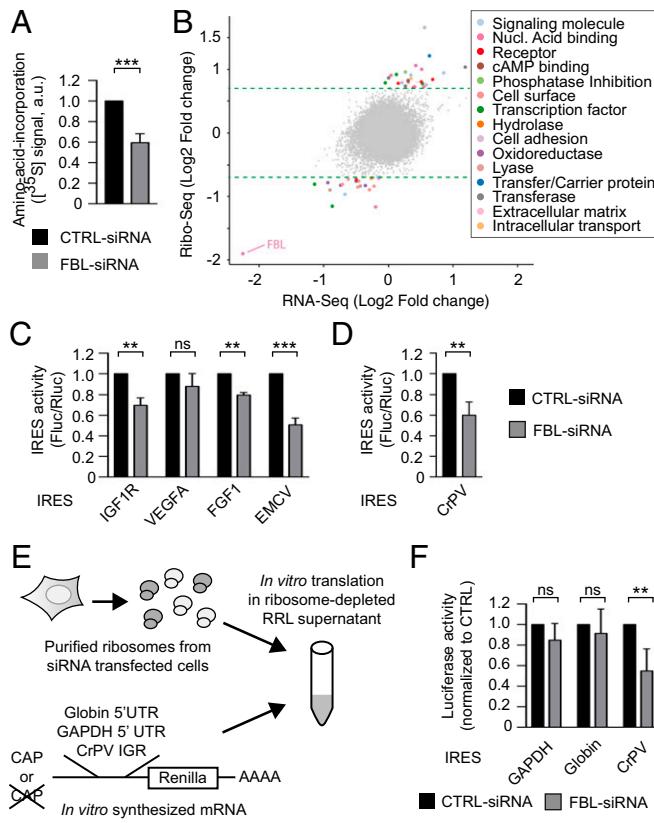


Fig. 4. 2'-O-Me inhibition selectively modifies the intrinsic capability of ribosomes to initiate translation from dicistrovirus IRES elements and not from the m⁷G-cap. (A) Global protein synthesis was measured by incorporation of [³⁵S]-methionine- [³⁵S]cysteine labeling following SDS/PAGE and counting of radioactive signals. Values are presented as mean \pm SD ($n = 4$). (B) Comparative mRNA translation by ribosome profiling on HeLa cells expressing either a CTRL-shRNA or a FBL-shRNA. Fold-changes at mRNA and translation levels are plotted along the y and the x axes, respectively. Translationally altered mRNA are colored according to their molecular function. Dotted green lines represent the significance threshold. (C) IRES-dependent translation (Fluc/Rluc) from several IRES elements was measured in HeLa cells transfected with CTRL-siRNA (black bars) or FBL-siRNA (gray bars). Values are presented as mean \pm SD ($n = 3$). (D) Identical experimental set-up as in C using a reporter construct carrying the CrPV-IGR IRES element. Values are presented as mean \pm SD ($n = 3$). (E) Schematic representation of the hybrid in vitro translation assay. (F) In vitro translation was evaluated by measuring luciferase activity produced with 1 μ g of ribosomes. Cap-dependent translation was evaluated using reporter constructs containing the 5'UTR of GAPDH or globin mRNA. IRES-dependent translation was evaluated using a Cap-less mRNA containing the CrPV-IGR IRES as 5'UTR. ns, not significant; ** $P \leq 0.01$; *** $P \leq 0.001$. See also Fig. S4.

not only individually, consistent with data obtained using snoRNA knockout yeast strains (12, 13).

FBL knockdown induced an unexpected site-specific alteration of 2'-O-Me (Fig. 3) by mechanisms that need to be further studied. Changes in single 2'-O-Me sites were not correlated with the global level of the corresponding snoRNA guide, further supporting that snoRNA expression by itself is not the main mechanism regulating 2'-O-Me (10). Possibly, the efficiency of methylation might be disproportionate among snoRNPs. The site-specific impact of *FBL* knockdown shows that modulating the expression of common components of the methylation machinery represents a means of regulating 2'-O-Me patterns. It follows that the steep down-regulation of *FBL* observed during neurogenesis and stem cell differentiation (20, 21) may affect rRNA 2'-O-Me patterns, with a direct impact on ribosome function. Conversely, overexpression of *FBL* in tumors and cancer cells might increase 2'-O-Me at selected sites, as suggested by our previous data (7). The moderate impact on ribosome production

(Fig. S1E) and absence of detectable consequences on ribosomal protein assembly and stoichiometry provides additional quantitative biochemical evidence that *FBL* regulates protein synthesis through its impact on 2'-O-Me plasticity (7). Therefore, *FBL* regulation may represent a means of modulating the 2'-O-Me pattern of rRNA without adversely impacting overall ribosome production.

In this study, we used translation initiation from Cap and IRES structures as functional assays to assess changes in behavior of ribosomes. The decrease in CrPV IRES activity in *in cellulo* and *in vitro* assays demonstrates that ribosomes with an altered 2'-O-Me pattern become intrinsically less efficient at initiating translation from IRES elements, in a manner independent of translation initiation factors. The decrease in EMCV IRES activity in our *in vitro* assay (Fig. S4J) reveals that 2'-O-Me impacts different types of IRESs, and further supports that 2'-O-Me is responsible for the *FBL*-dependent regulation of IRES-containing cellular mRNAs (7) (Fig. 4C). IRES elements recruit the 40S subunit through different interacting pathways involving eIF, but also ribosomal proteins, such as RPS25 (35). This raises the possibility that 2'-O-Me controls IRES translation via RPs, although our proteomic analysis demonstrates that 2'-O-Me alterations did not induce significant changes in RP composition, thus excluding that the decrease in IRES translation originated from a loss of RP, such as RPS25. Cap-independent translation of cellular mRNAs appears more widespread than anticipated, and comprises mechanisms based on direct interaction between mRNAs and 18S rRNA, in a Shine d'Algarno-like manner (32). Such a mechanism might thus be more sensitive to chemical modifications of rRNA. Importantly, rRNA 2'-O-Me provides selectivity to the translation machinery toward a subset of mRNAs (Fig. 4B and C). Additional studies are necessary to characterize mRNAs, the translation of which is regulated through rRNA 2'-O-Me.

The limited impact observed on translation from globin and GAPDH 5'UTR in the *in vitro* translation assay indicates that 2'-O-Me does not significantly modulate the ability of ribosomes to initiate Cap-dependent translation. This suggests that the decrease in global protein synthesis observed in cellular assays (Fig. 4A and Fig. S4C) is related to the lower ribosome production in *FBL* knockdown cells. Nevertheless, at this point we cannot exclude that 2'-O-Me affects some of the Cap-dependent pathways, and additional studies will be necessary to evaluate the impact of 2'-O-Me on the different mechanisms of Cap-dependent translation initiation. In addition, the limited impact of 2'-O-Me on Cap-dependent translation in the *in vitro* translation assay, also indicates that there was no major defect in translation elongation. Data from [³⁵S]-methionine- [³⁵S]cysteine pulse labeling, which reflect the rate of amino acid incorporation, and data from puromycylation assays, which reveal the number of ribosomes engaged in translation, both showed similar alterations upon *FBL* knockdown, and further indicate that elongation rate is similar in *FBL* knockdown cells compared with control cells. The impact of 2'-O-Me on synthesis of proteins, which are sensitive to translation elongation rate, remains to be studied. These observations unambiguously demonstrate that 2'-O-Me contributes to the translational activity of the ribosome.

The role of 2'-O-Me on ribosome structure and function is not known. Mapping of 2'-O-Me sites onto the ribosome structure revealed that 2'-O-Me can be modulated in several regions involved in intermolecular interactions, such as between tRNA and the A-site (Fig. 3B), intersubunit bridges (Fig. S3F), or around the peptide exit tunnel. The importance of 2'-O-Me at these locations was demonstrated in yeast and should now be explored in human models (11). Equally important are functional regions that did not display variations in 2'-O-Me. In particular, the 2'-O-Me sites of the PTC were unaltered, strongly indicating that this region is protected from 2'-O-Me variation. Therefore, our study supports the notion that 2'-O-Me comprises constitutively modified sites and regulated sites. How 2'-O-Me contributes to the molecular structure of the ribosome remains to be determined. Recent high-resolution crystal structures of the *Thermus thermophilus* ribosome and cryo-EM structures of human ribosomes, showed that the ribose 2'-O positions of several nucleotides are directly involved in

molecular interactions, both in a methylated and unmethylated state (29, 36). It can be anticipated that these interactions would be disrupted upon changes in methylation of these nucleotides, and may impact elongation and termination, in addition to initiation. However, technological advances in structural tools available today, such as cryo-EM and X-crystallography, are required to obtain a finer view of the structure–function relationship of human 2'-O-Me patterns.

In conclusion, 2'-O-Me plasticity reported herein expands the repertoire of ribosome composition and further demonstrates the existence of diversity in ribosome populations. The impact on the intrinsic ribosomal functioning establishes 2'-O-Me plasticity as a molecular mechanism modulating ribosomal activity, and further supports that modifications in rRNA chemical patterns, including pseudouridylation and base modifications, mediate ribosome functional specialization. These data expose the ribosomal RNA epitranscriptome as a new level of regulation of gene expression.

Materials and Methods

Detailed experimental procedures are described in *SI Materials and Methods*.

Ribosome Protein Composition. Ribosomes composition was analyzed by label-free quantitative proteomics as described previously (37, 38).

Analysis of rRNA Methylation. Site-specific rRNA methylation was determined by RiboMethSeq, as previously described (9).

Ribosome Structure Analysis. Methylated nucleotides were mapped on the cryo-EM structure of the human ribosome (PDB ID code 4UG0) (29). The reference structure of prokaryotic ribosome containing A-, P-, and E-site tRNAs plus mRNA was from *T. thermophilus* (PDB ID code 4V5C) (39).

1. Schwanhäusser B, et al. (2011) Global quantification of mammalian gene expression control. *Nature* 473:337–342.
2. Hinnebusch AG, Ivanov IP, Sonenberg N (2016) Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science* 352:1413–1416.
3. Hoerner TP, Erlacher MD (2017) Translating the epitranscriptome. *Wiley Interdiscip Rev RNA* 8:e1375.
4. Xue S, Barna M (2012) Specialized ribosomes: A new frontier in gene regulation and organismal biology. *Nat Rev Mol Cell Biol* 13:355–369.
5. Xue S, et al. (2015) RNA regulators in Hox 5' UTRs confer ribosome specificity to gene regulation. *Nature* 517:33–38.
6. Majzoub K, et al. (2014) RACK1 controls IRES-mediated translation of viruses. *Cell* 159: 1086–1095.
7. Marcel V, et al. (2013) p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. *Cancer Cell* 24:318–330.
8. Sharma S, Lafontaine DL (2015) 'View from a bridge': A new perspective on eukaryotic rRNA base modification. *Trends Biochem Sci* 40:560–575.
9. Marchand V, Blanloeil-Oillo F, Helm M, Motorin Y (2016) Illumina-based RiboMethSeq approach for mapping of 2'-O-Me residues in RNA. *Nucleic Acids Res* 44:e135.
10. Krogh N, et al. (2016) Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res* 44:7884–7895.
11. Decatur WA, Fournier MJ (2002) rRNA modifications and ribosome function. *Trends Biochem Sci* 27:344–351.
12. Liang XH, Liu Q, Fournier MJ (2007) rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol Cell* 28:965–977.
13. Baudin-Baillieu A, et al. (2009) Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids Res* 37:7665–7677.
14. Higa-Nakamine S, et al. (2012) Loss of ribosomal RNA modification causes developmental defects in zebrafish. *Nucleic Acids Res* 40:391–398.
15. Bortolin-Cavaillé ML, Cavaillé J (2012) The SNORD115 (H/MBII-52) and SNORD116 (H/MBII-85) gene clusters at the imprinted Prader-Willi locus generate canonical box C/D snoRNAs. *Nucleic Acids Res* 40:6800–6807.
16. Tollervey D, Lehtonen H, Jansen R, Kern H, Hurt EC (1993) Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell* 72:443–457.
17. Rodriguez-Corona U, Sobol M, Rodriguez-Zapata LC, Hozak P, Castano E (2015) Fibrillarin from Archaea to human. *Biol Cell* 107:159–174.
18. Watkins NJ, Bohnsack MT (2012) 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* 3:397–414.
19. Tessarz P, et al. (2014) Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature* 505:564–568.
20. Recher G, et al. (2013) Zebrafish midbrain slow-amplifying progenitors exhibit high levels of transcripts for nucleotide and ribosome biogenesis. *Development* 140:4860–4869.
21. Watanabe-Susaki K, et al. (2014) Biosynthesis of ribosomal RNA in nucleoli regulates pluripotency and differentiation ability of pluripotent stem cells. *Stem Cells* 32:3099–3111.
22. Su H, et al. (2014) Elevated snoRNA biogenesis is essential in breast cancer. *Oncogene* 33:1348–1358.
23. Basu A, et al. (2011) Requirement of rRNA methylation for 80S ribosome assembly on a cohort of cellular internal ribosome entry sites. *Mol Cell Biol* 31:4482–4499.
24. Belin S, et al. (2009) Dysregulation of ribosome biogenesis and translational capacity is associated with tumor progression of human breast cancer cells. *PLoS One* 4:e7147.
25. Sornjai W, et al. (2017) Hypermethylation of 28S ribosomal RNA in beta-thalassemia trait carriers. *Int J Biol Macromol* 94:728–734.
26. Kos M, Tollervey D (2010) Yeast pre-rRNA processing and modification occur co-transcriptionally. *Mol Cell* 37:809–820.
27. Birkedal U, et al. (2015) Profiling of ribose methylations in RNA by high-throughput sequencing. *Angew Chem Int Ed Engl* 54:451–455.
28. Incarnato D, et al. (2016) High-throughput single-base resolution mapping of RNA 2'-O-methylated residues. *Nucleic Acids Res* 45:1433–1441.
29. Khatter H, Myasnikov AG, Natchiar SK, Klaholz BP (2015) Structure of the human 80S ribosome. *Nature* 520:640–645.
30. Ben-Shem A, et al. (2011) The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* 334:1524–1529.
31. David A, et al. (2012) Nuclear translation visualized by ribosome-bound nascent chain puromylation. *J Cell Biol* 197:45–57.
32. Weingarten-Gabbay S, et al. (2016) Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science* 351:aad4939.
33. Fernández IS, Bai XC, Murshudov G, Scheres SH, Ramakrishnan V (2014) Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* 157:823–831.
34. Panthu B, Décimo D, Balvay L, Ohlmann T (2015) In vitro translation in a hybrid cell line free lysate with exogenous cellular ribosomes. *Biochem J* 467:387–398.
35. Hertz MI, Landry DM, Willis AE, Luo G, Thompson SR (2013) Ribosomal protein S25 dependency reveals a common mechanism for diverse internal ribosome entry sites and ribosome shunting. *Mol Cell Biol* 33:1016–1026.
36. Polikanov YS, Melnikov SV, Söll D, Steitz TA (2015) Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. *Nat Struct Mol Biol* 22:342–344.
37. Penzo M, et al. (2015) Human ribosomes from cells with reduced dyskerin levels are intrinsically altered in translation. *FASEB J* 29:3472–3482.
38. Casabona MG, Vandebrouck Y, Attree I, Couté Y (2013) Proteomic characterization of *Pseudomonas aeruginosa* PAO1 inner membrane. *Proteomics* 13:2419–2423.
39. Voorhees RM, Weixlbaumer A, Loakes D, Kelley AC, Ramakrishnan V (2009) Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat Struct Mol Biol* 16:528–533.
40. Bash-Imam Z, et al. (2017) Translational reprogramming of colorectal cancer cells induced by 5-fluorouracil through a miRNA-dependent mechanism. *Oncotarget* 8:46219–46233.
41. Jack K, et al. (2011) rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol Cell* 44:660–666.
42. Baudin-Baillieu A, et al. (2014) Genome-wide translational changes induced by the prion [PSI⁺]. *Cell Rep* 8:439–448.
43. Mi H, et al. (2017) PANTHER version 11: Expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res* 45:D183–D189.

Supporting Information

Erales et al. 10.1073/pnas.1707674114

SI Materials and Methods

Cell Culture and siRNA Transfection. HeLa cells (ATCC) were grown in Minimum Essential Medium supplemented with 10% FBS, glutamine, and nonessential amino acids at 37 °C with 5% CO₂. For siRNA experiments, three siRNA duplexes were used for fibrillarin silencing: 5'-GUCUCAUUUGUCGAGGAA-AdTdT-3'; 5'-UGGAGGACACUUUGUGAUUUdtdT-3'; and 5'-CUGUCAGGAUUGCAGAGAGAdTdT-3'.

The control siRNA does not target any human sequence (negative control siRNA duplex; Eurogentec). HeLa cells were transfected using the X-tremeGENE siRNA reagent (Roche) according to the manufacturer's instructions. Seventy-two hours after siRNA transfection, cells were plated according to future analyses.

Transfection and Plasmid Construction. For siRNA experiments, plasmids containing the bicistronic IRES luciferase constructs were described previously (7, 41).

Dual Luciferase Assays for in Cellulo Translation Assays. pIRES-FGF1, pIRES-EMCV, and pIRES-VEGFA were donated by A. C. Prats, Institut des Maladies Métaboliques et Cardiovasculaires, Toulouse, France and pRF-CrPV was a gift from D. Ruggero, University of California, San Francisco. Luciferase assays were performed 24 h after transfection with the reporter plasmids using X-tremeGENE 9 reagent (Roche). Dual luciferase assays were performed using the Dual-Glo luciferase reagent (Promega) according to the manufacturer's instructions, and a Tecan M1000 plate reader. IRES translation initiation is the ratio of background subtracted signal of Firefly luciferase over *Renilla* luciferase.

Global Protein Synthesis. Global protein synthesis analysis by puromycylation followed by puromycin detection was performed essentially as in ref. 31. Puromycin incorporation was detected by Western blot on whole-cell protein extracts. Global protein synthesis by isotope labeling was performed as previously described (46) by incubating cells for 30 min, with a [³⁵S]-methionine-[³⁵S]cysteine mix. Proteins were separated by SDS/PAGE on a 4–15% gradient polyacrylamide gel (Bio-Rad).

Real-time quantitative RT-PCR. Two-hundred nanograms of total RNA were reverse transcribed using the MMLV RT kit and random primers (Invitrogen), according to the manufacturer's instructions. Quantitative real-time PCR (RT-qPCR) was carried out using the Light cycler 480 II real-time PCR thermocycler (Roche). Expression of mRNAs was quantified using LightCycler 480 SYBR Green I Master Mix (Roche) and normalized using HPRT1 expression according to the 2- $\Delta\Delta$ C_t method.

Primers were as follows: LucR-Fwd 5'-AACGCGGCCCTCT-TCTTATT, LucR-Rev 5'-ACCAAGATTGCGCTGATTG, LucF-Fwd 5'-AACACCCCAACATCTTCGAC, LucF-Rev 5'-TTTCCGTATCGTCTTCC, HPRT1-Fwd 5'-TGACACTG-GCAAAACAATGCA, HPRT1-Rev 5'-GGTCCTTTCACCAAGCT.

Immunofluorescence analysis. Cells were grown on glass coverslips, fixed in 4% of paraformaldehyde in PBS before permeabilization with 0.5% Triton X-100 in PBS. Fibrillarin and nucleolin were detected using the anti-FBL rabbit polyclonal antibody (ab5821; Abcam) diluted at 1:2,000, and the anti-NCL mouse monoclonal antibody (ab13541; Abcam) at 1:4,000. Secondary antibodies were labeled with AlexaFluor488 or AlexaFluor555 (Molecular Probes), and used at 1:1,000. Coverslips were mounted using the

Fluoromount G mounting medium (EMS). Images were acquired on a Nikon NiE fluorescence microscope using a 60 \times Plan Apochromat immersion objective (NA 1.4) and a Flash 4.0 CMOS camera (Hammamatsu).

Ribosome purification. Ribosomes were purified as previously described (24). Briefly, cytoplasmic fractions were obtained by mechanical lysis of cells with a Dounce and centrifugation at 12,000 \times g for 10 min to pellet mitochondria. Cytoplasmic fractions were loaded onto a 1 M sucrose cushion in a buffer containing 50 mM Tris-HCl pH 7.4, 5 mM MgCl₂, 500 mM KCl, and 2 mM DTT, and centrifuged for 2 h at 240,000 \times g. The pellet containing the ribosomes was suspended in a buffer containing 50 mM Tris-HCl pH 7.4, 5 mM MgCl₂, and 25 mM KCl.

Ribosome Production. Ribosome production was measured as described previously (24). Briefly, cells were incubated for 1 h in methionine-cysteine-free DMEM supplemented with a [³⁵S]-methionine-[³⁵S]cysteine mix (GE Healthcare). Incorporation of radioactive amino acids was measured from one unit (OD_{260 nm}).

Two-dimensional gel electrophoresis. For 2D gel electrophoresis, 5 OD_{260 nm} units of ribosomes extracted from [³⁵S]-methionine-[³⁵S]cysteine-labeled cells were used. The method was previously described (24). Briefly, ribosomal proteins were extracted using acetic acid, extensively dialyzed against 1 M acetic acid and lyophilized. After lyophilizing, proteins were solubilized and reduced in 6 M guanidine hydrochloride, 0.5 M Tris-HCl pH 8.5 and 10 mM DTE. Proteins were then alkylated by adding 40 mM iodoacetamide in 6 M guanidine hydrochloride and 0.5 M Tris-HCl pH 8.5. Proteins were lyophilized and solubilized in sample buffer (20 mM Tris-boric acid pH 8.3, 1 mM EDTA and 8 M urea).

In the first dimension, proteins were separated according to their electric charge, in 4% polyacrylamide gels containing 0.2 M Tris-boric acid pH 8.6, 8 M urea, and 10 mM EDTA, placed in glass tubes. At the end of the first dimensional run, gels were extracted from the tube and equilibrated for 5 min in 0.625 M Tris-HCl pH 6.8, 2% SDS, and 0.002% Bromophenol blue. For the second dimension, the gels from the first dimension were placed in a 1.5-mm-thick gel cast made of 14% polyacrylamide (37.5:1 acrylamide: *N,N'*-methylenebisacrylamide). Protein separation was achieved using standard SDS/PAGE conditions. Proteins were stained by 0.1% Coomassie brilliant blue R 250. Gels were dried and exposed to PhosphoImaging screen. All chemicals were purchased from Sigma-Aldrich.

Mass spectrometry-based proteomic analyses of purified ribosomes. Label-free quantitative proteomics has been performed as in Casabona et al. (38). Ribosomes purified at 500 mM KCl were solubilized in Laemmli buffer, stacked in the top of a 4–12% NuPAGE gel (Invitrogen), and stained by R-250 Coomassie blue (Bio-Rad). Gel bands were excised and in-gel proteins were digested using trypsin (Promega). Resulting peptides were analyzed by nanoliquid chromatography coupled to tandem mass spectrometry (Ultimate 3000 coupled to LTQ-Orbitrap Velos Pro; Thermo Scientific) using a 120-min gradient. RAW files were processed using MaxQuant v1.5.3.30. Spectra were searched against the SwissProt database (*Homo sapiens* taxonomy, October 2016 version) and the frequently observed contaminants database embedded in MaxQuant. Trypsin was chosen as the enzyme and two missed cleavages were allowed. Precursor mass error tolerances were set respectively at 20 ppm and 4.5 ppm for first and main searches. Fragment mass error tolerance was set to 0.5 Da. Peptide modifications allowed during

the search were: carbamidomethylation (C, fixed), acetyl (Protein N-ter, variable) and oxidation (M, variable). Minimum peptide length was set at seven amino acids. Minimum number of peptides, razor + unique peptides and unique peptides were all set at 1. Maximum false-discovery rates (FDR)—calculated by employing a reverse database strategy—were set at 0.01 at peptide and protein levels. iBAQ values were calculated from MS intensities of unique + razor peptides and used for statistical analyses using ProStar. Proteins identified in the reverse and contaminant databases or exhibiting less than 5 iBAQ values in one condition were discarded from the matrix. After \log_2 transformation, iBAQ values were median-normalized, missing data imputation was carried out (replacing missing values by the 2.5-percentile value of each column) and statistical testing was conducted using the *limma* *t* test. Differentially expressed proteins were sorted out using a \log_2 (fold-change) cut-off of 2, and a FDR threshold on *P* values of 1% using the Benjamini-Hochberg method.

For generation of Fig. 1C, iBAQ values from ribosomal proteins were sorted out before column-wise normalization; to facilitate representation (data centering on 1), for each ribosomal protein the normalized iBAQ values were divided by the mean of the 10 values obtained.

Western blot analysis. Twenty micrograms of total protein lysates were run on a 12% SDS polyacrylamide gel and transferred onto a nitrocellulose membrane. The membrane was blocked with 3% nonfat milk in TBST. The antibodies used were the following: fibrillarin (ab166630; Abcam) at 1:2,000, Dyskerin (sc-48794; Santa Cruz) at 1:500, nucleophosmin (sc-5564; Santa Cruz) at 1:500, puromycin (clone 12D10; EMD Millipore) at 1:4,000, and Ku80 (ab3715; Abcam) at 1:2,000. Antibodies were incubated for 1 h in 1% milk-TBST. Proteins were detected by chemiluminescence using an anti-rabbit or anti-mouse peroxidase-conjugated antibody (Cell Signaling) diluted 1:10,000, and Clarity ECL substrate (Bio-Rad). Images were collected on a ChemiDoc XRS+ (Bio-Rad) and the signal analyzed using the Bio-Rad ImageLab software.

Northern blot analysis. Northern blot was performed as described in Belin et al. (24). The probes were obtained by oligonucleotide synthesis (Eurogenetec): ETS1-1399-5'-CGCTAGAGAAGGGCTT-TTCTC-3'; ITS1-5'-CCTCTTCGGGGACGCGCGCTGGCC-CCGA-3'; and ITS2-5'-GCGCGACGGCGGACGACACCGCG-GCGTC-3'.

Next, 50 pmoles of each oligonucleotide probe was labeled in the presence of 50 pmoles of [γ -³²P] ATP (Perkin-Elmer) and T4 polynucleotide kinase (New England Biolabs) for 30 min at 37 °C.

Three micrograms of nuclear RNAs were resolved on a 1% denaturing agarose gel and blotted onto a Hybond-N+ membrane (GE Healthcare). Signal detection was performed using a PhosphorImager (FLA 9500; GE Healthcare). Total 28S and 18S rRNA were visualized by fluorescence imaging following ethidium bromide staining, and were used as loading controls.

Analysis of rRNA methylation by isotope labeling. Cells were incubated for 1 h in DMEM supplemented with 10% heat-inactivated dialyzed FCS and [³H]-methyl-methionine at a final concentration of 15 μ Ci/mL (GE Healthcare). After 1-h labeling, cells were washed three times with ice-cold PBS, and scrapped. Nuclei were isolated by mechanical fractionation and nuclear RNAs were extracted and separated on formaldehyde agarose gels. RNAs were transferred onto nitrocellulose membranes. Radioactivity of the 45S/47S pre-rRNA was measured after exposure on a Phosphor-Imager screen, using a Typhoon scanner (GE Healthcare).

Analysis of rRNA methylation by RiboMethSeq. RiboMethSeq is based on protection of phosphodiester bonds against alkaline hydrolysis conferred by replacement of the 2'OH by a methyl group. The 3'-downstream nucleotide is thus absent at the 5'-end in the collection of RNA fragments. Partial alkaline hydrolysis of total RNA samples is followed by deep-sequencing and allows 5'-end counting for every fragment. The calculated MethScore repre-

sents the level of methylation of a given nucleotide in the ribosomal population.

RiboMethSeq was performed essentially as described previously (9), and is presented in detail below.

RNA fragmentation. RNA (1–250 ng) was subjected to alkaline hydrolysis in 50 mM bicarbonate buffer pH 9.2 for 4–10 min at 95 °C. The reaction was stopped by ethanol precipitation using 0.3 M Na-OAc pH 5.2 and glycoblue as a carrier in liquid nitrogen. After centrifugation, the pellet was washed with 80% ethanol and resuspended in nuclease-free water. The size of RNA fragments (30–200 nt) generated was assessed by capillary electrophoresis using a PicoRNA chip on Bioanalyzer 2100 (Agilent).

End repair of RNA fragments. Fragmented RNA without an additional fractionation step was 3'-end dephosphorylated using 5 U of Antarctic Phosphatase (New England Biolabs) for 30 min at 37 °C and after inactivation of the phosphatase for 5 min at 70 °C, RNA was phosphorylated at the 5'-end using T4 PNK and 1 mM ATP for 1 h at 37 °C. End-repaired RNA was then purified using RNeasy MinElute Clean-up kit (Qiagen) according to the manufacturer's recommendations, except that 675 μ L of 96% ethanol were used for RNA binding. Elution was performed in 10 μ L of nuclease-free H₂O.

Library preparation. RNA was converted to a library using NEBNext Small RNA Library kit (New England Biolabs) following the manufacturer's instructions. Library quality was assessed using a High Sensitivity DNA chip on a Bioanalyzer 2100. Library quantification was done using a fluorometer (Qubit 2.0 fluorometer; Invitrogen).

Sequencing. Libraries were multiplexed and subjected to high-throughput sequencing using an Illumina HiSeq. 1000 instrument with 50-bp single-end read runs. Since clustering of the short fragments was very efficient, libraries were loaded at a concentration of 6–8 pM per lane.

Bioinformatics pipeline. Adapter sequence trimming was done using Trimmomatic-0.32 with the following parameters: LEADING:30 TRAILING:30 SLIDINGWINDOW:4:15 MINLEN:17 AVGQUAL:30. Alignment to the reference rRNA sequence was done by Bowtie2 (v2.2.4) in End-to-End mode. The 5'-end counting was done directly on *.sam file using dedicated Unix script. Final analysis was performed by calculation of MethScore for quantification of 2'-O-Me residues.

Ribosome Structure Analysis. Methylated nucleotides were mapped on the cryo-EM structure of the human ribosome (PDB ID code 4UG0) (30). Reference structure of prokaryotic ribosome containing A-, P-, and E-site tRNAs plus mRNA was from *Thermus thermophilus* (PDB ID code 4V5C) (39). Observations of methylation site positions with regards to active sites of the ribosome were made after sequence alignment and structural superposition of the 23S rRNA of the *T. thermophilus* ribosome onto the 28S rRNA of human ribosome for the PTC and the CCA-end binding pocket of the E-site or sequence alignment and structural superposition of the 16S rRNA of the *T. thermophilus* ribosome onto the 18S rRNA of human ribosome for the decoding center and mRNA path. Figures and sequence alignment followed by structure superposition were performed using PyMOL 1.4 (Schrödinger; <https://pymol.org/2/>).

shRNA-Expressing Cells Lines. HeLa cell lines expressing an inducible shRNA were generated by lentiviral infection. Lentiviral particles were produced using the pTRIPZ-shRNA-NS and pTRIPZ-shRNA-FBL-351067 vectors, which were acquired from Open Biosystems (Dharmacon). Upon lentiviral infection, cell populations were selected using puromycin at 2 μ g/mL for 14 d. Cells were grown under selection at 1 μ g/mL and cultured without puromycin 24 h before any experiment. shRNA expression was induced by doxycycline treatment for 5 d at 1 μ g/mL. Induction efficiency was monitored by evaluating the tRFP expression

marker, by observation under fluorescent microscope or by flow cytometry. Change in FBL levels was verified by Western blotting.

Ribosome profiling

Ribosome protected fragments preparation and sequencing. HeLa ± FBL-shRNA cells were cultivated during 5 d with 1 μ g/mL doxycycline. Next, 150 million cells were treated with 100 μ g/mL cycloheximide for 10 min at 37 °C. Then, cells were suspended in lysis buffer (10 mM Tris-PO₄ pH7.5, 100 mM KCl, 10 mM Mg²⁺ acetate, 1% Triton X-100, 2 mM DTT, 100 μ g/mL cycloheximide). Glass beads were added and cells were lysed by vortexing 5 min at 4 °C. The supernatant was collected and digested with 15 U RNase I AMBION/OD₂₆₀ for 1 h at 25 °C. To purify monosomes, digested extracts were loaded on a 24% sucrose cushion (50 mM Tris-acetate pH 7.6, 50 mM NH₄Cl, 12 mM MgCl₂, 1 mM DTT) and centrifuged at 413,000 \times g for 2 h 15 min. Pellets were washed with lysis buffer and suspended in the same buffer. Ribosome protected fragments (RPF) were purified by phenol-chloroform extraction, and suspended in Superase-IN AMBION solution (1 U/ μ L).

RPFs were purified on 16% (vol/vol) acrylamide-bisacrylamide (19:1) gel made with 7 M urea and 1× Tris-acetate-EDTA (TAE) buffer. After staining for 30 min in 1× SYBR Gold (Life Technologies) diluted in 1× TAE, RNA contained in the 28- and 34-nt regions were separately excised from the gel. The gels slices were stored at least 2 h at -20 °C and then physically disrupted. RNAs were eluted overnight from gel fragments by passive diffusion in 500 μ L of RNA extraction buffer II (300 mM NaOAc pH 5.5, 1 mM EDTA) at 4 °C on a rotating wheel. After filtration, RPFs were precipitated in ethanol with 20 mg glycogen, suspended in 20 μ L Superase-IN AMBION solution (1 U/ μ L), and stored at -80 °C until use.

Ribosome footprints were depleted from ribosomal RNA with the Ribo-Zero Human kit (Illumina) following the manufacturer's recommendations. Sequencing libraries were prepared from an equal number of RPF with the TruSeq Small RNA library preparation kit (Illumina). Next generation sequencing was performed using a HisEq. 2500 single read 75.

Bioinformatic analysis: From raw data to alignment files. Raw data were first trimmed to remove the 3' adapter sequence with CutAdapt 1.9.1 configured with -e 0.12, -m 24, -M 35 (-M 51 for RNA-seq) and -a options to select a read size in a range from 24 to 35 nt allowing 12% of error. The trimmed reads were then mapped against the rRNA sequences of hg38 human genome with Bowtie 1.1.2 set up with -all and -un options. The -un option was used to select all unmapped reads. These filtered reads were finally mapped against the complete genome (hg38 human genome) and against the coding sequences only with Bowtie 1.1.2. The latter alignments were configured with two mismatches allowed (-n 2) and only uniquely mapped reads were selected (-m 1). The sam formatted files generated by the aligner were converted to sorted and indexed bam-formatted files using the Samtools program.

Differential expression. The number of reads for each gene was calculated with the featureCounts 1.5.0-p2 program and normalized with DESeq2 method through SARTools R Package. GO terms were identified for genes showing a significant expression variation using Panther (43). To represent RNA-seq and RIBO-seq profiles, we used the cross graphic representation where we display log₂ fold-changes in both data. The data were directly taken from the tables generated by SARTools. The colors were then selected based on the Panther groups and manually curated in case of no panther group was identified. The graphic was done using Bokeh v0.12.9 (<https://bokeh.pydata.org/en/0.12.9/>).

Ribosome footprints repartition. To identify reads as ribosome footprints, we did a metagene over the transcripts of hg38 human

assembly to study the periodicity and the proportion of reads mapped on the CDS, UTR5, and UTR3. This analysis was done using only the 28 Mers, which are the most abundant kmers. Only transcripts coding for proteins with "appris_principal" tag were kept which represent 26,176 transcripts. For the metagene, we selected the 100 last nucleotides of the UTR5, the 100 first nucleotides of the UTR3 and the 100 first and 100 last nucleotides of the CDS. All transcripts that didn't have one of these three requirements was discarded for the metagene analysis. With these filters, we analyzed 25,062 CDS and 14,859 UTR5/3. From this metagene, we counted the proportion of mapped reads (28Mers) and realized a Fourier transform using the scipy package with the scipy.fft.fft function to study the periodicity. Python Bokeh library v0.12.9 was used to plot the percentages of mapped 28Mers on each feature and their corresponding period detected from Fourier transform results.

In vitro hybrid translation. Hybrid in vitro translation assay was performed as described previously (35) and is summarized hereafter. After centrifugation of 1 mL of RRL for 2 h 15 min at 240,000 \times g, 900 μ L of ribosome-free RRL (named S100) was collected, frozen, and stored at -80 °C. The extent of ribosome depletion from reticulocyte lysate was checked by translating 27 nM of in vitro transcribed capped and polyadenylated globin-*Renilla* mRNA in the S100 RRL and validated when no luciferase activity could be detected. In parallel, transfected cells were lysed in hypotonic buffer R [Hepes 10 mM pH 7.5, CH₃CO₂K 10 mM, (CH₃CO₂)₂Mg 1 mM, DTT 1 mM] and potter homogenized (around 100 strokes). Cytoplasmic fraction was obtained by 13,000 \times g centrifugation for 10 min at 4 °C. The ribosomal pellet was then obtained by ultracentrifugation for 2 h 15 min at 240,000 \times g in a 1 M sucrose cushion and was rinsed three times in buffer R2 containing Hepes 20 mM, NaCl 10 mM, KCl 25 mM, MgCl₂ 1.1 mM, β -mercaptoethanol 7 mM and suspended in 30 μ L of buffer R2 to reach more than 10 μ g/ μ L ribosome concentration for optimal and long storage at -80 °C. The reconstituted lysate was then assembled by mixing 5 μ L of S100 RRL with a scale from 0.25 to 4 μ g of ribosomal pellet. Typically, the standard reaction contained 5 μ L of ribosome-free RRL with 1 μ g ribosomal pellet in a final volume of 10 μ L. Upon reconstitution, the translation mixture was supplemented with 75 mM KCl, 0.75 mM MgCl₂, and 20 μ M amino acid mix.

For in vitro translation assays, p0-*Renilla* vectors containing the β -globin, GAPDH 5'UTR, CrPV, DCV, or EMCV IRESs were described previously (38). mRNAs were obtained by in vitro transcription, using 1 μ g of DNA templates linearized at the AflII sites, 20 U of T7 RNA polymerase (Promega), 40 U of RNAsin (Promega), 1.6 mM of each ribonucleotide triphosphate, 3 mM DTT in transcription buffer containing 40 mM Tris-HCl (pH 7.9), 6 mM MgCl₂, 2 mM spermidine, and 10 mM NaCl. For capped mRNAs, the GTP concentration was reduced to 0.32 mM and 1.28 mM of m7GpppG cap analog (for β -globin mRNA) or m7GpppA (for CrPV mRNA) (New England Biolabs) was added. The transcription reaction was carried out at 37 °C for 2 h, the mixture was treated with DNase and the mRNAs were precipitated with ammonium acetate at a final concentration of 2.5 M. The mRNA pellet was then suspended in 30 μ L of RNase-free water and mRNA concentration was determined by absorbance using the Nanodrop technology. mRNA integrity was checked by electrophoresis on nondenaturing agarose gel.

Statistical Analysis. Statistical analysis was performed using the Prism software (v7.0. GraphPad). A two-tailed Student *t* test was used for evaluating significance, except for RiboMethSeq data for which a one-tailed *t* test was used.

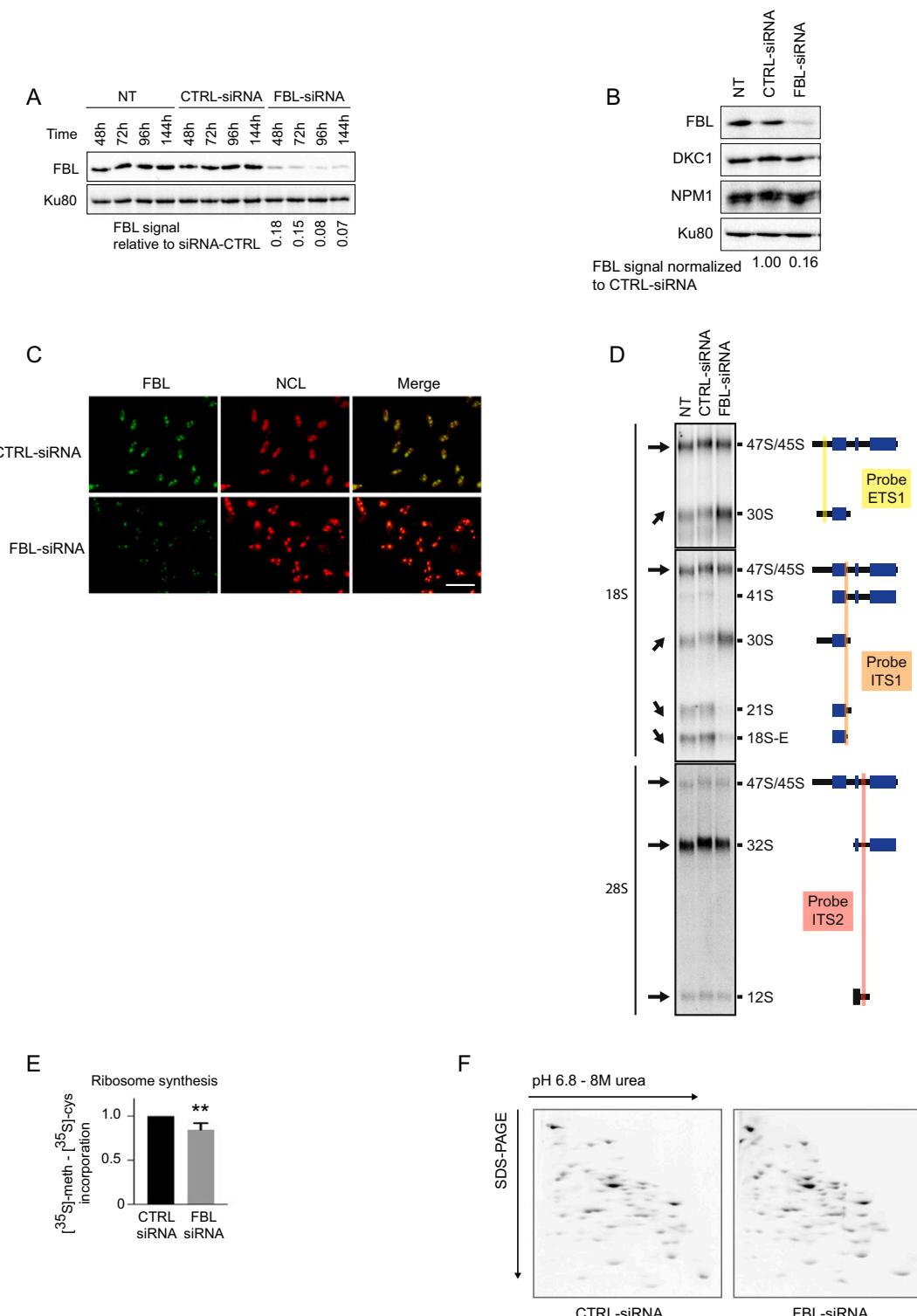


Fig. S1. Related to Fig. 1. Impact of FBL knockdown on major nucleolar markers. (A) Western blot analysis of *FBL* in HeLa cells transfected with three siRNAs targeting FBL for the periods of time indicated above each lane. Ku80 was used as a loading control. FBL signal was quantified and normalized against CTRL-siRNA (values are indicated below each condition). (B) Western blot analysis of dyskerin (DKC1) and nucleophosmin (NPM1) levels in FBL knockdown HeLa cells compared with nontransfected cells (NT) and cells transfected with a CTRL-siRNA. Analysis performed 72 h posttransfection. Ku80 was used as a loading control. FBL signal was quantified and normalized against the CTRL-siRNA condition. (C) Immunofluorescence detection of FBL (green) and nucleolin (NCL) (red) 72 h after transfection with CTRL or FBL-siRNA. (Scale bar, 10 μm .) (D) Northern blot analysis of pre-rRNA processing in nontransfected (NT) or siRNA transfected (CTRL-siRNA and FBL-siRNA) HeLa cells. The position of the probes used and the detected pre-rRNA species are indicated on the right. Arrows on the left indicate the trend of each species in FBL-siRNA cells to increase or decrease compared with CTRL-siRNA cells. (E) Ribosome synthesis rate. ${}^{35}\text{S}\text{-Meth-}{}^{35}\text{S}\text{-Cys}$ incorporation into purified cytoplasmic ribosomes 72 h after siRNA transfection in HeLa cells. Radioactivity was measured by liquid scintillation. Values are presented as mean \pm SD ($n = 2$). ** $P \leq 0.01$. (F) The 2D-PAGE analysis of 0.5 M KCl-purified ribosomes extracted from ${}^{35}\text{S}\text{-Meth-}{}^{35}\text{S}\text{-Cys}$ pulse-labeled cells. Proteins were separated according to their charge in the first dimension and according to their molecular weight in the second dimension. Images show the radioactive signal obtained by phosphor-imaging.

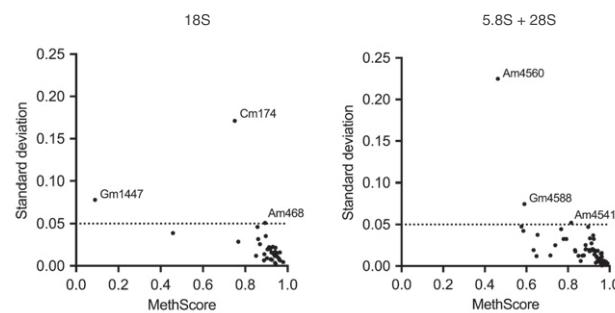


Fig. S2. Related to Fig. 2. Quantitative mapping of rRNA 2'-O-Me in human cells. Variability (SD) of RiboMethSeq data for each 2'-O-Me site, according to its level of methylation. Sites with variability greater than 5% are named on the graph. Most sites show a variability below 5% (dotted line). Values are represented as means of three independent biological replicates.

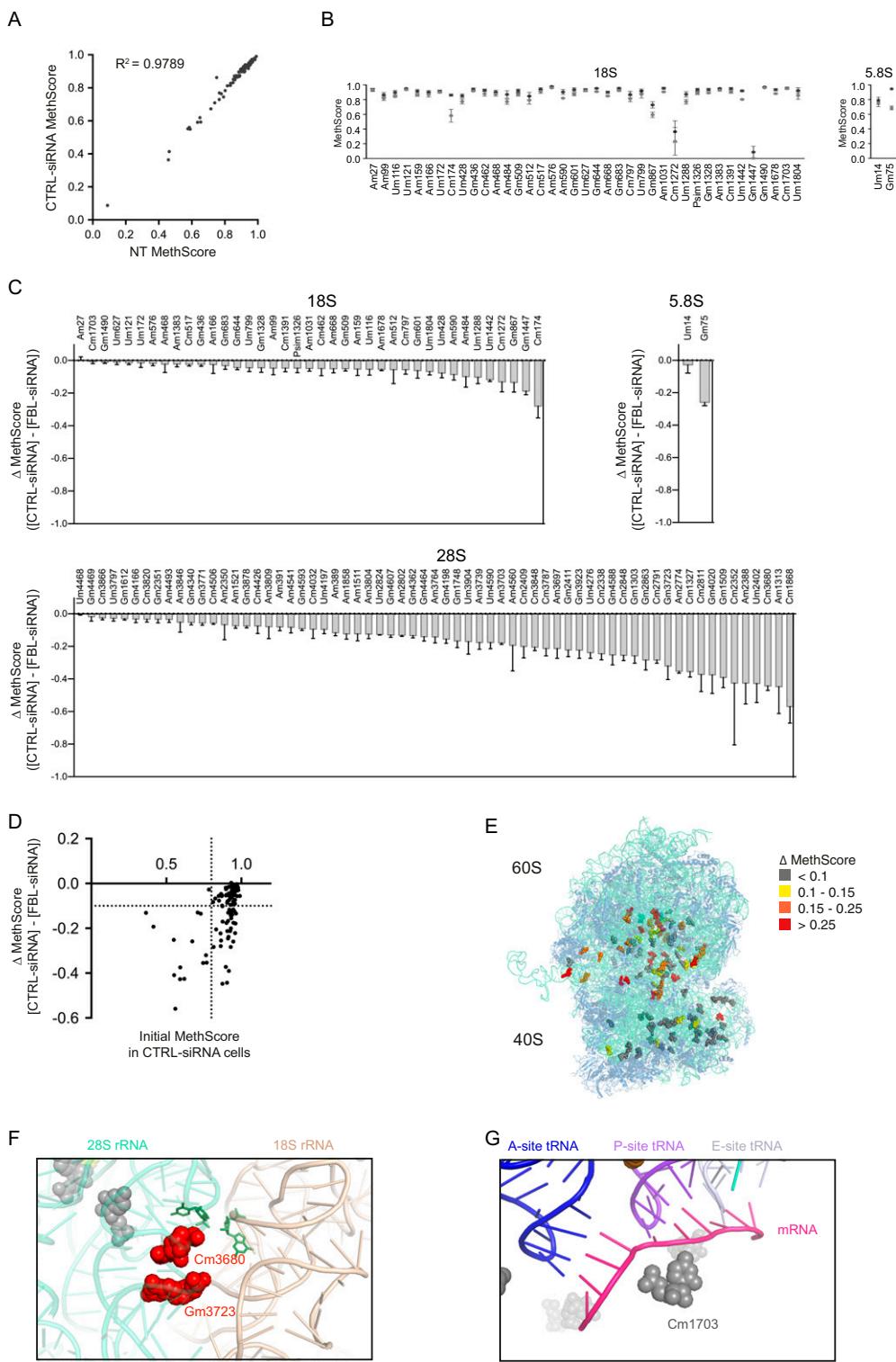


Fig. S3. Related to Fig. 3. Identification and localization of altered 2'-O-Me sites. (A) siRNA transfection has no effect on 2'-O-Me. Plot of the MethScore of each site in nontreated HeLa cells (NT) vs. HeLa cells transfected with control siRNA (CTRL-siRNA). Correlation coefficient (R^2) was calculated from the linear regression curve. (B) Alteration of 2'-O-Me upon *FBL* knockdown. Mean MethScore values \pm SD ($n = 3$ independent biological replicates) for each methylated nucleotide in 18S and 5.8S rRNA from HeLa cells transfected with CTRL-siRNA (black circle) or FBL-siRNA (gray circle). (C) FBL knockdown induces a decrease of MethScore at all but one site. For each site, the MethScore obtained in FBL-siRNA transfected HeLa cells was subtracted from the one from CTRL-siRNA transfected cells (same dataset as in Fig. 2B). Sites are shown in order of increasing difference in MethScore for the 18S, 5.8S, and 28S rRNAs. Error bars are SD from three independent biological replicates. (D) Distribution of 2'-O-Me sites according to their initial methylation frequency (x axis) and to their methylation variation upon FBL knockdown (y axis). Partially methylated sites in siRNA-CTRL cells (left of the vertical dotted line) were all down-methylated in siFBL-cells (Δ MethScore > 0.1). Note that the site 18S-Cm1447 was excluded, since it is not methylated in our HeLa cells. (E) Position of methylated nucleotides on the 3D structure of the human ribosome, shown as assembled 80S ribosome. Nucleotides are color-coded according to the variation in MethScore comparing FBL-siRNA cells with CTRL-siRNA cells. See also Datasets S2 and S3. (F) View of intersubunit bridge B2b [18S-1051 and 28S-3699 (shown as stick in green)], showing close proximity of Cm3680 and Gm3723 methylation sites (red), which methylation was decreased by 44.3% and 32.0%, respectively. (G) View of the decoding center, showing the anticodon loop of tRNAs interacting with the mRNA at the A, P, and E-sites. The Cm1703 nucleotide (gray, not affected by FBL knockdown) is close to the mRNA.

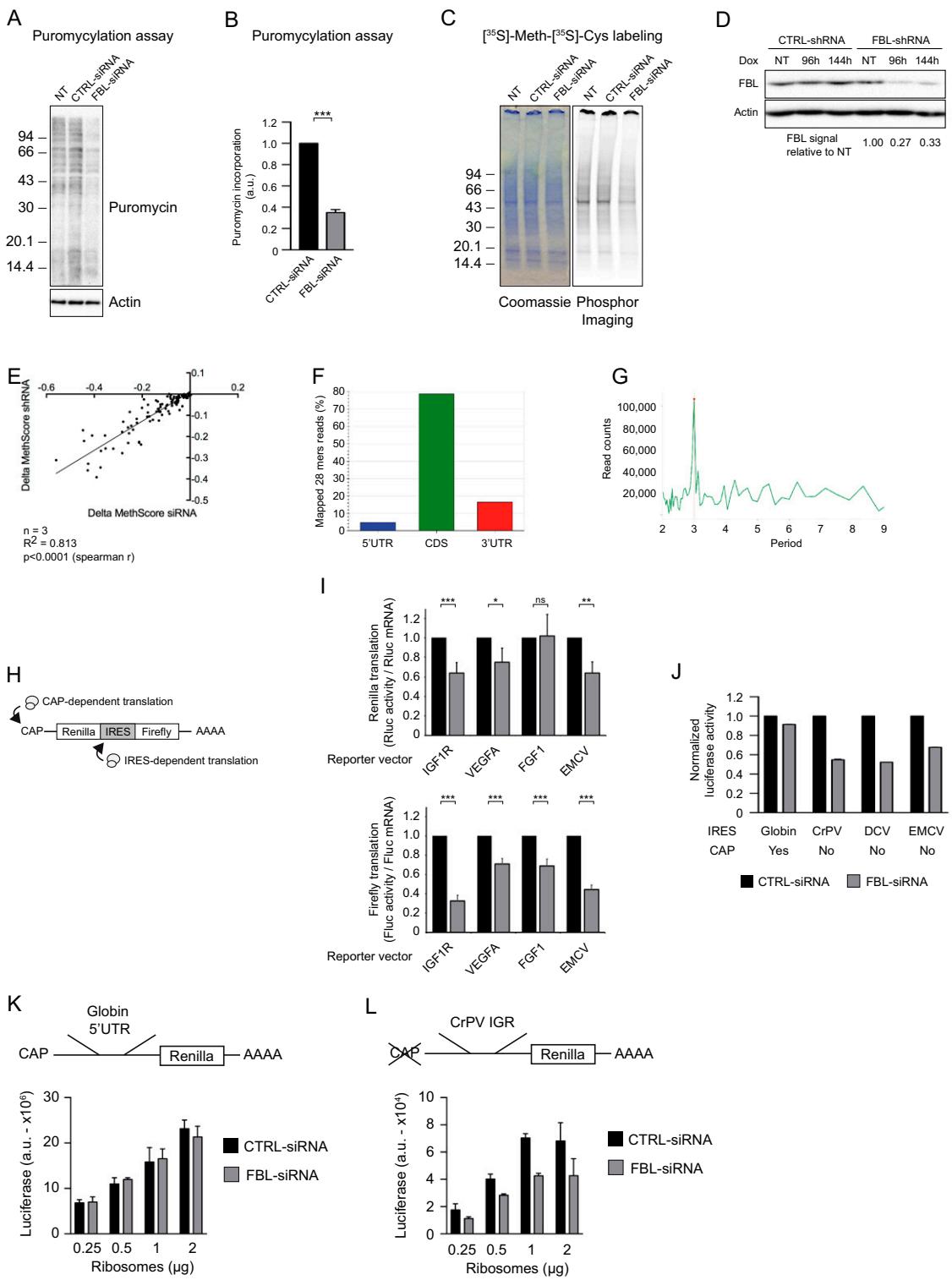


Fig. S4. Related to Fig. 4. rRNA methylation defect impairs IRES dependent translation from various viral IRES elements. (A) Global protein synthesis detection by incorporation of puromycin in nascent peptides. A representative Western blot with an anti-puromycin antibody is shown. Actin detection was used as loading reference. (B) Global protein synthesis was measured by incorporation of puromycin in nascent peptides using Western blot against puromycin. Puromycin signal was normalized against actin. Values are presented as mean \pm SD ($n = 3$). (C) Global protein synthesis detection by pulse labeling with $^{[35]}\text{S}$ -methionine and $^{[35]\text{S}}$ cysteine. A representative SDS/PAGE stained with Coomassie is shown (Left). The corresponding radioactive signal is shown upon phosphor-Imaging detection (Right). (D) Western blot analysis of FBL level in HeLa cells in which the expression of a CTRL shRNA or a FBL-shRNA was induced for 120 h by doxycycline (Dox) treatment. Actin was used as a loading control. The level of FBL upon shRNA expression is indicated below and normalized to CTRL-shRNA. (E) Plotting of the decrease of 2'-O-Me ($\Delta\text{MethScore}$) in FBL-shRNA expressing cells (y axis) and of the $\Delta\text{MethScore}$ in shRNA expressing cells compared with siRNA expression cells (Spearman r test, $P < 0.0001$). (F) Fraction of the mapped 28Mers on each feature of transcripts. In this metagene, transcripts with a CDS length lower than 200 nucleotides and a 5'

Legend continued on following page

UTR and 3'UTR lengths lower than 100 nt were discarded. (G) Period detected from the frequencies of RPF obtained by fast Fourier transform. The red dot shows the most frequent period found in the data. The vertical red line shows the expected period of 3 nt for ribosome footprints. This periodicity is not observed in UTRs regions indicating that the RPF correspond to active ribosome footprints. (H) Representation of in cellulo IRES-dependent translation assay using a bicistronic luciferase reporter construct. *Renilla* luciferase is translated in a Cap-dependent manner and Firefly luciferase is translated from the IRES element. (I) Translation efficiency of individual luciferase reporters was evaluated as the ratio of luciferase activity over mRNA levels. Luciferase activities were measured as in Fig. 4 and compared with luciferase mRNA levels measured by RT-qPCR. Data represent activity/mRNA ratios for Rluc (Upper, gray bars) and Fluc (Lower, gray bars), normalized against CTRL-siRNA (black bars). Values are presented as mean \pm SD ($n = 3$). (J) Ribosomes purified from HeLa cells transfected with either CTRL-siRNA or FBL-siRNA were used to translate mRNAs in the hybrid in vitro translation assay. Translation was evaluated on luciferase reporter mRNAs containing a 5'UTR originating from human globin mRNA, or containing IRES elements from the CrPV, the DCV, or EMCV. The presence of Cap on the mRNA is indicated below the graph. Data are presented as mean \pm SD ($n = 2$). (K and L) Identical experimental set-up as in Fig. 4F using a range of ribosome quantities. Data represent mean values (\pm SD, $n = 3$). Statistical significance was verified by conducting a Student *t* test. ns, not significant; * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$.

Other Supporting Information Files

[Dataset S1 \(XLSX\)](#)

[Dataset S2 \(XLSX\)](#)

[Dataset S3 \(XLSX\)](#)

Résultats & Discussion

Analyse des effets de la baisse de la méthylation des ARNr sur l'initiation de la traduction

Analyse des effets de la baisse de la méthylation des ARNr sur l'initiation de la traduction

I – Résultats

1 – Analyse globale à l'échelle du génome de l'initiation de la traduction

Les tests sur l'initiation IRES-dépendante que j'ai déjà effectués ont montré que la diminution du taux de fibrillarine et de méthylation des ARNr entraîne une baisse de l'initiation au niveau de certains IRES. Nous avons donc voulu réaliser une étude globale, à l'échelle du génome, mais ciblée sur l'initiation de la traduction.

Pour cela, je me suis intéressée à deux techniques, le GTI-Seq (Global Translation Initiation Sequencing) et le QTI-Seq (Quantitative Translation Initiation Sequencing), qui sont des adaptations du ribosome profiling et qui nécessitent l'utilisation de molécules qui bloquent spécifiquement les ribosomes en initiation, respectivement l'harringtonine³¹⁴ et la lactimidomycine^{315,316}.

Le GTI-Seq permet d'identifier spécifiquement les ribosomes en cours d'initiation. L'harringtonine, molécule utilisée pour bloquer les ribosomes à l'initiation, se fixe à la sous-unité 60S et bloque le ribosome lors de l'association des deux sous-unités. Elle est ajoutée dans le milieu de culture des cellules et laissée à incuber le temps que les ribosomes en élongation quittent les ARNm. Au cours de cette période d'incubation il se produit des événements de réinitiation de la traduction. Ces événements de réinitiation, qui correspondent à l'association de sous-unités ribosomiques libres aux ARNm libres, entraînent un enrichissement non quantitatif des ribosomes bloqués aux sites d'initiation. Le GTI-Seq ne permet donc pas une mesure quantitative de l'initiation de la traduction.

Le QTI-Seq est une technique qui permet elle-aussi d'identifier spécifiquement les ribosomes en cours d'initiation mais également de mesurer quantitativement l'initiation de la traduction³¹⁷. La lactimidomycine, molécule inhibitrice de l'initiation de la traduction utilisée en QTI-Seq, a un mode d'action similaire à la cycloheximide, mais elle est plus volumineuse et se fixe au site E ainsi qu'au site P du ribosome. Ainsi, elle ne peut se fixer qu'aux ribosomes qui ne comportent encore aucun ARNt au site E, c'est-à-dire aux ribosomes en initiation. Le protocole d'extraction des polysomes est ensuite adapté afin de s'affranchir des événements de réinitiation de la traduction qui ont lieu pendant la période d'incubation de la molécule. Pour cela, la traduction est tout d'abord bloquée en plaçant les cellules dans la glace, ce qui fige les ribosomes. Puis de la lactimidomycine est ajoutée aux cellules et va se fixer au niveau des ribosomes à l'initiation. Tous les ribosomes sont donc bloqués sur les ARNm, qu'ils soient en initiation ou en élongation. Pour

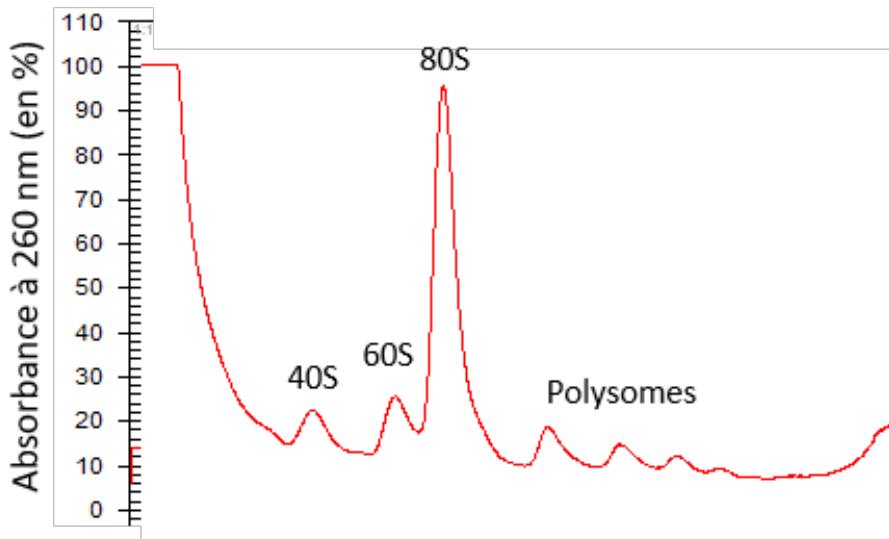


Figure 25 : profil de polysomes de cellules HeLa dont la traduction a été bloquée à la cycloheximide
 Les extraits de polysomes récoltés à partir de cellules HeLa traitées à la cycloheximide sont séparés sur gradients de sucre 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm.

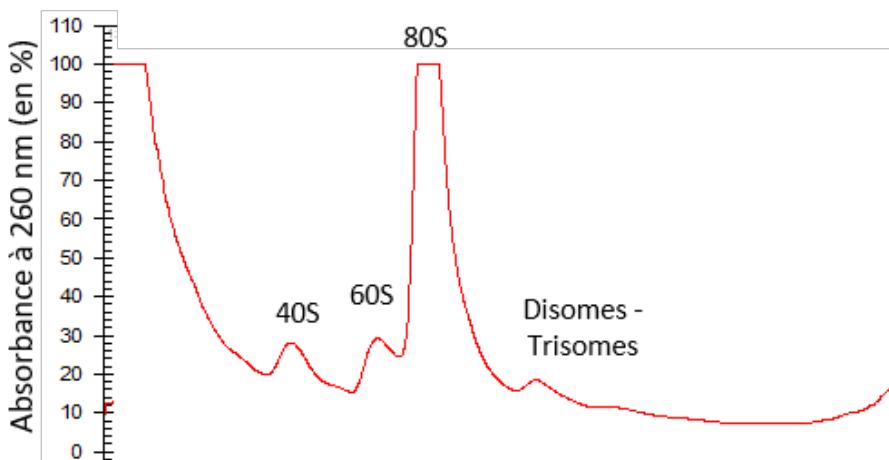


Figure 26 : profil de polysomes de cellules HeLa dont la traduction a été bloquée à l'harringtonine
 Les extraits de polysomes récoltés à partir de cellules HeLa traitées à l'harringtonine sont séparés sur gradients de sucre 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm.

décrocher les ribosomes en élongation et ne conserver que les ribosomes en initiation, il faut ajouter de la puromycine. Cependant, la puromycine n'agit que lorsque les ribosomes sont en cours de traduction. Il faut donc faire redémarrer la traduction *in vitro*, en ajoutant de la créatine phosphate, de la créatine kinase, de la spermidine et de l'ATP. La suite du protocole est identique à un ribosome profiling classique. Avec l'harringtonine, il est impossible d'adapter le protocole ainsi car elle se fixe initialement sur la sous-unité 60S et nécessite une traduction active pour bloquer le ribosome 80S sur le site d'initiation. Il n'est donc pas possible d'ajouter l'harringtonine après blocage des ribosomes et de la traduction par le froid.

Dans un premier temps j'ai tenté de réaliser la technique de QTI-Seq sur un échantillon préparé à partir de cellules HeLa. Alors que tous les contrôles préliminaires semblaient indiquer que le protocole utilisé était efficace, les reads obtenus après séquençage s'alignent tout le long des CDS sans enrichissement aux sites d'initiation de la traduction. La technique n'a donc pas fonctionné. J'ai par la suite testé l'action de la lactimidomycine par ajout dans le milieu de culture des cellules et je n'ai pas pu montrer un effet d'inhibition de la traduction. Il semblerait donc que la lactimidomycine, dans les conditions dans lesquelles je l'ai utilisée, ne bloque pas les ribosomes en initiation. Cette molécule étant très coûteuse, je n'ai pas pu essayer d'améliorer le protocole pour obtenir un résultat concluant.

J'ai alors réalisé la technique de GTI-Seq. J'ai tout d'abord vérifié son efficacité en réalisant des profils de polysomes à partir de cellules HeLa traitées avec de la cycloheximide ou de l'harringtonine pendant 10 min, durée durant laquelle les ribosomes en élongation, non bloquée, poursuivent la traduction jusqu'à quitter les ARNm. Sur ces profils, les pics de 40S, 60S et 80S sont corrects. En ce qui concerne les polysomes, on peut observer jusqu'à 5 ribosomes par ARNm avec la cycloheximide (Figure 25) et une importante perte des polysomes avec l'harringtonine, avec toutefois des disomes et quelques trisomes qui subsistent (Figure 26). Cela est dû au fait que l'harringtonine peut permettre un ou deux cycles d'élongation de la traduction. L'utilisation de l'harringtonine pour bloquer l'initiation de la traduction est donc validée.

2 – L'initiation de la traduction est affectée par la baisse de la méthylation des ARNr

J'ai réalisé la préparation de fragments d'ARNm protégés par les ribosomes à partir de cellules HeLa shFBL traitées ou non avec de la doxycycline pendant 5 jours. J'ai suivi le même protocole que pour le RiboSeq, en remplaçant la cycloheximide par l'harringtonine. Les RPFs sont

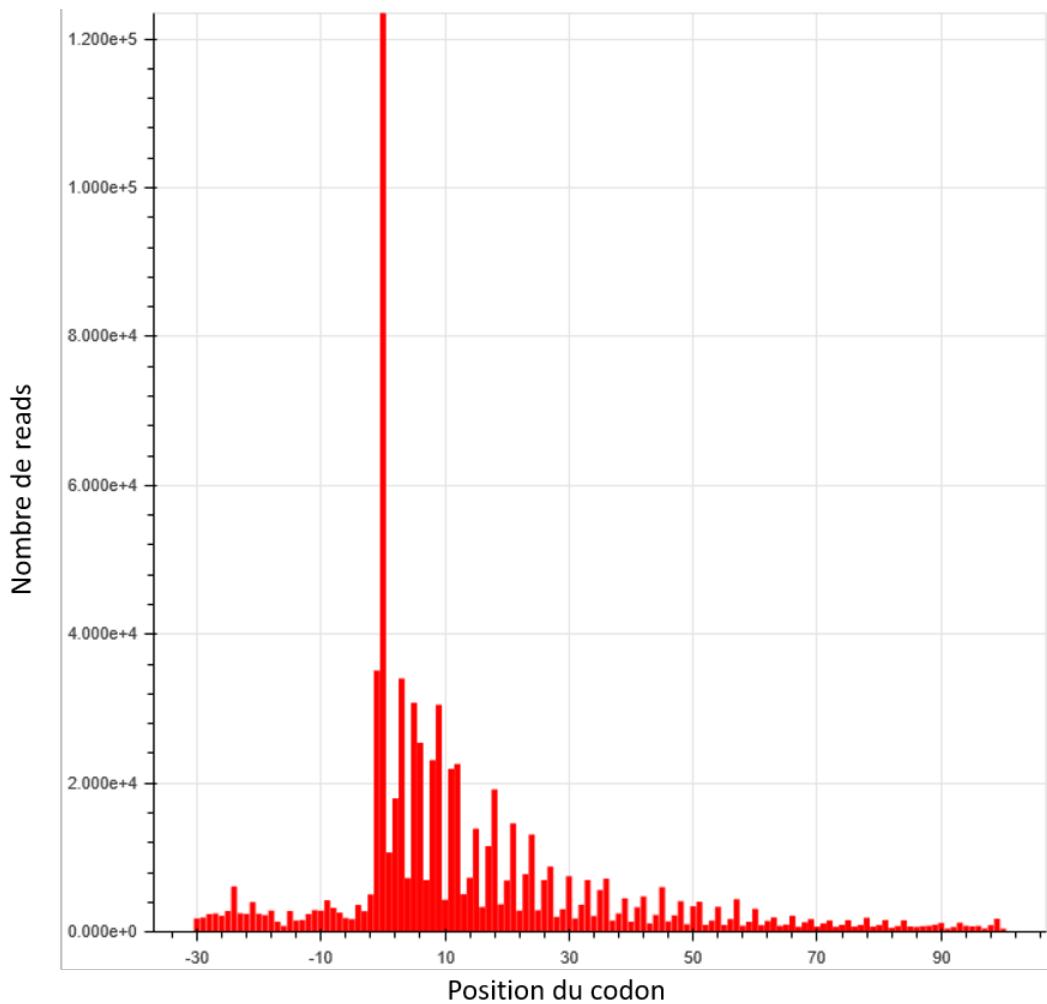


Figure 27 : Métagène des reads du GTI-Seq dans la région du codon d'initiation

Les reads de tous les gènes sont poolés, comptés et représentés codon par codon autour du codon d'initiation de la traduction.

ensuite utilisés pour la création de la banque, puis séquencés. Trois réplicats biologiques ont été réalisés.

Les données de séquençage sont traitées dans un premier temps comme pour le RiboSeq : élimination des adaptateurs, des reads trop courts et trop longs ainsi que des séquences ribosomiques. L'ensemble des reads est poolé et représenté sous la forme de métagènes, dans les régions autour du codon d'initiation annotée (Figure 27) et autour du codon de terminaison annotée (Figure 28). Ces graphes montrent un fort enrichissement des reads au codon d'initiation, bien qu'il y ait également un nombre important de reads sur les 30 premiers nucléotides. Puis le bruit de fond est très faible jusqu'au stop. La technique a donc bien fonctionné. De plus, la périodicité est bien visible (Figure 27), ce qui indique que les reads correspondent à des fragments d'ARNm protégés par des ribosomes en traduction.

Il a fallu ensuite distinguer les pics d'initiation du bruit de fond. Pour cela, nous nous déplaçons avec une fenêtre de 7 nucléotides et nous regardons si le nombre de reads à ce nucléotide central est un maximum local au sein de la fenêtre. Si ce n'est pas le cas, la fenêtre se déplace d'un nucléotide en 3'. Si c'est le cas, comme précédemment publié³¹⁶, nous calculons le R-Score défini comme tel :

$$R\text{-Score} = \frac{\frac{\text{Nombre de reads à cette position en GTI-Seq}}{\text{Nombre total de reads sur tout le gène en GTI-Seq}}}{\frac{\text{Nombre de reads à cette position en Ribo-Seq}}{\text{Nombre total de reads sur tout le gène en Ribo-Seq}}}$$

Si $R\text{-Score} \geq 0,05$, alors le maximum local est défini comme un pic. La caractérisation d'un pic nécessite d'avoir une bonne couverture du gène en RiboSeq.

Afin de déterminer des variations de l'initiation de la traduction entre la condition témoin et la condition d'hypométhylation des ARNr, nous avons appliqué une méthode du « Tout ou rien ». En effet, le GTI-Seq n'étant pas une technique quantitative, nous avons listé les sites d'initiation qui présentent un pic dans une condition et pas dans l'autre condition. Nous avons obtenu une liste de 66 sites (Figure 29) : 57 qui présentent un pic en condition d'hypométhylation des ARNr, et 9 qui présentent un pic en condition témoin. Cette liste ne comporte aucun gène en commun avec la liste de gènes candidats du RiboSeq. Ces deux techniques nous permettent donc de regarder deux niveaux de régulation différents : avec le RiboSeq on regarde uniquement l'elongation de la traduction tandis qu'avec le GTI-Seq on regarde uniquement l'initiation. Les sites ont été caractérisés suivant leur localisation sur le transcript. Un site d'initiation déjà annoté comme tel est un aTis

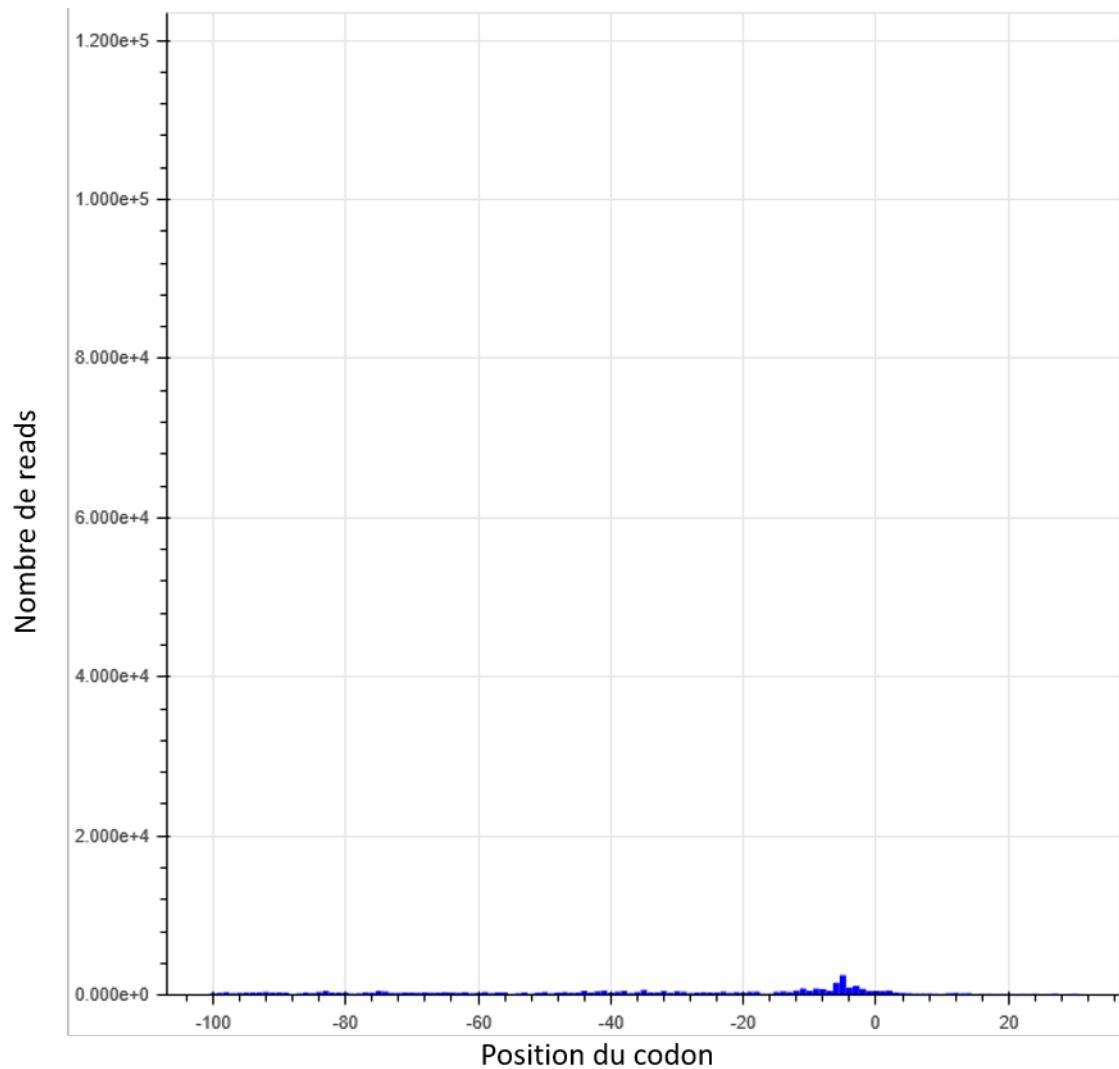


Figure 28 : Métagène des reads du GTI-Seq dans la région du codon de terminaison

Les reads de tous les gènes sont poolés, comptés et représentés codon par codon autour du codon de terminaison de la traduction.

(annotated Translation initiation site), un site d'initiation non annoté et situé en amont du site annoté est un uTis (upstream Tis) et un site non annoté et situé en aval du site annoté est un dTis (downstream Tis). Parmi les 66 sites, un seul est un aTis, 35 sont des uTis et 30 sont des dTis.

J'ai vérifié si les sites d'initiation correspondent à des IRES connus, mais sans résultat positif. Puis j'ai effectué une analyse « Gene Ontology » sur les gènes correspondants aux sites d'initiation, sans identifier d'enrichissement fonctionnel.

II – Discussion

Suite aux résultats obtenus avec les mesures de l'initiation IRES-dépendante, qui montraient un effet spécifique sur certains IRES de la baisse de la méthylation des ARNr, j'ai recherché les effets sur l'initiation de la traduction à l'échelle du génome. Je me suis intéressée à deux techniques de ribosome profiling qui ciblent l'initiation de la traduction par l'utilisation de drogues qui bloquent les ribosomes en initiation. De cette manière, nous faisions l'hypothèse de révéler des IRES et éventuellement des uORF qui présentent des variations de l'initiation de la traduction entre la condition normale et la condition d'hypométhylation des ARNr. J'ai tout d'abord tenté de mettre en application une technique quantitative, le QTI-Seq. Malheureusement, je n'ai pas réussi à obtenir les résultats attendus. J'ai donc ensuite réalisé un GTI-Seq, méthode non quantitative, sur des échantillons de RPFs préparés à partir de cellules HeLa shFBL traitées ou non à la doxycycline pendant 5 jours.

Le GTI-Seq nous a permis de localiser les sites d'initiation de la traduction dans les cellules HeLa shFBL en condition normale et en condition d'hypométhylation des ARNr. Cette méthode n'étant pas quantitative, j'ai choisi d'appliquer une méthode « tout ou rien » en listant les sites d'initiation qui présentent un pic dans une condition mais pas dans l'autre. J'ai ainsi identifié une liste de 66 sites.

Parmi ces sites d'initiation mis en évidence, 57 présentent un pic en condition d'hypométhylation des ARNr et 9 présentent un pic en condition témoin. De plus, parmi ces sites, un seul est un aTis, 35 sont des uTis et 30 sont des dTis. Ces informations révèlent que la baisse de la méthylation des ARNr est à l'origine de l'initiation de la traduction au niveau de sites où il n'y a pas de traduction en condition normale. Ces sites sont majoritairement des sites d'initiation de la traduction non annotés. La baisse de la méthylation des ARNr entraîne donc une modification des capacités d'initiation des ribosomes, plutôt au niveau d'éléments régulateurs autour de la CDS qu'au niveau de la CDS elle-même. Cependant, cette altération de l'initiation semble être compensée au niveau de l'elongation, ce qui suggère que les taux protéiques ne sont pas affectés.

Nom du gène	Position relative à l'ATG	Condition dans laquelle le pic est présent	Caractérisation
AGFG1	-56	FBL-	uTis
ALDH3A2	-27	FBL-	uTis
ALDH3A2	-118	FBL-	uTis
ANKS6	85	FBL-	dTis
ASCC3	-13280	FBL-	uTis
BUD13	33	FBL-	dTis
CAMKMT	15	FBL-	dTis
CBSL	-3588	FBL-	uTis
CBSL	-3588	FBL-	uTis
CDK8	-386	FBL-	uTis
CLCC1	-12939	FBL-	uTis
CLCN7	2485	FBL-	dTis
CTXN1	-391	FBL-	uTis
DNAJC16	-2124	FBL-	uTis
EPHX1	-18556	FBL-	uTis
EPHX1	-3181	FBL-	uTis
EPHX1	-18110	FBL-	uTis
FLAD1	-532	FBL-	uTis
GALT	4	FBL-	dTis
GPN2	-96	FBL-	uTis
GRB14	-92	FBL-	uTis
HAUS2	324	FBL-	dTis
HECTD1	1446	FBL-	dTis
HECTD1	1071	FBL-	dTis
IPPK	28	FBL-	dTis
IQCH	6	FBL-	dTis
MDP1	38	FBL-	dTis
MEN1	-329	FBL-	uTis
MRPL41	-42	FBL-	uTis
NAA25	34	FBL-	dTis
NHEJ1	-2351	FBL-	uTis
PAQR7	-6838	FBL-	uTis
PARP12	20	FBL-	dTis
PDP1	-4083	FBL-	uTis
POLG	-237	FBL-	uTis
POLG	-249	FBL-	uTis
PTTG1IP	-46	FBL-	uTis
RFX1	-12304	FBL-	uTis
RNF130	250	FBL-	dTis
RP11-195F19.29	4	FBL-	dTis
RP11-33O4.2	1440	FBL-	dTis
RP11-475E11.9	98	FBL-	dTis
SDCCAG8	10	FBL-	dTis
SH2B2	443	FBL-	dTis
SLC4A1AP	160	FBL-	dTis
SLC4A1AP	160	FBL-	dTis
SMIM19	-4716	FBL-	uTis
SMIM19	-4266	FBL-	uTis
SMIM19	-4048	FBL-	uTis
SMIM19	-5222	FBL-	uTis
SMIM19	47	FBL-	dTis
TECPR2	-13551	FBL-	uTis
TIGD5	2217	FBL-	dTis
TOR1A	84	FBL-	dTis
UPF2	163	FBL-	dTis
UPF2	-7309	FBL-	uTis
ZNF609	4402	FBL-	dTis
HSF1	-1	FBL+	aTis
NSA2	81	FBL+	dTis
PLD5	6930	FBL+	dTis
PLEC	614	FBL+	dTis
PPP1R14B	323	FBL+	dTis
RAB31	-38	FBL+	uTis
RPS27	-18	FBL+	uTis
SMARCD2	-84	FBL+	uTis
TUBA1B	-63	FBL+	uTis

Figure 29 : Liste des pics d'initiation présents dans une condition et absents dans l'autre

Le GTI-Seq effectué sur des cellules HeLa shFBL traitées ou non avec 1 µg/mL de doxycycline pendant 5 jours a révélé une liste de 66 pics d'initiation présents dans une condition et absents dans l'autre.

Par la suite, j'ai réalisé une analyse « Gene Ontology » sur les gènes correspondant aux sites d'initiation révélés par GTI-Seq, qui n'a révélé aucun enrichissement fonctionnel. Au cours de ma thèse, je n'ai pas pu approfondir plus la recherche sur ces sites d'initiation. Par la suite, la première vérification à effectuer sera de confirmer que la traduction s'initie au niveau des sites d'initiation détectés par GTI-Seq. Les régions autour des sites seront récupérées et insérées dans un système double rapporteur. Le travail sera poursuivi sur les sites pour lesquels une initiation de la traduction aura été confirmée, par exemple par la recherche de séquences nucléotidiques ou de structures secondaires qui seraient partagées par ces régions qui présentent un site d'initiation affecté par la baisse de la méthylation des ARNr.

Résultats & Discussion

Délétions ciblées de 2'-O-méthylations localisées dans le tunnel de sortie du peptide naissant chez *S. cerevisiae*

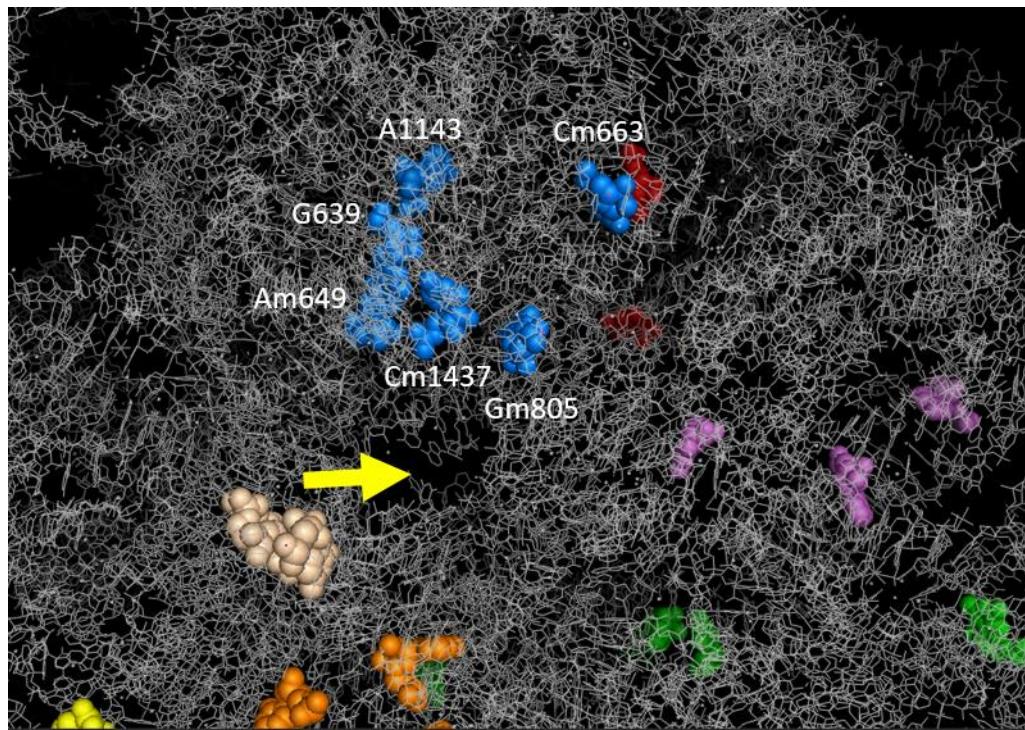


Figure 30 : Visualisation Pymol des positions méthylées autour du tunnel de sortie du peptide naissant sur le ribosome de levure *Saccharomyces cerevisiae*, vue du dessus

Les six positions méthylées chez l'homme, en bleu, sont localisées sur le ribosome de levure *S. cerevisiae*. La flèche jaune indique le tunnel de sortie du peptide naissant. Les positions bleues sont donc situées dans la région du tunnel de sortie.

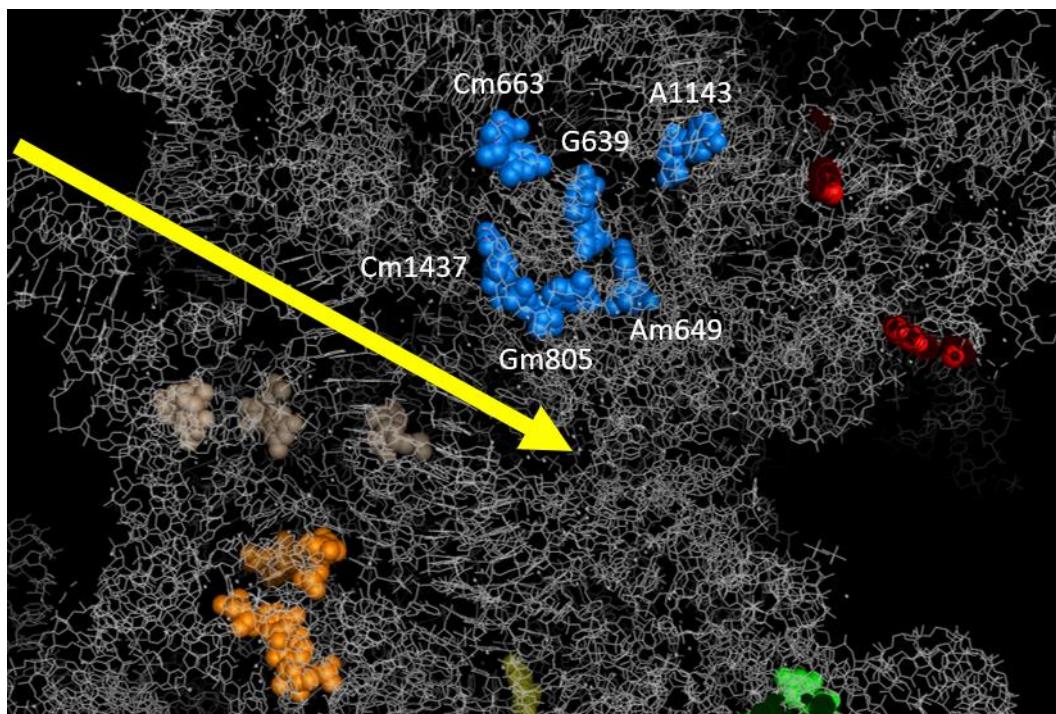


Figure 31 : Visualisation Pymol des positions méthylées autour du tunnel de sortie du peptide naissant sur le ribosome de levure *Saccharomyces cerevisiae*, vue latérale

Les six positions méthylées chez l'homme, en bleu, sont localisées sur le ribosome de levure *S. cerevisiae*. La flèche jaune indique le tunnel de sortie du peptide naissant. Les positions bleues sont donc situées dans la région du tunnel de sortie.

Délétions ciblées de 2'-O-méthylation localisées dans le tunnel de sortie du peptide naissant chez *S. cerevisiae*

I – Résultats

1 – Localisation des 2'-O-méthylation affectées par la baisse du taux de fibrillarine dans les cellules humaines

A la suite de l'étude de la baisse de la 2'-O-méthylation des ARNr causée par la baisse du taux de fibrillarine, nous avons localisé au sein du ribosome de la levure *Saccharomyces cerevisiae* les positions les plus touchées, c'est-à-dire celles pour lesquelles la différence de RiboMethScore entre la condition témoin et la condition d'hypométhylation des ARNr est égale ou supérieure à 0,1, et nous les avons groupés par région. Nous nous sommes intéressés plus particulièrement à un groupe de méthylation, conservées entre la levure et l'homme, et situées au niveau du tunnel de sortie du peptide (Figures 30 et 31). Nous avons localisé dans cette région six méthylation chez l'homme, dont quatre sont conservées chez la levure (Figure 32). Le génome de levure est plus facile à manipuler que le génome de cellules humaines et les méthylation sont conservées entre les deux organismes. Nous avons donc décidé de supprimer, dans le génome de la levure *S. cerevisiae*, les quatre méthylation présentes au niveau du tunnel de sortie du polypeptide.

2 – Construction de souches de levure déletées pour des snoARNs

Pour cela je délète dans le génome de la souche FY1679 (*MATa ura3-52 trp1Δ63 leu2Δ1 his3Δ200*) de *Saccharomyces cerevisiae* les quatre snoARNs responsables de ces méthylation : snR39B pour la méthylation Gm805, snR58 pour la méthylation Cm663, snR24 dans l'intron du gène *ASC1* pour la méthylation Cm1437, et snR18 dans l'intron du gène *EFB1* pour la méthylation Am649. Dans le cas des snoARNs snR24 et snR18, je délète également les gènes *ASC1* et *EFB1*. La protéine Asc1p est une protéine plateforme associée au ribosome et aux rôles multiples, non essentielle. La protéine Efb1p est une protéine de l'elongation de la traduction, qui est essentielle. Pour la délétion des snoARNs, j'utilise un système qui permet l'excision des marqueurs de sélection insérés aux locus à chaque délétion afin de ne pas les multiplier dans le génome de la souche³¹⁸. La stratégie repose sur l'amplification, à partir d'un plasmide, d'une séquence comprenant un marqueur de sélection, flanquée de part et d'autres par des séquences de reconnaissance de l'endonucléase I-SceI et de séquences répétées. Après délétion de chacun des snoARNs par une telle construction et sélection sur milieu approprié, l'endonucléase est introduite sur plasmide et son expression est induite. Elle effectue une coupure au niveau de ses séquences de reconnaissance, et cette coupure

Position chez l'homme	Position chez la levure	snoARN correspondant
Am1313	Am649	snR18 (intron de EFB1)
Cm1327	Cm663	snR58
Gm1509	Gm805	snR39B
Cm2338	Cm1437	snR24 (intron de ASC1)

Figure 32 : Liste des positions méthylées conservées entre la levure et l'homme au niveau du tunnel de sortie

Le cadre orange indique les snoARNs délétés dans la souche FYΔ(39B-58-24) pASC1 et le cadre bleu indique les snoARNs délétés dans la souche FYΔ(39B-58-18) pEFB1.

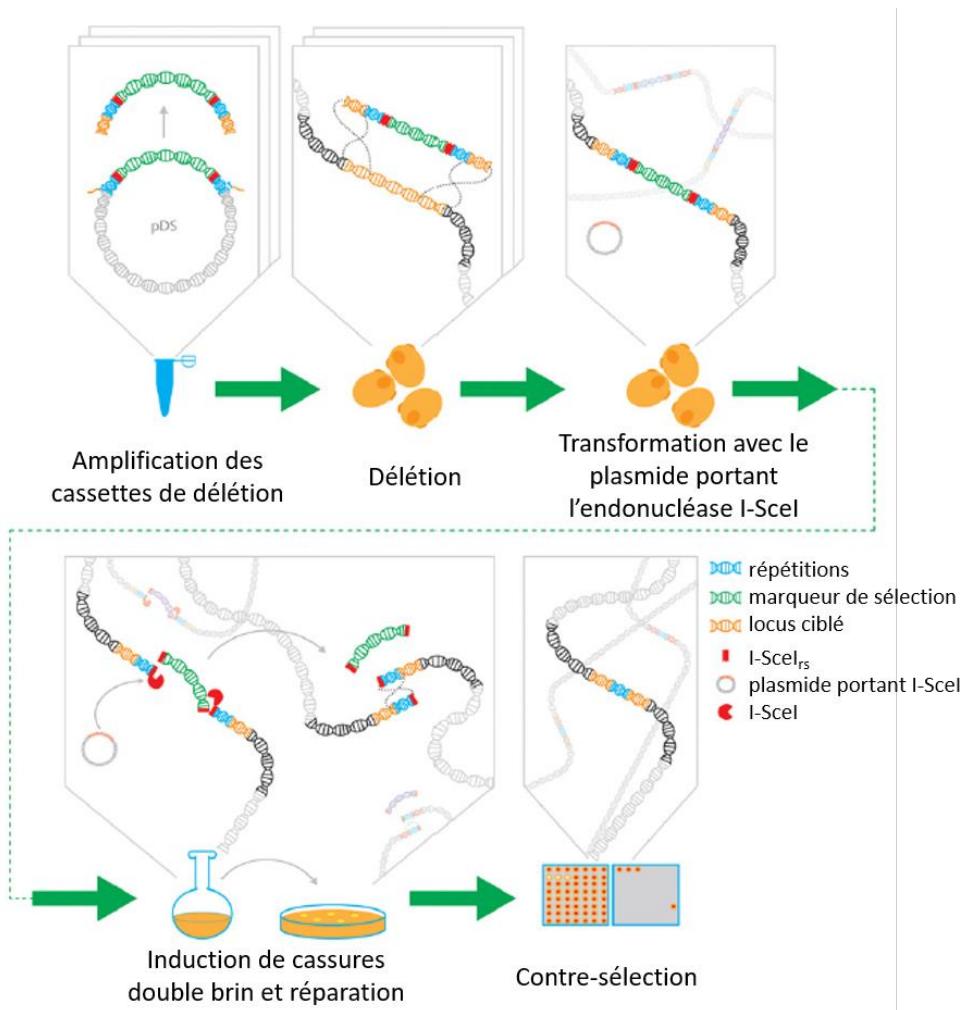


Figure 33 : Stratégie d'excision des marqueurs de sélection pour la délétion des snoRNAs

La stratégie repose sur l'amplification de la cassette de délétion à partir d'un plasmide. Cette cassette comprend la séquence d'un marqueur de sélection, flanquée de part et d'autres par des séquences de reconnaissance de l'endonucléase I-SceI et de séquences répétées. Après délétion de chacun des snoRNAs par une telle construction et sélection sur milieu approprié, l'endonucléase est introduite sur plasmide et son expression est induite. Elle effectue une coupure au niveau de ses séquences de reconnaissance, et cette coupure est réparée par homologie de séquence entre les séquences répétées présentes à chaque extrémité de la coupure. Ainsi, le snoARN a été déleté et le marqueur excisé, ce qui le rend de nouveau disponible pour une construction ultérieure.

est réparée par homologie de séquence entre les séquences répétées présentes à chaque extrémité de la coupure. Ainsi, le snoARN a été délégué et le marqueur excisé, ce qui le rend de nouveau disponible pour une construction ultérieure (Figure 33).

La déletion de snR39B a été effectuée de manière classique, avec insertion du marqueur de sélection *HIS3*. Les autres snoARNS ont été délégués suivant la stratégie d'excision de marqueur de sélection que je viens de décrire. J'ai d'abord délégué snR58, puis le gène *ASC1* avec snR24 dans l'intron. Cette souche est nommée FYΔ(39B-58-24). La séquence codante d'*ASC1* est ensuite apportée sur un plasmide, dans la souche FYΔ(39B-58-24) pASC1. Une autre souche a été construite, qui elle est déléguée pour les trois snoARNS snR39B, snR58 et snR18. Le gène *EFB1* qui contient snR18 étant essentiel, il a fallu que j'apporte d'abord dans la souche sa séquence codante sur plasmide. Puis j'ai délégué le gène *EFB1* avec snR18 dans l'intron, au locus, grâce à une cassette dont les extrémités ne sont pas présentes sur le plasmide contenant la séquence codante du gène *EFB1*. Ainsi, seule la séquence au locus peut être déléguée. Cette souche est nommée FYΔ(39B-58-18) pEFB1. La souche déléguée pour les quatre snoARNS n'est pas encore construite.

3 – Impact sur la croissance cellulaire

J'ai effectué différents tests sur les deux souches FYΔ(39B-58-24) pASC1 et FYΔ(39B-58-18) pEFB1.

Tout d'abord, j'ai cherché à démontrer si la déletion de trois snoARNS responsables des méthylations dans le tunnel de sortie du peptide affectait la croissance cellulaire. J'ai mis en culture ces deux souches ainsi que la souche témoin FY1679 et mesuré à intervalles réguliers la densité optique des cultures. J'ai ainsi obtenu des courbes de croissance et calculé le temps de génération en phase exponentielle pour les trois souches (Figure 34). Dans les conditions de culture de cette expérience, la souche FY1679 a un temps de génération de 3,2h et les souches FYΔ(39B-58-24) pASC1 et FYΔ(39B-58-18) pEFB1 respectivement de 4,3h et 4h. On observe donc une légère augmentation du temps de génération des souches pour lesquelles j'ai supprimé trois méthylations à proximité du tunnel de sortie du polypeptide.

J'ai ensuite vérifié la biogenèse et l'assemblage des ribosomes dans les souches FYΔ(39B-58-24) pASC1 et FYΔ(39B-58-18) pEFB1 avec des profils de polysomes. Sur le profil de la souche sauvage FY1679, on observe un profil normal : le pic de 60S est deux fois supérieur au pic de 40S (l'ARNr 28S constituant du 60S est deux fois plus grand que l'ARNr 18S constituant de la 40S) ce qui indique une biogenèse correcte des sous-unités des ribosomes. On distingue aussi un grand pic de monosomes 80S et des polysomes, ce qui confirme la formation de polysomes actifs (Figure 35). Pour la souches FYΔ(39B-58-24) pASC1 on observe une diminution du pic de 60S, ainsi que des épaulements appelés

Courbes de croissance des souches FY1679, FY Δ (39B-58-24) pASC1 et FY Δ (39B-58-18) pEFB1

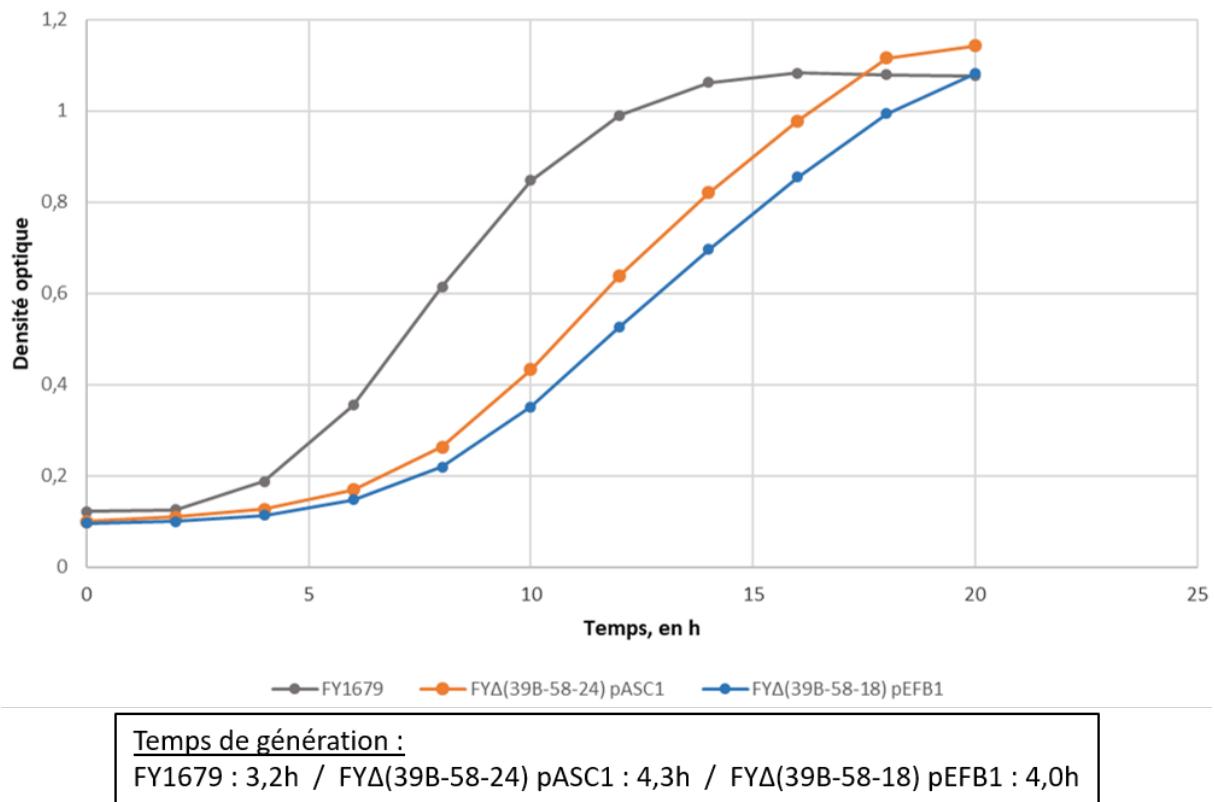


Figure 34 : Courbes de croissance des souches FY1679, FY Δ (39B-58-24) pASC1 et FY Δ (39B-58-18) pEFB1

A partir des courbes de croissance des souches FY1679, FY Δ (39B-58-24) pASC1 et FY Δ (39B-58-18) pEFB1, les temps de génération ont été calculés en phase exponentielle de croissance.

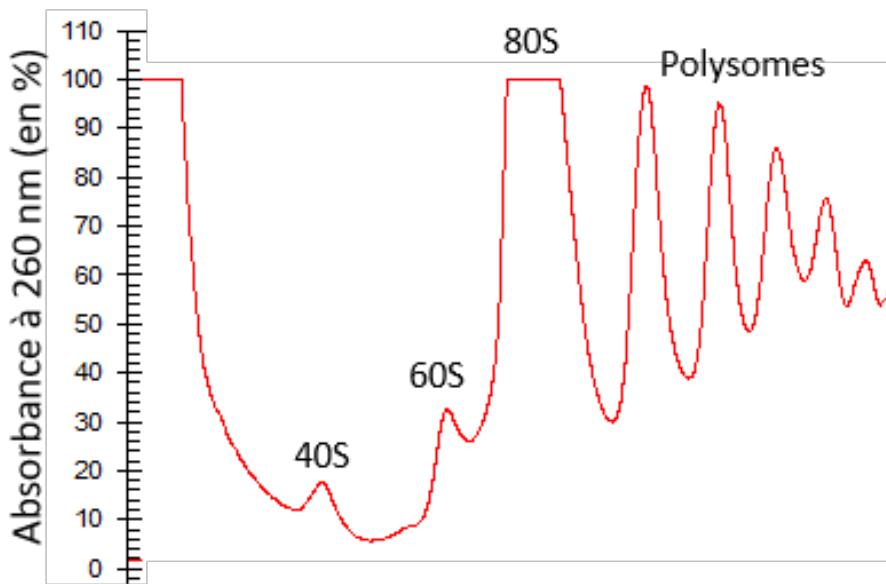


Figure 35 : profil de polysomes de la souche FY1679

Des extraits de polysomes sont récoltés à partir de la souche FY1679 après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucrose 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm.

"halfmers", ce qui indique des défauts de biogenèse de la sous-unité 60S (Figure 36). Il a déjà été démontré que la délétion de snR24 entraîne des défauts de biogenèse de la sous-unité 60S³¹⁹. Ici, les défauts observés sont probablement dus à la délétion de snR24 mais ils pourraient également être causés par l'absence de l'un ou de plusieurs des autres snoARNs que j'ai délétés. Pour vérifier cela, j'ai réexprimé snR24 sur plasmide dans la souche FYΔ(39B-58-24) pASC1 et j'ai fait un profil de polysomes (Figure 37). Sur ce profil, on peut voir des pics de 40S et de 60S qui indiquent une biogenèse correcte des sous-unités et une disparition des « halfmers ». La réexpression de snR24 est donc suffisante pour corriger les défauts de biogenèse de la sous-unité, c'est bien uniquement la délétion de ce snoARN qui cause ces défauts. Pour la souche FYΔ(39B-58-18) pEFB1, j'observe un ratio normal entre les pics de 40S et de 60S et un grand pic de monosomes 80S, qui indique une biogenèse des sous-unités et un assemblage du ribosome corrects. On distingue également des polysomes qui montrent une traduction active (Figure 38).

Enfin, j'ai réalisé des tests de résistance à différents antibiotiques. J'ai choisi de tester la cycloheximide, un inhibiteur de la traduction ; la paromomycine, qui induit des erreurs d'incorporation des acides aminés au cours de la traduction et l'hygromycine B, qui induit également des erreurs d'incorporation des acides aminés. J'ai tout d'abord établi les concentrations optimales de chaque antibiotique avec la souche sauvage FY1679, de telle sorte qu'à cette concentration, un léger effet sur la croissance cellulaire soit visible. Pour chaque antibiotique, les trois souches FY1679, FYΔ(39B-58-24) pASC1 et FYΔ(39B-58-18) pEFB1 sont déposées en goutte sur la boîte, de façon à obtenir cinq gouttes pour chaque souche, chaque goutte étant diluée par un facteur 10 par rapport à la goutte précédente. La gamme s'étend ainsi de 10^5 cellules à 10^1 cellules (Figure 39). Pour les trois antibiotiques testés, on n'observe aucun effet sur la croissance cellulaire. Toutefois, le profil de croissance de la souche témoin FY1679 est différent de celui obtenu dans l'expérience présentée dans la partie suivante (« La protéine Asc1p est impliquée dans la translecture du codon stop ») et lors de la mise au point (Figure 40). L'expérience est donc à refaire afin de valider ou d'infirmer ce résultat préliminaire.

4 – Etude de l'impact sur la fidélité de la traduction

Après avoir étudié l'impact de la suppression des 2'-O-méthylations du tunnel de sortie du peptide sur la biogenèse du ribosome et la croissance cellulaire, j'ai testé l'impact sur la fidélité de la traduction à travers des mesures des taux de translecture du TMV (codon UAG) et de décalage de cadre de lecture -1 IBV et +1 EST, à la fois dans la souche FY1679 en contrôle et dans les souches FYΔ(39B-58-24) pASC1 et FYΔ(39B-58-18) pEFB1.

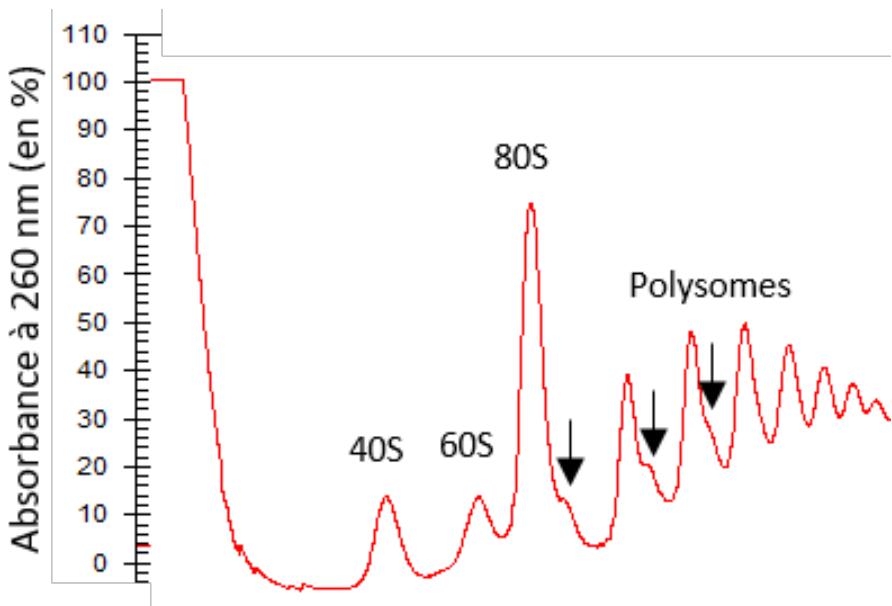


Figure 36 : profil de polysomes de la souche FYΔ(39B-58-24) pASC1

Des extraits de polysomes sont récoltés à partir de la souche FYΔ(39B-58-24) pASC1 après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucrose 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm. Les flèches noires indiquent des « halfmers ».

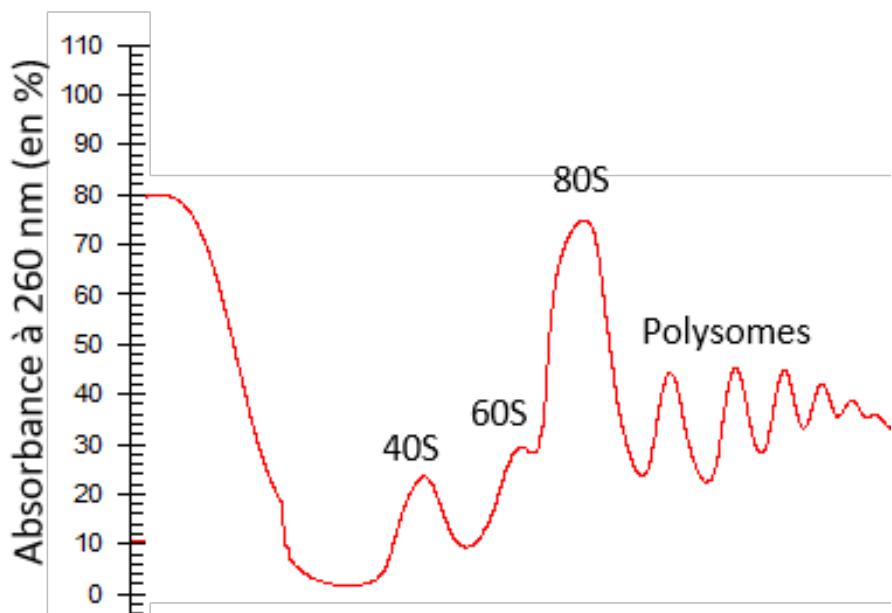


Figure 37 : profil de polysomes de la souche FYΔ(39B-58-24) pASC1 pU24

Des extraits de polysomes sont récoltés à partir de la souche FYΔ(39B-58-24) pASC1 pU24 après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucrose 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm.

J'ai effectué tout d'abord des mesures sur la souche FYΔ(39B-58-24) pASC1 avec la souche FY1679 en témoin (Figure 41). Le taux de translecture du stop TMV UAG est de 23,9% pour la souche FYΔ(39B-58-24) pASC1 et 26,2% pour la souche FY1679 ; le taux de décalage de cadre de lecture -1 IBV est de 19% pour la souche FYΔ(39B-58-24) pASC1 et 19,9% pour la souche FY1679 ; et le taux de décalage de cadre de lecture +1 EST est de 5,5% pour la souche FYΔ(39B-58-24) pASC1 et 5,1% pour la souche FY1679. En appliquant le test statistique de Wilcoxon-Mann-Whitney, ces valeurs ne présentent aucune variation significative entre la souche FYΔ(39B-58-24) pASC1 et la souche FY1679. J'ai ensuite effectué des mesures sur la souche FYΔ(39B-58-18) pEFB1 avec la souche FY1679 en témoin (Figure 42). Le taux de translecture du stop TMV UAG est de 18,8% pour la souche FYΔ(39B-58-18) pEFB1 et 23,3% pour la souche FY1679 ; le taux de décalage de cadre de lecture -1 IBV est de 21,2% pour la souche FYΔ(39B-58-18) pEFB1 et 19,5% pour la souche FY1679 ; et le taux de décalage de cadre de lecture +1 EST est de 5,4% pour la souche FYΔ(39B-58-18) pEFB1 et 5,8% pour la souche FY1679. En appliquant le test statistique de Wilcoxon-Mann-Whitney, ces valeurs ne présentent aucune variation significative entre la souche FYΔ(39B-58-18) pEFB1 et la souche FY1679.

Ainsi, les résultats ne montrent aucun effet significatif sur ces événements de recodage, et donc sur la fidélité de la traduction, ni dans la souche FYΔ(39B-58-24) pASC1 ni dans la souche FYΔ(39B-58-18) pEFB1.

II – Discussion

L'extinction de la fibrillarine par un shRNA entraîne une baisse globale mais partielle de la méthylation des ARNs ribosomiques. Après avoir réalisé des études sur des cellules HeLa dans lesquelles les taux de méthylation des ARNr sont globalement diminués, j'ai choisi de me focaliser sur un groupe restreint de méthylations. Pour cela, j'ai choisi de changer d'organisme et de mener cette étude chez la levure *Saccharomyces cerevisiae*. Dans ce but, j'ai localisé les méthylations affectées par la baisse du taux de fibrillarine sur une structure du ribosome de *S. cerevisiae*. Je me suis intéressée plus particulièrement à un groupe de méthylations conservées entre la levure et l'homme et regroupées dans la même région, au niveau du tunnel de sortie du polypeptide. Ce tunnel joue un rôle essentiel sur le repliement correct des protéines, qui affecte leur activité. La perte de ces méthylations pourrait donc avoir comme effet principal un repliement incorrect des protéines, mais aussi affecter la croissance cellulaire ou la fidélité de la traduction. J'ai alors déléte trois des quatre snoARNs correspondants à ces méthylations suivant deux combinaisons différentes et j'ai effectué des tests sur les souches construites.

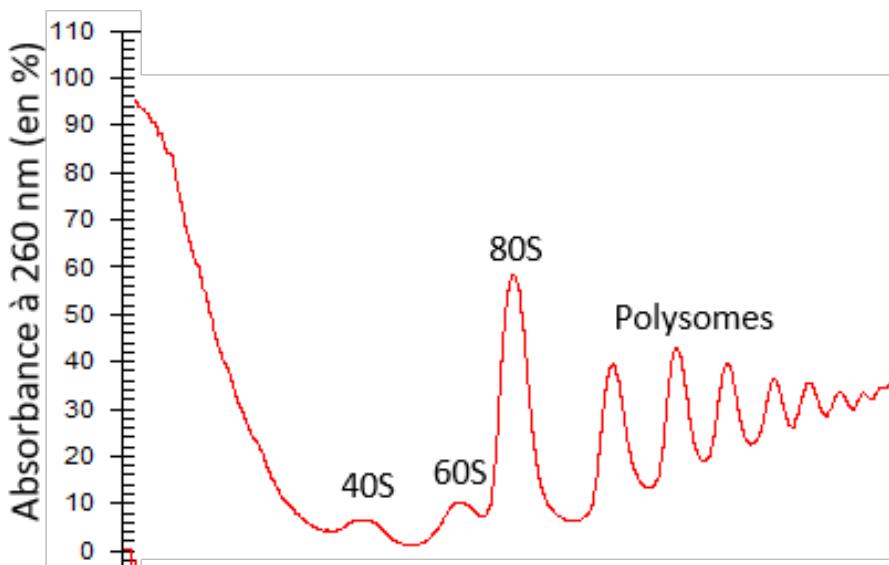


Figure 38 : profil de polysomes de la souche FYΔ(39B-58-18) pEFB1

Des extraits de polysomes sont récoltés à partir de la souche FYΔ(39B-58-18) pEFB1 après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucrose 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm.

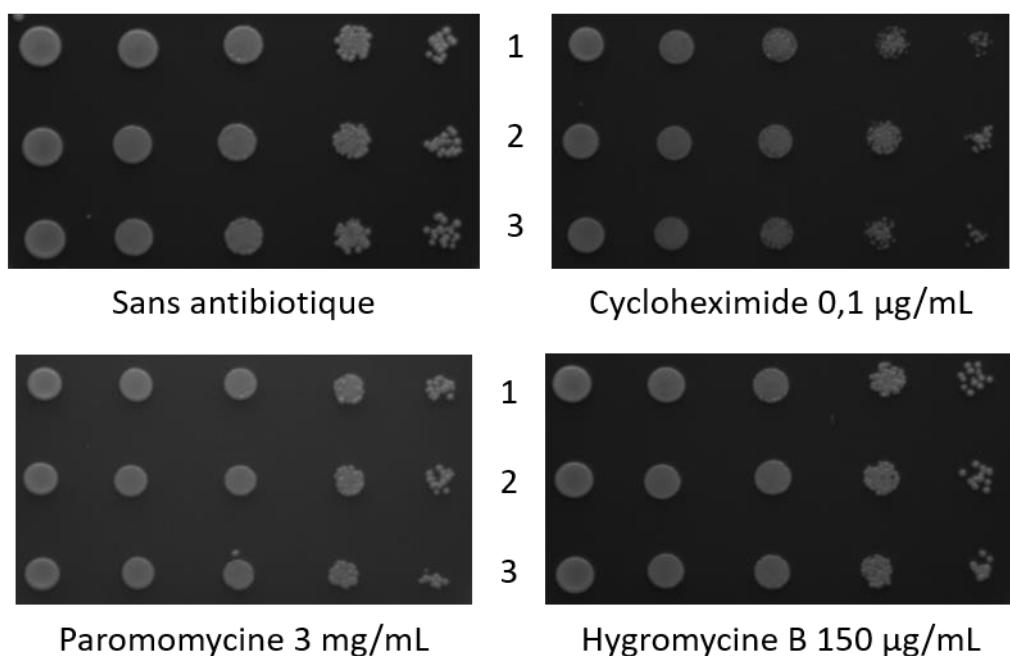


Figure 39 : Test en goutte de la sensibilité des souches FY1679, FYΔ(39B-58-24) pASC1 et FYΔ(39B-58-18) pEFB1 à différents antibiotiques

Les souches FY1679 (1), FYΔ(39B-58-24) pASC1 (2) et FYΔ(39B-58-18) pEFB1 (3) ont été déposées à différentes concentrations sur un milieu solide avec différents antibiotiques. Ces différentes concentrations sont, de gauche à droite, 10^5 , 10^4 , 10^3 , 10^2 et 10^1 .

Par ces tests, j'ai tout d'abord confirmé que la délétion du snoARN snR24 entraînait des défauts de biogenèse de la sous-unité 60S et j'ai montré que la perte de trois méthylations localisées au niveau du tunnel de sortie du peptide augmentait légèrement le temps de génération des cellules. Puis j'ai montré que la sensibilité aux antibiotiques cycloheximide, paromomycine et hygromycine B n'était pas affectée. Enfin, j'ai constaté que la fidélité de la traduction à travers la mesure des taux de translecture et de décalage de cadre de lecture n'était pas touchée. Il semblerait donc que la perte de trois des quatre méthylations au niveau du tunnel de sortie du polypeptide n'ait pas d'effet mesurable sur la traduction par les techniques que j'ai utilisées. Il faut donc maintenant que je réalise les mêmes tests sur la souche délétée pour les quatre snoARNs.

Les tests que j'ai effectués ne mesurent pas directement le repliement des protéines. Il serait important de développer des systèmes qui permettent de tester l'impact de la perte des méthylations autour du tunnel de sortie du peptide sur le repliement des protéines.

Tout d'abord, j'ai réalisé un test de sensibilité à des molécules qui bloquent le ribosome ou entraînent des misincorporations. Parmi les composés chimiques qui existent, on peut trouver la canavanine, qui est connue pour induire un repliement incorrect des protéines. Il serait donc intéressant d'effectuer un test de sensibilité à la canavanine sur les souches délétées pour les snoARNs responsables des méthylations autour du tunnel de sortie du peptide.

Il est connu qu'en cas de repliement incorrect des protéines au niveau du réticulum endoplasmique (RE), la voie de réponse UPR est activée³²⁰. L'accumulation de protéines mal conformées au RE induit l'activité endonucléase de la protéine senseur Ire1p³²¹, qui intervient dans l'épissage de l'ARNm du gène *HAC1*, ce qui active sa traduction³²². La protéine Hac1p est un facteur de transcription qui active l'expression des gènes de la voie UPR³²³. Nous pouvons alors définir plusieurs moyens de vérifier si le repliement des protéines est correct ou non. La première technique consisterait à mesurer l'activité activatrice de la transcription de la protéine Hac1p avec un système rapporteur qui comporte la séquence codante d'un gène rapporteur (par exemple la luciférase) sous contrôle d'un promoteur induit par Hac1. Un second test serait de mesurer les niveaux des transcrits des protéines impliquées dans la voie de réponse UPR, par RT-qPCR par exemple.

De plus, il est connu que le Cadmium (Cd^{2+}) induit des défauts de repliement des protéines au niveau du réticulum endoplasmique³²⁰. Nous pourrions donc envisager un test de sensibilité au cadmium similaire aux tests de sensibilité aux antibiotiques.

Un autre moyen de mesurer le repliement des protéines est d'utiliser un système rapporteur avec la protéine GFP (Green Fluorescent Protein). L'activité de fluorescence de la GFP nécessite un repliement correct de la protéine. Un système rapporteur avec la GFP en fusion avec une protéine,

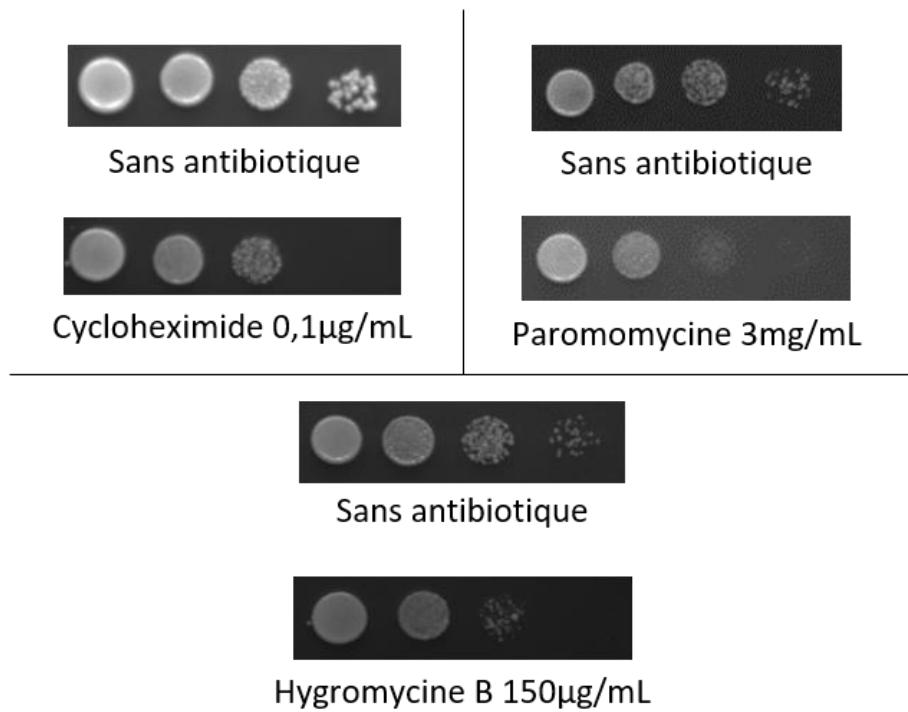


Figure 40 : Mise au point des tests en goutte de la sensibilité à différents antibiotiques

La souches FY1679 a été déposée à différentes concentrations sur un milieu solide avec les antibiotiques cycloheximide, paromomycine et hygromycine B. Les différentes concentrations de levure sont, de gauche à droite : 10^5 , 10^4 , 10^3 , 10^2 et 10^1 .

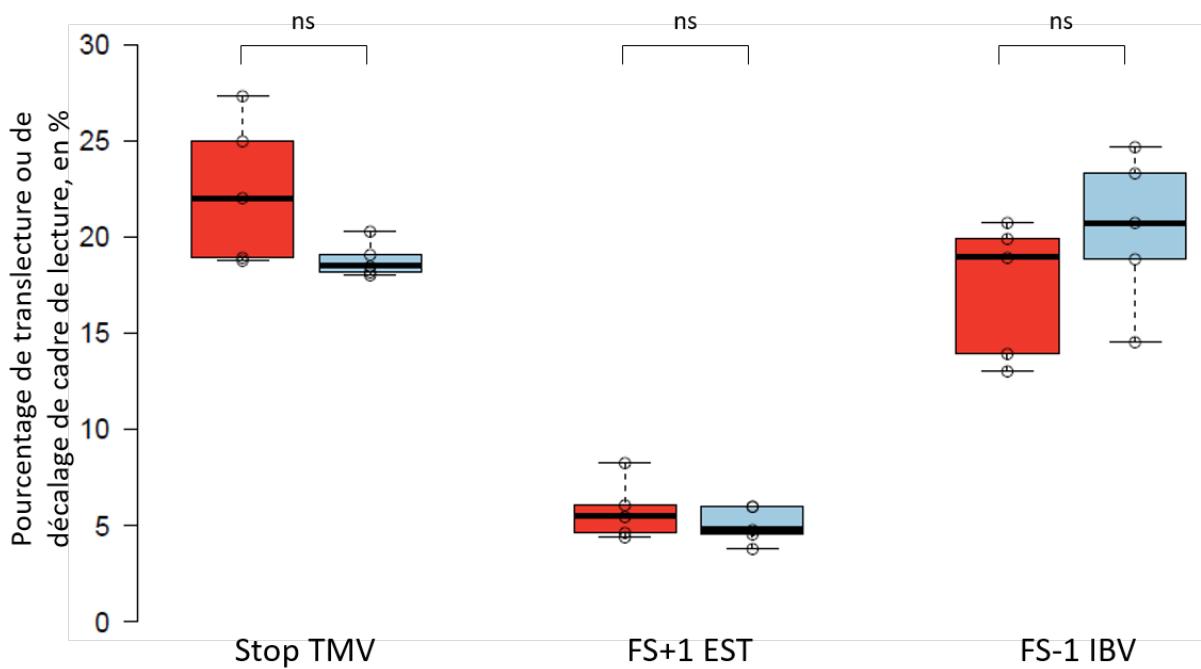


Figure 41 : Translecture du codon stop TMV (UAG) et décalages de cadre de lecture +1 EST et -1 IBV dans les souches FY1679 et FYΔ(39B-58-24) pASC1

La souche FY1679 est en rouge tandis que la souche FYΔ(39B-58-24) pASC1 est en bleu. Un test statistique de Wilconxon-Mann-Whitney a été appliqué sur les données. ns : non significatif

dont le repliement influe sur l'activité de la GFP, a été mis au point chez les bactéries³²⁴ puis adapté chez la levure³²⁵ et dans des cellules de mammifères³²⁶. Ainsi la mesure de la fluorescence de la GFP rend compte d'un éventuel défaut de repliement des protéines.

L'utilisation de ces systèmes de mesure directe du repliement des protéines et la caractérisation du phénotype de la souche déletée pour les quatre snoARNs responsables des quatre méthylations autour du tunnel de sortie du peptide donneraient davantage d'informations pour déterminer si la délétion de ces quatre méthylations induit des effets sur la traduction.

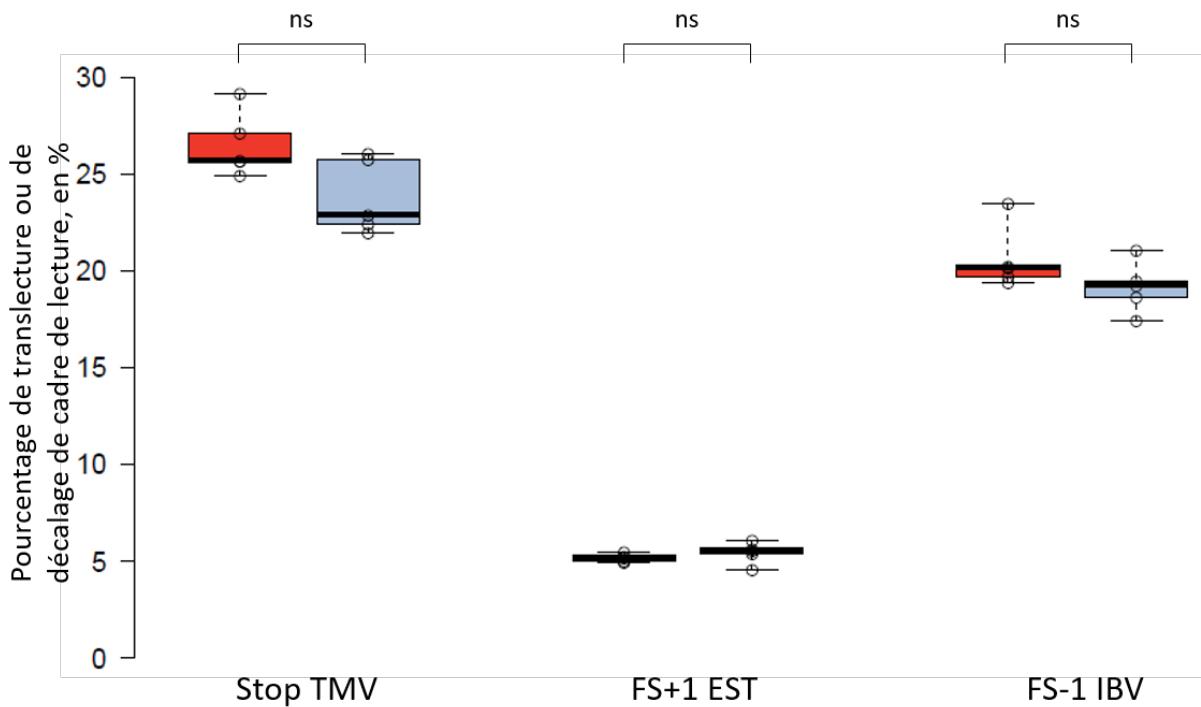


Figure 42 : Translecture du codon stop TMV (UAG) et décalages de cadre de lecture +1 EST et -1 IBV dans les souches FY1679 et FYΔ(39B-58-18) pEFB1

La souche FY1679 est en rouge tandis que la souche FYΔ(39B-58-18) pEFB1 est en bleu. Un test statistique de Wilconxon-Mann-Whitney a été appliqué sur les données. ns : non significatif

Résultats & Discussion

La protéine Asc1p est impliquée
dans la translecture du codon stop

La protéine Asc1p est impliquée dans la translecture du codon stop

I - Résultats

Au cours de l'étude des méthylations dans le tunnel de sortie du peptide, j'ai été amenée à déléter la séquence codante du gène *ASC1*, puisque son intron contient le snoARN snR24 (aussi appelé U24) responsable de la méthylation en position 1437. Les résultats obtenus dans cette partie font l'objet d'un article en cours d'écriture.

1 – La protéine Asc1p est une protéine plateforme conservée chez les eucaryotes et impliquée dans de nombreux processus traductionnels et cellulaires

Asc1p est une protéine plateforme associée à la sous-unité 40S du ribosome dans une stoechiométrie 1:1, parfois considérée comme une protéine ribosomique³²⁷. Elle est située au niveau du tunnel de sortie de l'ARNm, à proximité des protéines qui interagissent avec l'ARNm³²⁸.

Elle possède un domaine avec sept répétitions WD40 (section d'environ 40 acides aminés souvent terminée par le dipeptide tryptophane-aspartate) qui adopte une conformation d'hélice tripale³²⁹ et qui sert de plateforme pour les interactions protéines-protéines. Il a été montré qu'Asc1p est impliquée dans des voies de signalisation cellulaire et est associée à la synthèse protéique à travers sa localisation à la tête de la sous-unité 40S du ribosome.

Une des premières fonctions proposée pour Asc1p est une répression de la traduction car en son absence, l'activité traductionnelle des ribosomes augmente³²⁷. Asc1p est nécessaire pour certains événements traductionnels spécifiques. Une étude globale par RiboSeq a montré qu'en son absence, la traduction d'ARNm courts est plus affectée que celle des autres ARNm³³⁰. La délétion du gène *ASC1* permet aux ribosomes de traduire des régions qui induisent normalement leur blocage, comme des signaux polylysines (queue poly(A))^{331,332} et des suites d'arginine (acide aminé chargé positivement) CGA^{333,334}. En l'absence d'Asc1p, les ribosomes continuent à décoder les codons CGA mais ne maintiennent pas le cadre de lecture, provoquant ainsi des changements du cadre de lecture. Asc1p participe donc au maintien du cadre de lecture en induisant le blocage des ribosomes au niveau des répétitions CGA³³⁴. Des études plus récentes suggèrent qu'elle participe également à la reconnaissance des ribosomes bloqués afin d'induire les voies de contrôle qualité des ARNm (NGD, NSD). Il a ainsi été montré qu'Asc1p, avec Hel2p et Slh1p, participe à la reconnaissance des ribosomes en pause et au recrutement du complexe RQC (Ribosome-bound Quality Control) pour la dégradation du peptide

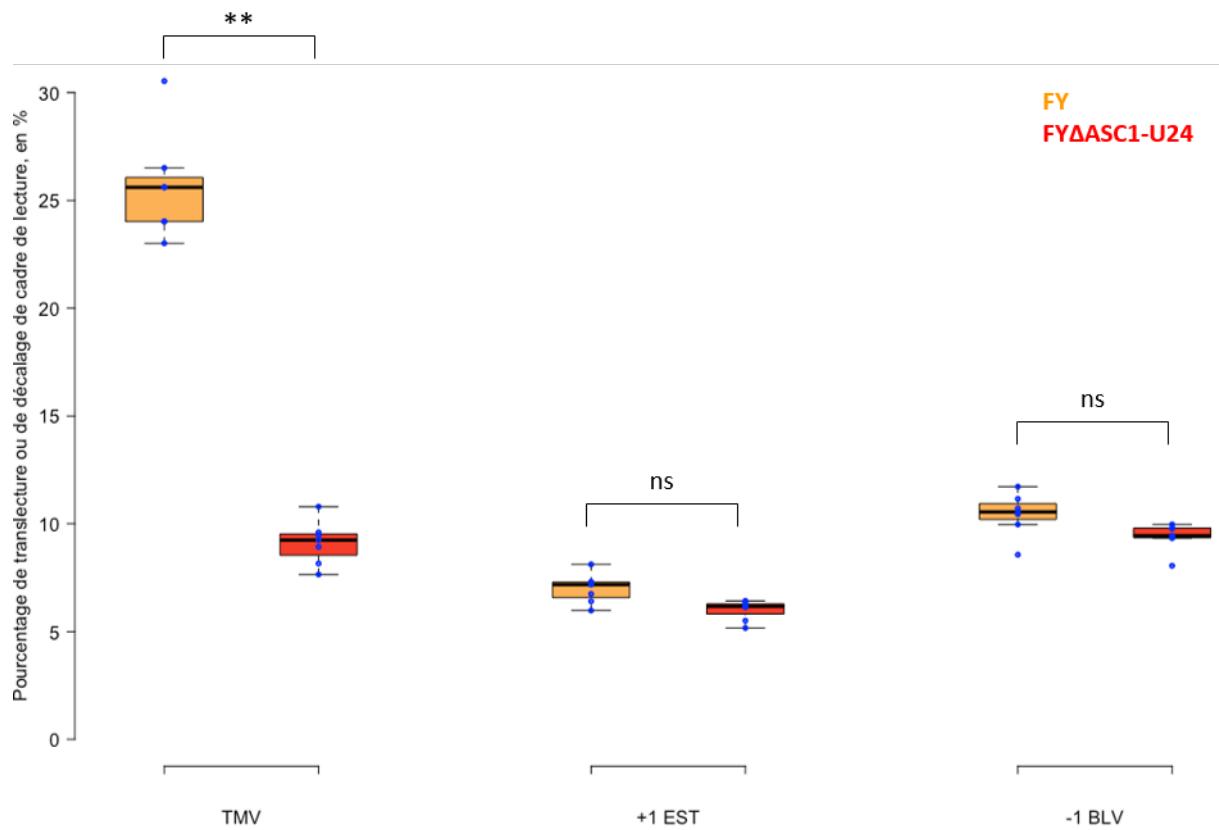


Figure 43 : Translecture du codon stop TMV (UAG) et décalages de cadre de lecture +1 EST et -1 BLV dans les souches FY1679 et FYΔ(ASC1-U24)

Un test statistique de Wilcoxon-Mann-Whitney a été appliqué sur les données. ** $p \leq 0,01$; ns : non significatif.

naissant³³⁵. Asc1p est aussi impliquée dans le clivage endonucléolytique à l'extrémité 3' des ARNm qui ne présentent pas de codon stop³³⁶ et dans la dégradation des ARNr 18S non-fonctionnels (qui fait intervenir les mêmes facteurs que la voie de dégradation NGD), probablement à travers son interaction avec Rps3 qui lui-même interagit avec Hbs1^{337,338}. De plus, il a été montré qu'Asc1p affecte la traduction de certains ARNm à travers leur région 5'UTR³³⁹ et à travers son interaction avec des protéines qui se lient aux ARNm, comme Scp160p⁴⁷. Il a aussi été suggéré qu'elle pourrait avoir un rôle de régulation de l'initiation de la traduction. En effet, il a été montré qu'Asc1p interagit avec la sous-unité eIF3c³⁴⁰ et qu'Asc1p intervient dans la phosphorylation de protéines de la traduction, dont des facteurs de l'initiation³⁴¹.

En plus de son implication dans la phosphorylation de protéines de la traduction, Asc1p, qui est elle-même phosphorylée à de nombreux sites, influence la phosphorylation d'au moins 90 protéines sur 120 sites à travers son rôle de plateforme³⁴¹. La plupart de ces protéines sont des régulateurs de processus fondamentaux : traduction, turnover des protéines, transport des protéines, organisation du cytosquelette, signalisation cellulaire.

Une autre étude fondée sur des analyses protéomiques et transcriptomiques d'une souche déletée pour le gène *ASC1* a révélé des groupes fonctionnels de protéines et d'ARNm qui présentent une expression altérée, et a déterminé des processus cellulaires qui sont affectés³³⁹. Ces processus sont la traduction, l'homéostasie du fer, la réponse aux phéromones, l'intégrité de la paroi cellulaire et le métabolisme énergétique (glycolyse, respiration, fermentation, stress oxydatif). Il a été montré qu'Asc1p régule post-traductionnellement les taux de facteurs de transcription impliqués dans les voies de signalisation MAPK de la croissance invasive et dans la réponse aux phéromones. Elle régule également l'initiation de la traduction des facteurs de transcription (comme Ste12p, Tec1p et Phdp1) responsables de la majorité des variations transcriptionnelles observées dans la souche Δ ASC1, à travers leurs régions 5'UTR. Cela suggère une interaction directe d'Asc1p avec les facteurs impliqués dans l'initiation de la traduction comme démontré pour eIF3³⁴⁰.

L'environnement protéique d'Asc1p a été étudié grâce à la technique BioID, qui permet le marquage des protéines à proximité d'une protéine d'intérêt, ici Asc1p³⁴². En phase exponentielle, une grande partie des voisins d'Asc1p sont des protéines qui lient l'ARNm (Sro9p, Gis2p, Scp160p par exemple). On retrouve également des facteurs d'initiation de la traduction (eIF4E, sous-unité a d'eIF3), des facteurs de transcription (Spt5p, Mbf1p) et de façon surprenante la protéine Stm1p. Cette protéine maintient l'association des sous-unités du ribosome en l'absence d'ARNm, et ne peut pas être présent au centre de décodage lorsqu'un ARNm y est déjà. Une hypothèse avancée par les auteurs est que Stm1p reste à proximité du ribosome en traduction active au niveau d'Asc1p, dans l'attente d'une situation qui nécessiterait son recrutement au centre de

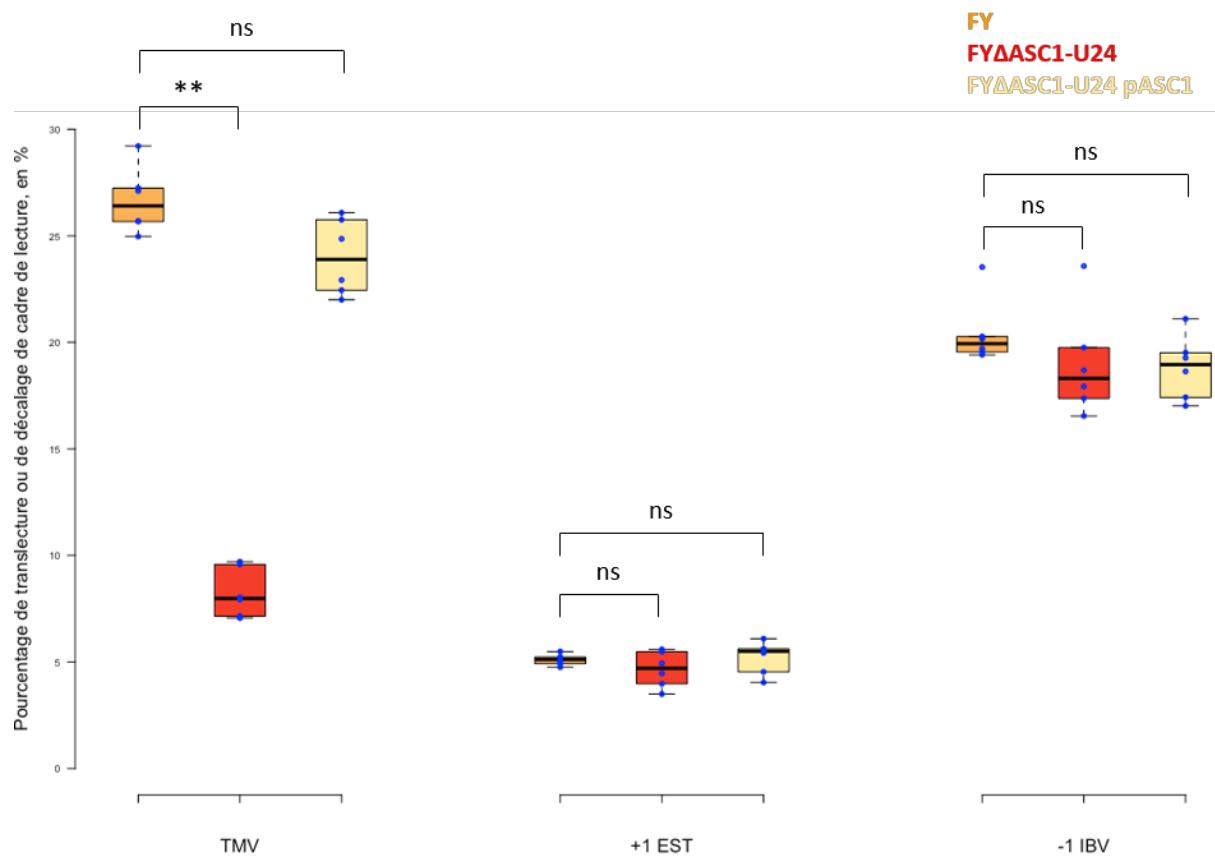


Figure 44 : Translecture du codon stop TMV (UAG) et décalages de cadre de lecture +1 EST et -1 IBV dans les souches FY1679, FYΔ(ASC1-U24) et FYΔ(ASC1-U24) pASC1

Un test statistique de Wilcoxon-Mann-Whitney a été appliqué sur les données. ** $p \leq 0,01$; ns : non significatif.

décodage du ribosome. L'environnement protéique d'Asc1p a également été étudié en condition de privation de glucose. La plupart des protéines qui colocalisent avec Asc1p durant la phase exponentielle disparaissent et les protéines ribosomiques des deux sous-unités s'accumulent dans l'entourage d'Asc1p, ce qui suggère une agrégation des ribosomes.

Asc1p est une protéine hautement conservée chez les eucaryotes. Chez les mammifères, son orthologue est RACK1, pour Receptor for Activated Protein C Kinase 1³²⁷. Tout comme Asc1p, RACK1 se lie à des facteurs de signalisation, comme des kinases et des phosphatases, et régule ainsi leur activité et la liaison à leurs cibles³⁴³. Chez les animaux et les plantes, l'absence de RACK1 est responsable de la létalité aux stades précoces de l'embryogenèse³⁴⁴ tandis que la délétion du gène *ASC1* n'a pas d'impact sur la croissance cellulaire³⁴⁵. De ce fait, il est plus facile d'étudier le rôle de cette protéine en délétant Asc1p chez la levure. Il a également été montré que RACK1 est nécessaire à la traduction au niveau de certains IRES viraux qui nécessitent des facteurs d'initiation³⁴⁶.

Au cours de ma thèse, je me suis intéressée à la fidélité de la traduction par la mesure d'événements de recodage que sont la translecture du codon stop et le décalage de cadre de lecture. Jusqu'à présent, aucune donnée n'a été publiée sur Asc1p concernant ces événements. J'ai alors cherché à savoir si la délétion du gène *ASC1* a un effet sur la translecture et le décalage de cadre de lecture.

2 – L'absence de la protéine Asc1p entraîne une diminution de la translecture des codons stop

J'ai comparé l'efficacité de translecture et de décalage de cadre de lecture entre une souche sauvage et une souche dans laquelle le gène *ASC1* a été délété. Ces dosages ont révélé une baisse d'un facteur environ 2,7 de la translecture du codon stop TMV (UAG) en l'absence d'Asc1p (24,8% dans la souche sauvage FY1679 et 9,3% dans la souche délétée pour *ASC1*) tandis que les événements de décalage de cadre de lecture -1 et +1 sont insensibles à la délétion du gène *ASC1* (10,6% et 9,6% de décalage de cadre de lecture -1 BLV, et 7,0% et 6,1% de décalage de cadre de lecture +1 EST respectivement dans les souches FY1679 et FYΔ(*ASC1*-U24)) (Figure 43). Dans un premier temps, afin de confirmer cet effet et de vérifier s'il était uniquement dû à l'absence de la protéine Asc1p et non à la délétion du snoARN snR24, j'ai effectué de nouvelles mesures du taux de translecture du codon stop TMV (UAG) et du décalage de cadre de lecture -1 IBV et +1 EST, dans les souches FY, FYΔ(*ASC1*-U24) et FYΔ(*ASC1*-U24) pASC1 (l'ADNc du transcript d'*ASC1* a été cloné dans un

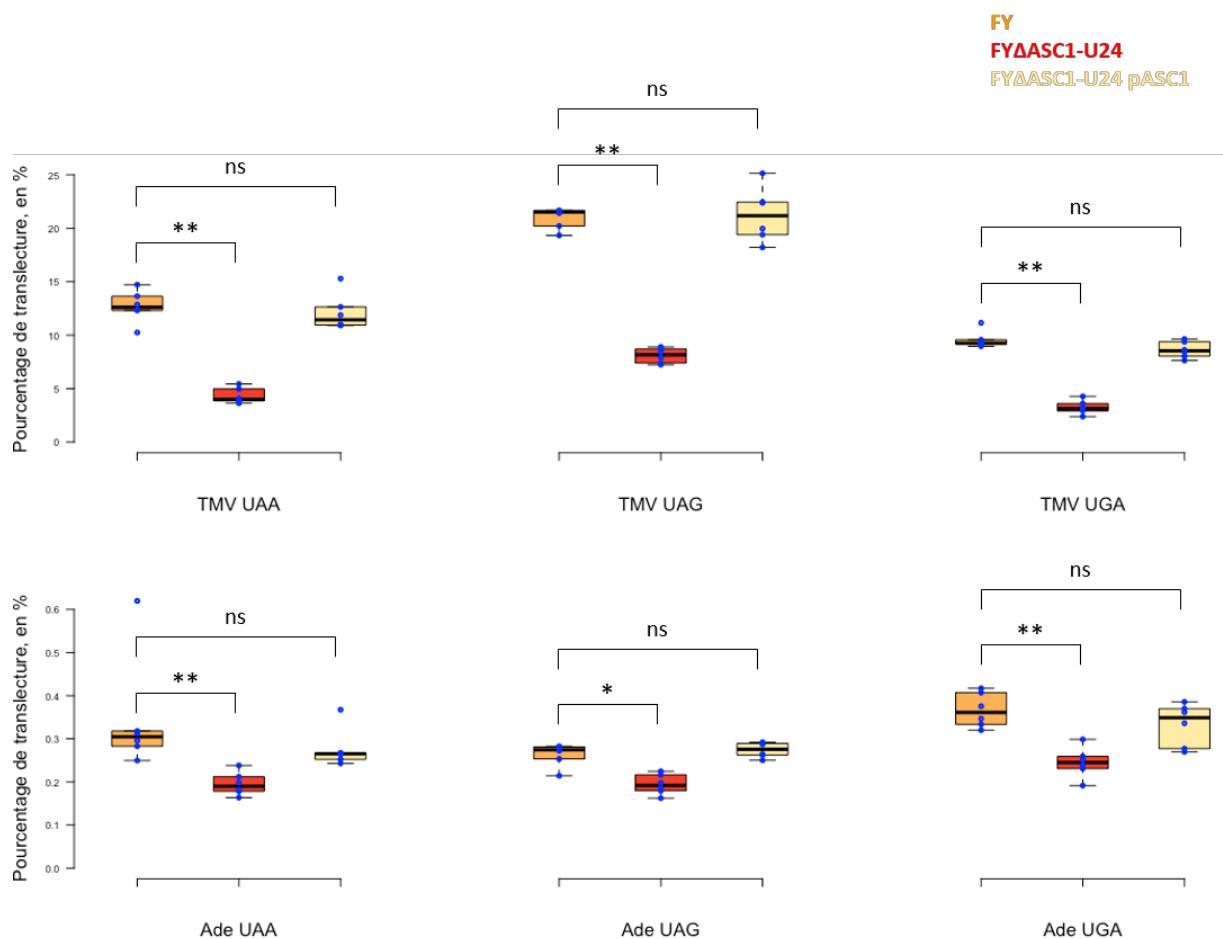


Figure 45 : Translecture du codon stop TMV (UAG) et décalages de cadre de lecture +1 EST et -1 IBV dans les souches FY1679, FY Δ (ASC1-U24) et FY Δ (ASC1-U24) pASC1

Un test statistique de Wilcoxon-Mann-Whitney a été appliqué sur les données. * $p \leq 0,05$; ** $p \leq 0,01$; ns : non significatif.

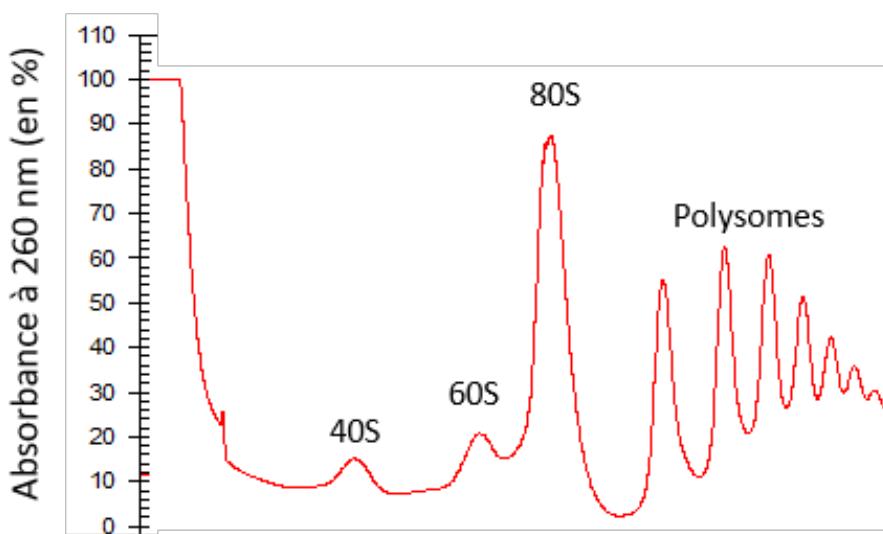


Figure 46 : Profil de polysomes de la souche FY1679

Des extraits de polysomes sont récoltés à partir de la souche FY1679 après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucrose 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm.

plasmide). Le taux de translecture du codon stop TMV est de 26,4% dans la souche FY, de 8,0% dans la souche FYΔ(ASC1-U24) et de 23,9% dans la souche FYΔ(ASC1-U24) pASC1 ; le taux de décalage de cadre de lecture -1 IBV est de 19,9% dans la souche FY, de 18,3% dans la souche FYΔ(ASC1-U24) et de 19,0% dans la souche FYΔ(ASC1-U24) pASC1 et le taux de décalage de cadre de lecture +1 EST est de 5,1% dans la souche FY, de 4,7% dans la souche FYΔ(ASC1-U24) et de 5,5% dans la souche FYΔ(ASC1-U24) pASC1 (Figure 44). Le test statistique de Wilcoxon-Mann-Whitney effectué sur ces mesures révèle une baisse significative de la translecture dans la souche FYΔ(ASC1-U24) par rapport à la souche FY et aucune différence entre les souches FY1679 et FYΔ(ASC1-U24) pASC1. Il ne montre également aucune différence significative de décalage de cadre de lecture -1 et +1 entre les trois souches. J'ai ainsi confirmé que la délétion du gène *ASC1* entraîne une baisse de la translecture du codon stop TMV (UAG) et n'a aucun effet sur le décalage de cadre de lecture -1 et +1. De plus, le taux initial de translecture est restauré quand le gène *ASC1* est ré exprimé sur plasmide, ce qui confirme que c'est bien l'absence d'Asc1p qui est responsable de cette baisse et non la délétion de snR24.

Par la suite, j'ai voulu savoir si cet effet sur la translecture était spécifique du codon stop TMV ou était généralisable à tous les codons stop. En effet, le codon stop TMV est un stop où la translecture est programmée, c'est à dire que le contexte nucléotidique a été sélectionné pour permettre un niveau de translecture très élevé.³⁴⁷ J'ai donc testé la translecture des trois codons stop UAA, UAG et UGA dans le contexte TMV et dans un contexte dit "anonyme" qui est celui du codon stop présent dans l'allèle *ade1-14* du gène *ADE1*. Pour les trois codons stops dans le contexte TMV, le taux de translecture diminue significativement quand le gène *ASC1* est délété, d'un facteur 2,6 à 3,1 selon le codon stop, et est restauré quand le gène *ASC1* est exprimé sur plasmide (Figure 45). Ce résultat montre que la délétion affecte la translecture des trois stops UAA, UAG et UGA dans le contexte TMV de translecture programmée. Pour les trois codons stops dans le contexte *ADE1*, j'ai également démontré que le taux de translecture diminue significativement quand le gène *ASC1* est délété, d'un facteur 1,4 à 1,6 selon la nature du codon stop, et est restauré quand le gène *ASC1* est exprimé sur plasmide (Figure 45). Ainsi, l'absence d'Asc1p affecte la translecture des codons stops quel que soit le contexte nucléotidique dans lequel ils se trouvent.

3 – La délétion du gène *ASC1* augmente la sensibilité à des antibiotiques

Asc1p est une protéine associée à la sous-unité 40S du ribosome. Je me suis donc demandé si son absence entraîne des défauts de biogenèse des sous-unités et d'assemblage du ribosome. Pour répondre à cette question, j'ai réalisé des profils de polysomes des souches FY, FYΔ(ASC1-U24), FYΔ(ASC1-U24) pASC1 et FYΔ(ASC1-U24) pACT-U24 (le snoARN snR24 est exprimé sur plasmide, sa

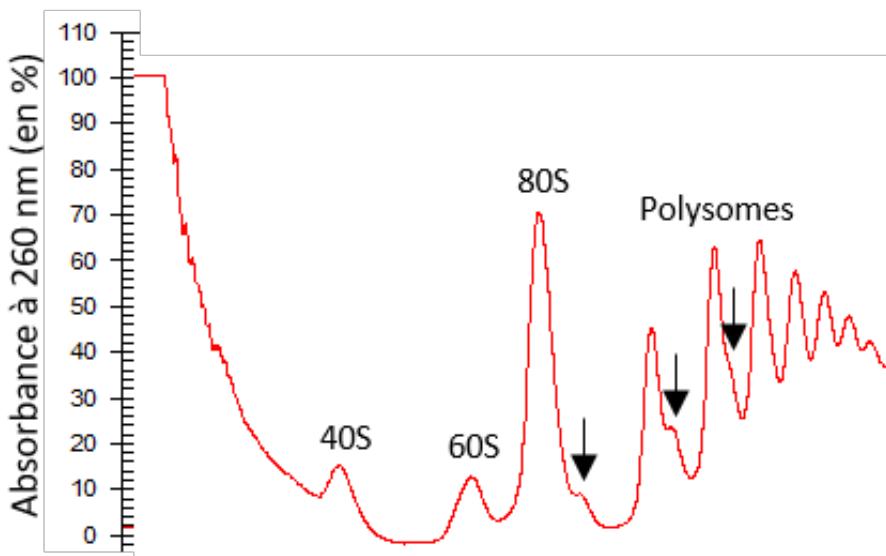


Figure 47 : Profil de polysomes de la souche FYΔ(ASC1-U24)

Des extraits de polysomes sont récoltés à partir de la souche FYΔ(ASC1-U24) après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucrose 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm. Les flèches noires indiquent des halfmères.

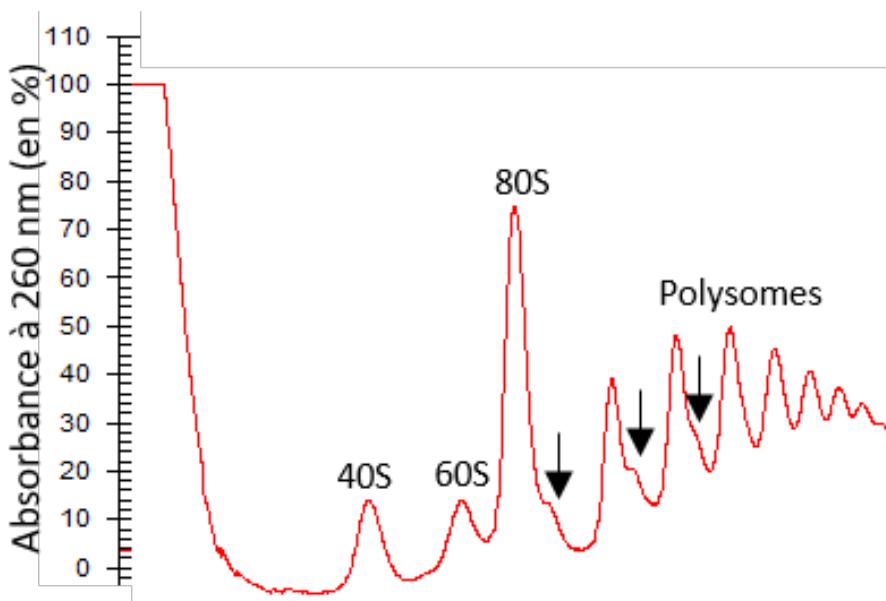


Figure 48 : Profil de polysomes de la souche FYΔ(ASC1-U24) pASC1

Des extraits de polysomes sont récoltés à partir de la souche FYΔ(ASC1-U24) pASC1 après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucrose 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm. Les flèches noires indiquent des halfmères.

séquence est insérée dans l'intron de la séquence codante de l'actine et leur expression est sous contrôle du promoteur ADH1). Les souches FYΔ(ASC1-U24) et FYΔ(ASC1-U24) pASC1 présentent une réduction du pic correspondant à la sous-unité 60S ainsi que des « halfmers » au niveau du pic de ribosomes 80S et des polysomes. Cela indique un défaut de biogénèse de la sous-unité 60S (Figures 46 à 48). Ce défaut est corrigé sur le profil de polysomes de la souche FYΔ(ASC1-U24) pACT-U24 (Figure 49). Ces résultats montrent que la délétion de snR24 entraîne un défaut de synthèse de la sous-unité 60S tandis que la délétion de la séquence codante du gène *ASC1* n'a aucun impact sur la biogénèse des sous-unités et l'assemblage des ribosomes et confirment ce qui avait déjà été montré³¹⁹.

La souche FYΔ(ASC1-U24) présente un taux de translecture plus faible que celui de la souche sauvage FY, ce qui signifie que la fidélité de la traduction est affectée par la délétion d'*ASC1*. Cette altération a-t-elle un impact sur la croissance cellulaire et sur la résistance à certains antibiotiques ? Dans un premier temps, j'ai effectué des courbes de croissance des souches FY1679 et FYΔ(ASC1-U24) pU24. Le temps de génération mesurée en phase exponentielle de croissance est de 4,4h pour la souche FY1679 et 4,1h pour la souche FYΔ(ASC1-U24) pU24 (Figure 50). La délétion du gène *ASC1* n'entraîne donc pas de différence significative du temps de génération.

J'ai ensuite testé la sensibilité de ces souches à différents antibiotiques : cycloheximide, paromomycine et hygromycine B. Les deux souches FYΔ(ASC1-U24) et FYΔ(ASC1-U24) pACT-U24 présentent une sensibilité accrue à chacun de ces antibiotiques par rapport à la souche sauvage FY. Le phénotype sauvage de sensibilité est restauré dans la souche FYΔ(ASC1-U24) pASC1. Ces résultats démontrent que la délétion du gène *ASC1* entraîne une augmentation de la sensibilité aux antibiotiques cycloheximide, paromomycine et hygromycine B et que la délétion de snR24 n'a aucun effet (Figure 51).

II – Discussion

Au cours de l'étude sur le rôle des 2'-O-méthylation au niveau du tunnel de sortie du peptide naissant, j'ai montré pour la première fois que l'absence de la protéine Asc1p entraîne une diminution du taux de translecture TMV. Cet effet a été confirmé pour les trois codons stop quel que soit le contexte testé, ce qui montre qu'Asc1p influence à la fois le taux de translecture programmée et le taux de translecture non programmée. La diminution du taux de translecture varie d'un facteur 2,6 à 3,1 pour le contexte TMV et d'un facteur 1,4 à 1,6 pour le contexte ADE1, avec le codon stop UAA qui est le plus affecté, puis UGA et enfin UAG, dans les deux contextes. Afin de mieux caractériser l'effet de la délétion du gène *ASC1* sur la translecture, nous pourrions réaliser de

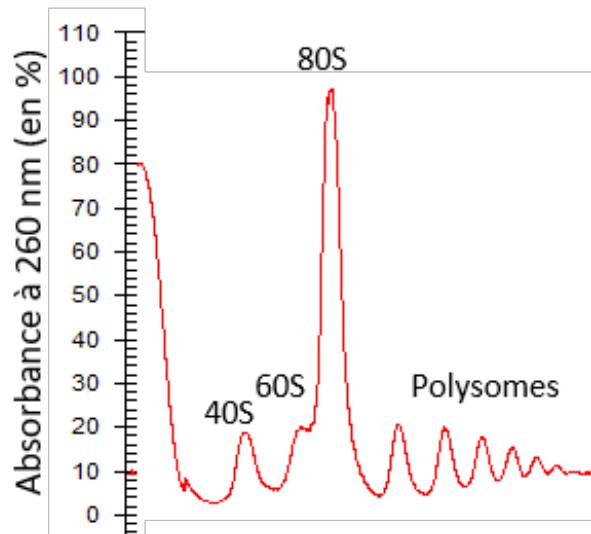


Figure 49 : profil de polysomes de la souche FYΔ(ASC1-U24) pACT-U24

Des extraits de polysomes sont récoltés à partir de la souche FYΔ(ASC1-U24) pACT-U24 après traitement à la cycloheximide. Ces extraits sont séparés sur gradients de sucre 7-47%. Les gradients sont passés dans un fractionneur qui mesure l'absorbance à 260 nm.

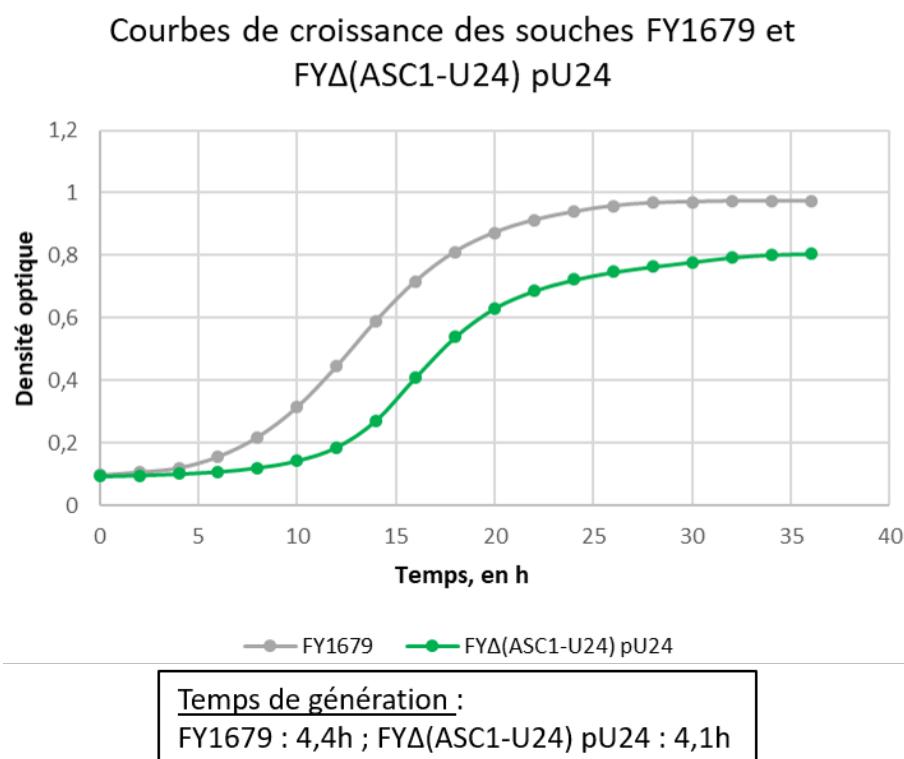


Figure 50 : Courbes de croissance des souches FY1679 et FYΔ(ASC1-U24) pU24

A partir des courbes de croissance des souches FY1679 et FYΔ(ASC1-U24) pU24, les temps de génération ont été calculés en phase exponentielle de croissance.

nouvelles mesures dans d'autres contextes, et également réaliser des constructions afin de déterminer si cet effet peut varier en fonction des éléments structuraux et/ou nucléotidiques de l'ARNm. Nous pourrions alors confirmer ou infirmer une hiérarchie dans les codons stops en fonction de l'impact de l'absence d'Asc1p sur leur taux de translecture (UAA plus affecté qu'UGA, lui-même plus affecté qu'UAG).

La découverte de cet effet sur la translecture causé par la perte d'Asc1p amène à se poser une question importante : comment Asc1p, liée à la sous-unité 40S du ribosome au niveau du tunnel de sortie de l'ARNm, a une action sur le complexe de terminaison au site A ? Jusqu'à présent, aucune interaction entre Asc1p et eRF1 (Sup45 chez la levure) ou eRF3 (Sup35 chez la levure) n'a été démontrée. Cependant, il a été montré qu'Asc1p interagit avec la protéine ribosomique Rps3, qui elle-même interagit avec Hbs1, homologue d'eRF3^{337,338}. Rps3 est localisée principalement au niveau du tunnel d'entrée de l'ARNm adjacent au site A et interagit avec le domaine globulaire N-terminal de Hbs1, qui est rattaché au corps de la GTPase par une région flexible³⁴⁸. Une extension C-terminale de Rps3 s'étend le long de la surface extérieure de la petite sous-unité du ribosome pour entrer en contact avec une des répétitions WD40 d'Asc1p³⁴⁹. Ainsi Asc1p interagit indirectement, par l'intermédiaire de Rps3, avec Hbs1 (Figure 52). Hbs1 étant un homologue du facteur de terminaison eRF3, il est possible que Rps3 interagisse également avec ce dernier, établissant un pont avec Asc1p. L'absence d'Asc1p pourrait altérer ces interactions et ainsi avoir un impact sur l'efficacité de la terminaison de la traduction. Une autre étude a suggéré une interaction entre Rps3 et le complexe de terminaison eRF1-eRF3 car l'hydrolyse du GTP porté par ce complexe nécessite des changements conformationnels qui impliquent Rps3³⁵⁰. Pour tester l'implication d'une interaction entre Asc1p et Rps3 dans la diminution de la translecture des codons stop, il pourrait être intéressant de toucher Rps3. Cependant, Rps3 étant essentielle, il n'est pas possible de la déléter. Un mutant de délétion de l'extension C-terminale qui interagit avec Asc1p pourrait alors être envisagé. L'existence d'une interaction entre Rps3 et eRF3 et/ou eRF1 pourrait également être recherchée.

De plus, il a été montré que le facteur d'initiation eIF3 interagit avec le facteur de terminaison eRF1 et que des mutants d'eIF3 présentent une diminution de la translecture des codons stop³⁵¹. Il est également connu qu'eIF3 interagit avec Asc1p³⁴⁰, et que les sous-unités a et g d'eIF3 interagissent avec Rps3^{352,353}. Ce réseau d'interaction ainsi que le phénotype de diminution de la translecture commun aux mutants d'eIF3 et au mutant Δ ASC1 suggèrent que les protéines Asc1p, Rps3 et eIF3 participent à la modulation de l'efficacité de la terminaison de la traduction. Des études supplémentaires doivent être menées afin de vérifier cette hypothèse, par exemple sur des mutants d'eIF3 dont l'interaction avec Asc1p est rompue.

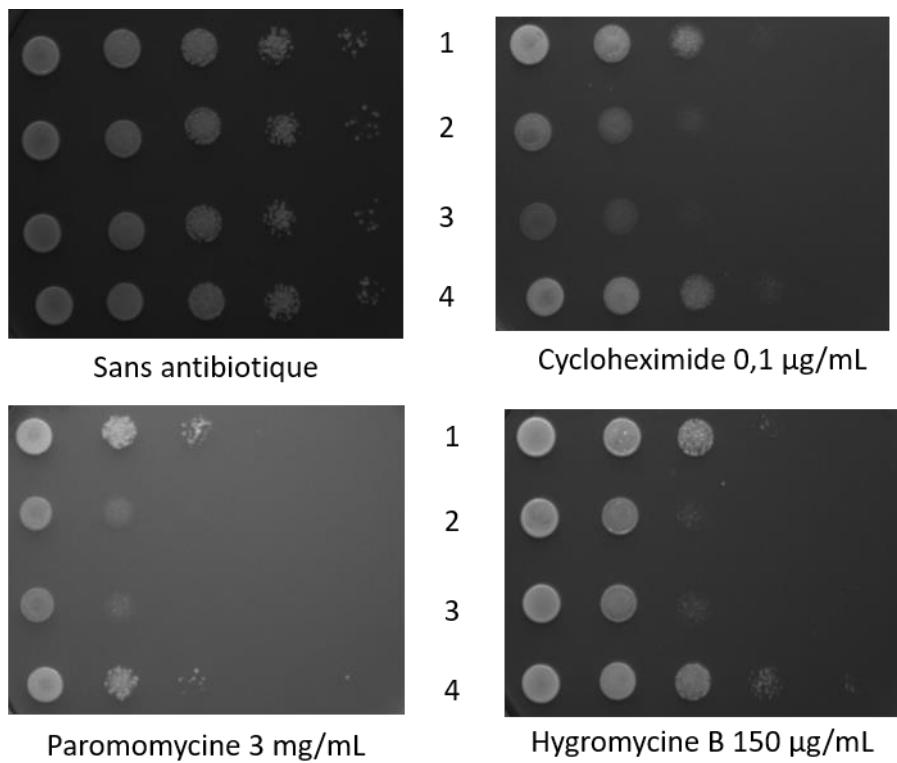


Figure 51 : Test en goutte de la sensibilité des souches FY1679, FY Δ (ASC1-U24), FY Δ (ASC1-U24) pACT-U24 et FY Δ (ASC1-U24) pASC1 à différents antibiotiques

Les souches FY1679 (1), FY Δ (ASC1-U24) (2), FY Δ (ASC1-U24) pACT-U24 (3) et FY Δ (ASC1-U24) pASC1 (4) ont été déposées à différentes concentrations sur un milieu solide avec différents antibiotiques. Ces différentes concentrations sont, de gauche à droite, 10⁵, 10⁴, 10³, 10² et 10¹.

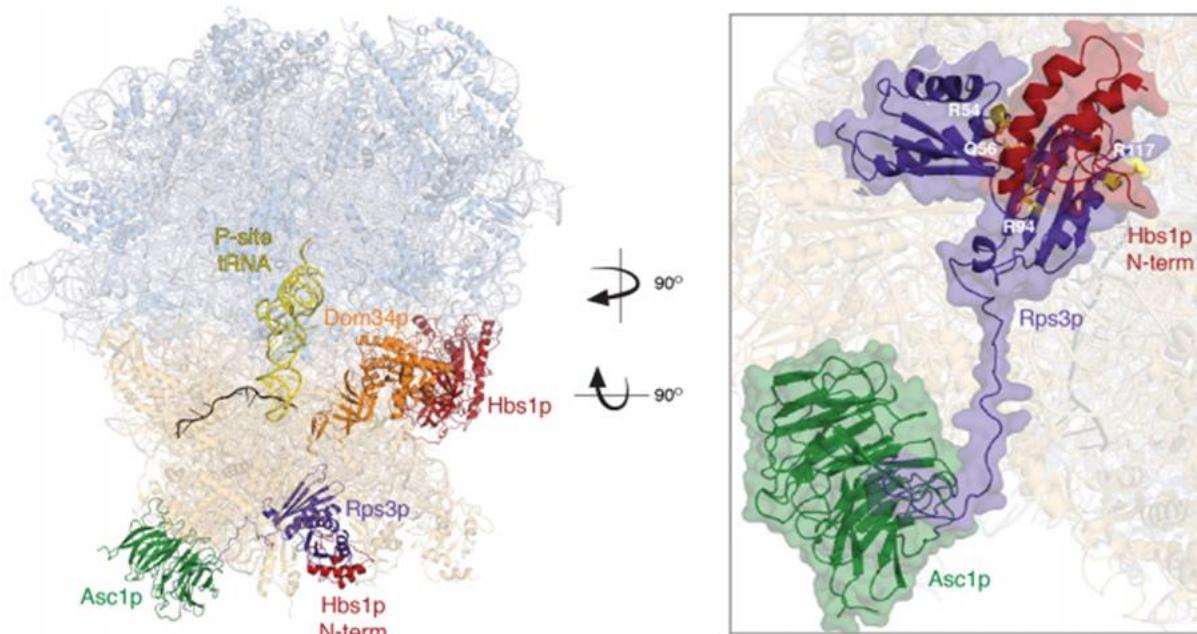


Figure 52 : Interactions entre Asc1p, Rps3 et Hbs1 chez *S. cerevisiae*

A gauche, structure cryo-EM du ribosome 80S de levure en complexe avec Dom34:Hbs1 (respectivement en orange et rouge), l'ARNt au site A (jaune) et un ARNm sans stop (noir) (d'après Hilal T. et. al, 2016). On distingue également Asc1p (en vert) et Rps3 (en violet). A droite, zoom sur les protéines Hbs1, Rps3 et Asc1p et leurs interactions. Les positions des mutations dans Rps3, à l'interface avec Hbs1, sont représentées en jaune. Adapté de Limoncelli et. al, 2017³³⁷.

J'ai ensuite montré que la délétion du gène *ASC1* n'avait aucun impact sur la croissance cellulaire et la biogenèse des ribosomes, contrairement à la délétion de snR24 présent dans l'intron d'*ASC1*, qui entraîne une diminution de la biogenèse de la sous-unité 60S. Mon travail confirme donc l'effet de la délétion de snR24.

Puis, j'ai réalisé un test de sensibilité aux antibiotiques cycloheximide, paromomycine et hygromycine B sur les souches FY, FYΔ(*ASC1*-U24), FYΔ(*ASC1*-U24) pACT-U24 et FYΔ(*ASC1*-U24) pASC1. J'ai montré que la délétion du gène *ASC1* entraîne une augmentation de la sensibilité à ces antibiotiques. *Asc1p* étant impliquée dans plusieurs processus traductionnels et cellulaires, l'impact de l'absence de cette protéine sur la sensibilité aux antibiotiques ciblant le ribosome n'est probablement pas dû à son rôle dans la translecture des codons stops mais plutôt dans d'autres mécanismes, comme la reconnaissance des ribosomes en pause. En effet, les antibiotiques sont utilisés à des doses sublétale et entraînent un blocage d'une partie des ribosomes, tandis qu'une autre partie est toujours disponible pour la traduction, ce qui assure la survie cellulaire. En cas d'absence d'*Asc1p*, si le mécanisme de reconnaissance des ribosomes en pause, et donc de récupération de ces ribosomes, est altéré, la traduction n'est plus maintenue et les cellules meurent.

Enfin, j'ai mesuré l'impact de la délétion d'*Asc1p* sur la fidélité de la traduction par des mesures de la translecture du codon stop et du décalage de cadre de lecture. Le taux de décalage de cadre de lecture s'est révélé insensible à la délétion du gène *ASC1* tandis que le taux de translecture diminue. Des hypothèses pour expliquer l'action d'*Asc1p* sur l'efficacité de la terminaison de la traduction font intervenir des interactions indirectes avec les facteurs de terminaison. Cependant, la terminaison de la traduction est un événement de compétition entre les facteurs de terminaison et des ARNt proche-cognats pour la reconnaissance du codon stop au site A. Il est donc possible qu'*Asc1p* agisse plutôt sur les ARNt plutôt que sur les facteurs de terminaison. Afin de tester cette hypothèse, je pourrais évaluer l'effet de la délétion du gène *ASC1* sur la misincorporation des acides aminés. Pour cela je pourrais utiliser un système rapporteur avec la séquence de la luciférase comportant un codon erroné dans le site actif. La protéine synthétisée à partir de cette séquence n'est pas active. Seule la misincorporation à ce codon de l'acide aminé de la séquence sauvage rétablit l'activité de la luciférase, rendant ainsi compte du taux de misincorporation. Si le taux de misincorporation n'est pas affecté, cela renforcera l'hypothèse d'une action de la protéine *Asc1p* sur les facteurs de terminaison. En revanche, si le taux de misincorporation varie en absence d'*Asc1p*, cela confortera l'hypothèse d'une action sur les ARNt.

Il a été montré qu'*Asc1p* intervient dans plusieurs mécanismes de pause des ribosomes³³¹⁻³³⁸. Une de ces études a révélé qu'*Asc1p* intervient avec *Hel2* et *Slh1* dans la reconnaissance

de ribosomes en pause et le recrutement du complexe RQC³³⁵. Elle a également montré l'implication de Cue3 et d'Ykr023w, une protéine dont la fonction n'est pas connue. Ykr023w possède un domaine en doigts de zinc, caractéristique d'une liaison aux acides nucléiques. Les quatre protéines Hel2, Slh1, Cue3 et Ykr023w interagissent physiquement entre elles et participent au recrutement du complexe RQC. Elles ont récemment été renommées Rqt1 (Hel2), Rqt2 (Slh1), Rqt3 (Cue3) et Rqt4 (Ykr023w) et il a été démontré que trois d'entre elles (Slh1, Cue3 et Ykr023w) appartiennent au même complexe RQT (RQC-Trigger, « trigger » pour déclenchement)³⁵⁴. Asc1p ne semble pas faire partie de ce complexe mais intervient tout de même dans le même processus, peut-être à travers son rôle de protéine plateforme. Le lien entre Asc1p et les protéines Rqt1-4 reste à déterminer, et la fonction d'Ykr023w (Rqt4) est également à caractériser.

Asc1p est connue pour être impliquée dans de nombreux processus traductionnels et cellulaires, mais un effet sur la translecture des codons stop n'avait jamais été décrit. Ici je mets en évidence pour la première fois la diminution de la translecture de codons stop programmés et non programmés en absence d'Asc1p. Tout comme pour une grande partie des processus dans lesquels elle est impliquée, son rôle exact reste à déterminer. Du fait de sa caractéristique de protéine plateforme, il est très probable qu'elle participe par le recrutement, et ainsi le rapprochement, de protéines interagissant entre elles. Son absence entraînerait une perte de ces interactions et serait à l'origine des phénotypes observés.

Conclusions & Perspectives

Conclusions et Perspectives

La traduction est une étape de l'expression des gènes fortement régulée, dont un niveau de régulation est le ribosome lui-même. La notion de « ribosome spécialisé », qui suggère que l'hétérogénéité de la composition des ribosomes participe à la régulation de la traduction de groupes spécifiques d'ARNm, reste toutefois à mieux caractériser. L'implication de certaines protéines ribosomiques et modifications chimiques de ces protéines dans la traduction de groupes d'ARNm a déjà été montrée. En revanche, le rôle des modifications chimiques des ARNr dans la fidélité et la régulation de la traduction n'a pas encore été révélé. Jusqu'à présent, il est connu que des délétions uniques de ces modifications n'a pas d'impact sur la traduction et qu'ainsi leur effet est combinatoire. Ces modifications sont également conservées, ce qui suggère une importance fonctionnelle.

Au cours de ma thèse, je me suis intéressée au rôle de la modification 2'-O-méthylation dans la fidélité de la traduction. Tout d'abord, j'ai étudié l'hétérogénéité et la plasticité des méthylations des ARNr au sein d'une population de ribosomes dans des cellules humaines. Ensuite j'ai analysé et essayé de caractériser les variations globales de la traduction en réponse à une diminution des 2'-O-méthylation, puis j'ai étudié l'implication de ces méthylations dans des événements traductionnels plus spécifiques, dont l'initiation IRES-dépendante et le recodage, toujours dans des cellules humaines. Enfin j'ai réalisé une étude ciblée sur un groupe de méthylations situées au niveau du tunnel de sortie du peptide, cette fois chez la levure.

Les 2'-O-méthylation des ARNr sont effectuées par des complexes snoRNPs composés d'un ARN guide, d'une méthyltransférase, la fibrillarine, et de protéines accessoires. Jusqu'à présent, le rôle des 2'-O-méthylation avait été étudié principalement de manière spécifique sur un groupe ciblé de méthylations, par la délétion des snoARNs responsables de ces méthylations. Une baisse globale de la méthylation des ARNr avait également été obtenue par l'extinction de la protéine L13a qui interagit avec la fibrillarine mais l'étude s'était limitée à l'initiation IRES-dépendante. Dans une étude précédente de nos collaborateurs de l'équipe de JJ Diaz, il avait été montré qu'une altération du taux de la fibrillarine entraînait une variation du taux de 2'-O-méthylation des ARNr. Nous avons donc choisi d'affecter les 2'-O-méthylation par l'extinction de la fibrillarine. La fibrillarine étant une protéine essentielle, nous avons réalisé une extinction par un shRNA dirigé contre l'ARNm de la fibrillarine et dont l'expression est inductible. J'ai mis au point les conditions optimales d'induction de l'expression du shRNA et analysé les variations des taux des 2'-O-méthylation causées par la baisse du taux de fibrillarine. Pour cela nous avons utilisé une technique nouvellement mise au point,

le RiboMethSequencing, qui permet de quantifier le taux de 2'-O-méthylation à chaque position des ARNr au sein d'une population de ribosomes. Nous avons ainsi validé la baisse globale de la méthylation des ARNr en réponse à l'extinction de la fibrillarine. Cependant, de manière surprenante, cette baisse varie de façon site-spécifique. Nous avons déterminé que les sites les plus touchés sont ceux qui présentent un taux de méthylation initial inférieur à 80% et avons avancé deux hypothèses pour l'expliquer : 1) à ces sites, l'ajout d'un groupement méthyle est moins efficace et est donc plus sensible à la baisse du taux de fibrillarine ; 2) les variations au niveau des autres positions (taux initial supérieur à 80%) sont faiblement visibles car seules les sous-unités qui présentent des méthylations à ces positions se forment correctement et sont donc purifiées. Cette étude nous a également permis de montrer que parmi les sites les plus touchés, certains étaient localisés dans des régions fonctionnelles du ribosome. En revanche, les méthylations du centre de décodage et du centre peptidyl-transférase sont peu affectés, ce qui suggère un rôle important de ces modifications. Cette étude nous a donc permis de mieux caractériser l'altération du taux de méthylation des ARNr en réponse à l'extinction de la fibrillarine. Elle soulève également une question importante : pourra-t-on affecter plus fortement le taux de 2'-O-méthylation ? Deux axes d'analyses pourraient être envisagés. Le premier se focaliserait sur l'optimisation de la baisse globale de la méthylation. Le problème principal de cette approche réside dans le fait que la fibrillarine est une protéine essentielle impliquée dans le clivage des ARNr. La piste la plus prometteuse aujourd'hui serait de muter la protéine fibrillarine afin qu'elle perde uniquement sa fonction de méthyltransférase, préservant ainsi la fonction de clivage des ARNr. Néanmoins, la méthylation des ARNr est essentielle pour un clivage correct chez les mammifères, il n'est donc pas à exclure qu'un tel mutant présente tout de même des défauts de biogénèse des ribosomes. La déméthylation des ARNr semblerait donc plus adaptée mais nous n'avons actuellement aucune technique à notre disposition pour l'envisager. Le second axe d'étude résiderait dans la perte de méthylations localisées dans des régions fonctionnelles définies, comme le DC et le PTC. En effet, ce sont les deux régions les plus importantes fonctionnellement et les ribosomes matures hypométhylés présentent presque tous les méthylations de ces régions. Nous pourrions envisager de déléter dans le génome les snoARNs correspondants à ces méthylations par la technique CRISPR-Cas9. Il n'est toutefois pas certains que les ARNr qui en résultent soient intégrés dans des ribosomes fonctionnels. Dans ce cas, nous pourrions conclure sur la nécessité des 2'-O-méthylation dans le DC et le PTC. Si des ribosomes fonctionnels sont synthétisés, nous pourrions alors étudier les impacts de la perte des méthylations dans le DC et le PTC sur la traduction et sa fidélité.

L'étude des effets d'une baisse globale de la méthylation des ARNr n'avait jusqu'à présent été réalisée que sur des cibles spécifiques. Au cours de ma thèse, j'ai pu mettre en évidence l'impact

de la baisse globale des 2'-O-méthylations due à la baisse du taux de fibrillarine sur la traduction d'un groupe d'ARNm et sur l'initiation de la traduction à certains sites non annotés. L'analyse du groupe d'ARNm n'a pour le moment pas permis de dégager d'élément fonctionnel ou moléculaire commun qui permettrait d'établir un mécanisme moléculaire de l'implication des méthylations des ARNr dans la traduction de ce groupe spécifique de transcrits. Concernant l'étude ciblée sur l'initiation de la traduction, l'analyse des sites d'initiation affectés par la baisse de la méthylation des ARNr a été abordée mais nécessite d'être approfondi. Tout d'abord, la liste des sites affectés doit être vérifiée par l'utilisation de systèmes double rapporteurs, puis une analyse sera effectuée dans le but de déterminer des éléments fonctionnels ou moléculaires communs à ces sites. Cette analyse tiendra compte du fait qu'à l'exception d'un seul, tous les sites d'initiation affectés par la baisse des 2'-O-méthylations ne sont pas annotés comme des sites d'initiation de l'ORF principale, ce qui suggère plutôt un rôle de régulation des méthylations.

Ces analyses globales sur la traduction et l'initiation de la traduction ont pour la première fois montré un rôle spécifique des 2'-O-méthylations dans la traduction d'un groupe d'ARNm et dans l'initiation à certains sites. L'analyse de ces résultats prometteurs est à approfondir afin de comprendre les mécanismes impliqués dans la régulation de la traduction par les 2'-O-méthylations.

L'impact de la baisse des méthylations des ARNr sur la fidélité a également été étudié à travers les événements de recodage que sont la translecture du codon stop et le décalage de cadre de lecture, et l'initiation IRES-dépendante. J'ai ainsi montré une augmentation des taux de translecture et de décalage de cadre de lecture, ainsi qu'une baisse de l'initiation au niveau d'IRES.

Ces résultats, bien que prometteurs, soulèvent quelques limites techniques et biologiques. Tout d'abord, la technique de ribosome profiling présente une très bonne sensibilité mais pourrait toutefois ne pas détecter des variations traductionnelles très faibles. Ensuite, la baisse des 2'-O-méthylations n'étant pas complète, son impact pourrait être limité. Enfin, les effets d'une baisse de la méthylation des ARNr sur la traduction pourraient être limités en conditions normales de culture cellulaire. Le premier point d'amélioration concerne la baisse partielle des méthylations et a été développé dans le paragraphe précédent. Le second point concerne les conditions de culture des cellules. L'impact de la diminution des 2'-O-méthylations sur la traduction pourrait être plus important en condition de stress, comme le stress oxydatif, l'hypoxie ou encore l'infection par un agent pathogène. En terme pratique, le stress oxydatif semble être le stress le moins compliqué à induire, par l'ajout d'agents oxydatifs dans le milieu de culture, tandis que les conditions d'hypoxie et d'infection des cellules nécessiteraient des équipements plus spécifiques et plus coûteux.

Mon équipe, toujours en collaboration avec l'équipe de JJ Diaz, prévoit d'étudier l'influence des 2'-O-méthylations dans la cinétique de l'elongation de la traduction en collaboration avec l'équipe de N. Westbrook et K. Perronet à l'Institut d'Optique. Pour cela, un outil d'analyse

dynamique de la traduction en molécule unique sera utilisé. Cet outil repose sur de la traduction *in vitro* et permet d'obtenir pour chaque ribosome des informations sur la cinétique de la traduction. Des ribosomes hypométhylés ou qui présentent une perte de méthylation ciblée seront purifiés et utilisés pour cette expérience en molécule unique afin de mesurer la vitesse de l'elongation de la traduction de ces ribosomes.

Aujourd'hui, l'étude structurale du ribosome par cryo-EM est suffisamment précise pour permettre de visualiser les modifications chimiques des ARNs ribosomiques. On pourrait alors envisager d'étudier le rôle des 2'-O-méthylation avec une vision plus structurale et utiliser cette technique pour observer l'impact d'une baisse globale ou d'une perte ciblée de méthylations dans une région fonctionnelle sur la structure du ribosome.

Les 2'-O-méthylation les plus touchées ont été localisées sur le ribosome de levure et groupées en régions fonctionnelles. Je me suis ainsi intéressée à un groupe de méthylations conservées entre la levure et l'Homme (quatre sont présentes chez la levure sur les six présentes chez l'Homme) et situées au niveau du tunnel de sortie du peptide. Afin de réaliser une étude ciblée de l'impact de la perte de ces méthylations, j'ai choisi de changer d'organisme et d'effectuer cette étude chez la levure *S. cerevisiae*, dans laquelle il est plus facile de faire des délétions au sein du génome. J'ai donc délétré dans un premier temps trois des quatre snoARNs responsables de ces quatre méthylations suivant deux combinaisons différentes et étudié l'impact sur la croissance cellulaire, la biogenèse des ribosomes, la fidélité de la traduction et la sensibilité aux antibiotiques. Cependant je n'ai mesuré aucun effet de la perte de trois des quatre méthylations autour du tunnel de sortie du peptide sur ces processus. Je dois donc à présent effectuer les mêmes tests dans la souche délétrée pour les quatre snoARNs. De plus, les tests que j'ai effectués ne permettent pas une mesure directe du repliement des protéines. Il serait important de développer des moyens de mesurer le repliement, à l'aide d'un système rapporteur comprenant la GFP par exemple.

Au cours de cette étude sur la perte des méthylations autour du tunnel de sortie du peptide, j'ai révélé de façon surprenante et inédite un effet de la délétion du gène *ASC1* sur la translecture des codons stop. La protéine Asc1p est une protéine plateforme associée à la sous-unité 40S du ribosome et impliquée dans de nombreux processus traductionnels et cellulaires. Son implication a été démontrée dans tous ces processus mais les mécanismes précis par lesquels elle intervient reste aujourd'hui encore très largement incompris. En ce qui concerne l'effet sur la translecture que j'ai montré, nous avons avancé plusieurs hypothèses qui font intervenir des interactions entre la protéine Asc1p et les protéines Rps3 et eIF3 qui interagiraient à leur tour avec les facteurs de terminaison eRF1 et eRF3. Pour tester ces hypothèses, il serait intéressant de construire des mutants

des protéines Asc1p, Rps3 et/ou eIF3 qui auraient perdu la capacité d'interaction avec une ou plusieurs autres de ces protéines et de déterminer si ces mutants présentent le même phénotype de diminution de la translecture des codons stop que la souche délétée pour le gène *ASC1*. La caractérisation du mécanisme par lequel Asc1p régule l'efficacité de la terminaison de la traduction permettrait à la fois d'améliorer la compréhension de cette étape ainsi que le rôle d'Asc1p dans ce processus et peut-être dans d'autres processus traductionnels.

Mes travaux sur les 2'-O-méthylation des ARNr ont permis de mieux caractériser le rôle de ces modifications dans la régulation et la fidélité de la traduction. Tout d'abord, la baisse de ces méthylations en réponse à une diminution du taux de fibrillarne a été caractérisée. Puis j'ai démontré pour la première fois que la baisse des 2'-O-méthylation avait un impact sur la traduction d'un groupe spécifique d'ARNm et de sites d'initiation. Ces résultats prometteurs sont aujourd'hui à approfondir dans le but d'établir un mécanisme moléculaire du rôle des méthylations des ARNr dans la traduction. J'ai également montré que la fidélité de la traduction et l'initiation IRES-dépendante étaient affectés par la baisse des méthylations. Enfin, j'ai révélé chez la levure l'implication de la protéine Asc1p dans la translecture des codons stop. L'étude du mécanisme moléculaire mis en jeu pourrait permettre de mieux comprendre la modulation de l'efficacité de la terminaison de la traduction.

Matériels & Méthodes

Matériels et Méthodes

I - Les lignées cellulaires et les souches de levure

Lignée HeLa shFBL

La lignée HeLa shFBL a été construite par intégration d'une construction comprenant le shRNA dirigé contre l'ARNm de la fibrillarine, de façon stable dans le génome de cellules HeLa, par un système lentiviral. Cette construction est composé d'un marqueur de sélection de mammifère (résistance à la puromycine) et du transactivateur du système d'induction (reverse tetracycline transactivator 3 noté rtTA3) placés sous contrôle d'un promoteur constitutif ; et du shFBL en fusion transcriptionnelle avec la turboRFP (ou tRFP), sous contrôle du promoteur TRE (Tetracycline Response Element) inducible en présence de doxycycline. Le transactivateur TA3 se fixe au promoteur TRE en présence de doxycycline, qui est un analogue plus stable de la tétracycline, et active la transcription. C'est une version modifiée du transactivateur sauvage qui a subi différentes mutations, lui permettant de se fixer au promoteur TRE en présence de doxycycline plutôt qu'en son absence (chez le sauvage), et d'activer la transcription et non de la réprimer. La tRFP et le shFBL sont co-exprimés à partir du même promoteur TRE, ce qui permet de suivre l'expression du shFBL par la mesure de la fluorescence de tRFP.

Souche FY1679

Le génotype de la souche FY1679 est : *MATa ura3-52 trp1Δ63 leu2Δ1 his3Δ200*.

Souche FYΔ(39B-58-24) pASC1

Cette souche est dérivée de la souche FY1679 et son génotype est : *MATa ura3-52 trp1Δ63 leu2Δ1 his3Δ200 snR39B :: HIS3 ΔsnR58 ΔASC1-snR24*. La délétion de la séquence codante du gène *ASC1* n'étant pas mon objectif, elle est apportée dans la souche par le vecteur centrométique pRS416-*ASC1ORF* (URA3) noté pASC1.

Souche FYΔ(39B-58-18) pEFB1

Cette souche est dérivée de la souche FY1679 et son génotype est : *MATa ura3-52 trp1Δ63 leu2Δ1 his3Δ200 snR39B :: HIS3 ΔsnR58 ΔEFB1-snR18*. La délétion de la séquence codante du gène *EFB1* étant létale, elle est apportée dans la souche par le vecteur centrométique pFL39-EFB1ORF (TRP1) noté pEFB1.

Souche FYΔ(ASC1-U24)

Cette souche est dérivée de la souche FY1679 et son génotype est : *MATa ura3-52 trp1Δ63 leu2Δ1 his3Δ200 ASC1-snR24 :: LEU2*. *snR24* est aussi noté U24.

Souche FYΔ(ASC1-U24) pACT-U24

Cette souche est dérivée de la souche FY1679 et son génotype est : *MATa ura3-52 trp1Δ63 leu2Δ1 his3Δ200 ASC1-snR24 :: LEU2*. Afin de n'étudier que la délétion de la séquence codante du gène *ASC1* et de s'affranchir des effets de la délétion de *snR24*, celui-ci est apporté dans la souche par le vecteur centromérique pRS413-ACT-U24 (URA3) noté pACT-U24 et à partir duquel *snR24* est exprimé au sein d'une séquence intronique de la séquence codante de l'actine ACT.

Souche FYΔ(ASC1-U24) pASC1

Cette souche est dérivée de la souche FY1679 et son génotype est : *MATa ura3-52 trp1Δ63 leu2Δ1 his3Δ200 ASC1-snR24 :: LEU2*. La séquence codante du gène *ASC1* est apportée dans la souche par le vecteur centrométique pRS416-ASC1ORF (URA3) noté pASC1.

II – Les plasmides

Plasmide pAC99 (mammifères)

Ce plasmide comporte la séquence codante *LacZ* de la β -galactosidase et la séquence codante *Luc* de la Firefly-luciférase, exprimées sous la forme d'une protéine de fusion. Entre les séquences de *LacZ* et de *Luc* se trouve le site de clonage *MscI*, auquel sont insérées les séquences à tester, de translecture ou de décalage de cadre de lecture. Les caractéristiques de la cible insérée déterminent le nom de chaque plasmide.

Plasmide pIRES (mammifères)

Ce plasmide est un vecteur bicistronique dont le premier cistron, la séquence codante de la Renilla-luciférase, est sous le contrôle du promoteur CMV, et le second cistron, la séquence codante de la Firefly-luciférase, est sous le contrôle d'un IRES (IGF-1R, c-Myc ou FGF1). Ces plasmides m'ont été donnés par Anne-Catherine Prats pour les plasmides pIRES c-Myc et FGF1¹⁰¹ et par Jean-Jacques Diaz pour le plasmide pIRES IGF-1R³⁵⁵.

Plasmide rapporteur pAC PGK (levure)

Ce plasmide est un vecteur centromérique qui comporte la séquence codante *LacZ* de la β -galactosidase et la séquence codante *Luc* de la Firefly-luciférase, exprimées sous la forme d'une

protéine de fusion. Cette construction est sous le contrôle du promoteur constitutif PGK1 (promoteur de la phosphoglycérate kinase). Entre les séquences de *LacZ* et de *Luc* se trouve le site de clonage Mscl, auquel sont insérées les séquences à tester, de translecture ou de décalage de cadre de lecture. Les caractéristiques de la cible insérée déterminent le nom de chaque plasmide. Le marqueur de sélection du plasmide est en général LEU2 mais peut être modifié.

Plasmide pDS4 (levure)

Ce plasmide porte le promoteur, le terminateur et la séquence codante de KILEU2, flanqués de part et d'autre par des séquences de reconnaissance par l'endonucléase Scel et des séquences répétées AB. Ce plasmide est AmpR.

Plasmide pUDC073 (levure)

Ce plasmide porte la séquence codante de l'endonucléase Scel, sous contrôle d'un promoteur inductible en milieu galactose. Il est également URA+ et AmpR.

III – Les milieux

Milieu Arg 0,2 :

Ce milieu correspond à un milieu YNB classique dans lequel le sulfate d'ammonium, source classique d'azote, a été remplacé par de l'arginine.

Composition : Yeast Nitrogen Base (YNB) 1,7g/L ; Glucose 20g/L ; Arginine 2g/L ; (Agar 20g/L).

IV – Les protocoles

Profils de polysomes

Les profils de polysomes sont effectués à partir d'extraits de polysomes réalisés comme présenté dans le protocole de ribosome profiling. Puis les échantillons sont déposés sur des gradients de sucre 7-47%. Ces gradients ont la composition suivante : 31% de glucose ; 50mM de Tris-acétate pH 7,6 ; 50mM de NH₄Cl ; 12mM de MgCl₂ ; 1mM de DTT et ont subi 3 cycles de congélation/décongélation afin de former le gradient. Les gradients sont centrifugés 3h à 39000 rpm à 4°C, puis passés dans un fractionneur qui mesure la densité optique à 260nm. On visualise ainsi les

sous-unités 40S et 60S, les monosomes 80S et les polysomes (séparés dans le gradient suivant leurs poids moléculaires.

Ribosome profiling

La préparation de chaque échantillon a nécessité 150 millions de cellules à hauteur de 4,5-5 millions de cellules par boîte de culture.

1) Extraction des polysomes :

Dans chaque boîte de cellules, ajouter 100 µg/mL de cycloheximide à 50 mg/mL pour bloquer les ribosomes sur les ARNm. Laisser 10 min à 37°C. Décoller les cellules à la trypsine + cycloheximide 100 µg/mL, à froid. Centrifuger les cellules pendant 10 min à 300g et 4°C. Resuspendre le culot dans du tampon de lyse + cycloheximide 100 µg/mL (1 mL pour 10 boîtes de culture).

Tampon de lyse : 10 mM Tris-HCl pH7,5 ou 10 mM Tris-PO₄ pH7,5 ; 100 mM KCl ; 10 mM Mg²⁺ acétate ; 1% Triton X100 ; 2 mM DTT

Transférer dans un tube eppendorf de 2 mL. Ajouter des billes de verre et vortexer 5 min à 4 °C. Centrifuger 10 min à 3000 rpm à 4 °C. Transférer le surnageant dans un tube eppendorf et faire une mesure de la concentration en ARN (OD à 260nm). Congeler rapidement dans l'azote liquide et congeler à -80°C.

2) Digestion RNase I :

Traiter les extraits de polysomes à la RNase1 AMBION (qui ne digère pas les ARNs simples brins) pendant 1h à 1h30min à 25°C à raison de 15 unités de RNase1/OD_{260nm}.

3) Purification des monosomes sur coussin de saccharose 24% :

Coussin saccharose 24%, 50 mM Tris-acétate pH7,6 ; 50 mM NH4Cl ; 12 mM MgCl₂ : soit 17 mL saccharose 70% ; 2,5 mL Tris-Ac 1M pH7,6 ; 2,5 mL NH4Cl 1M ; 0,6 mL MgCl₂ 1M ; 0,05 mL DTT 1M ; eau qsp 50 mL.

Déposer environ 1 mL d'extraits digérés à la RNase I sur 3 mL de coussin en tube de polycarbonate 13X56 mm pour centrifuger 2h15min à 100000 rpm avec le rotor TLa11 dans l'ultracentrifugeuse OptimaMAX-XP. Rincer 2 fois les culots par 1 mL de tampon d'extraction de polysomes (tampon de

lyse) 1X avec de la cycloheximide à 50 µg/mL, puis reprendre avec 750 µL du même tampon et soit congeler les monosomes à -20°C, soit extraire les ARNs par le phénol.

4) Purification des ARNs au phénol et précipitation :

Cette étape permet de retirer les protéines des ribosomes au phénol/chloroforme et on purifie l'ARN en le précipitant par le traitement à l'acétate et à l'alcool.

Si besoin, décongeler les échantillons.

Ajouter un volume de phénol acide (aquaphenol non tamponné) sous la hotte avec des gants. Agiter fortement 1h à 65°C sous la hotte à l'aide d'un vortex. Centrifuger 10 min sous la hotte. Prélever la phase aqueuse (phase supérieure).

Ajouter le même volume de chloroforme sous la hotte. Vortexer 1 min. Centrifuger 10 min sous la hotte. Prélever la phase aqueuse (phase supérieure).

Ajouter 0,1 volume d'acétate de sodium pH5,2 3M. Mélanger et ajouter 3 volumes d'EtOH 100%. Mélanger et conserver une nuit à -20°C pour la précipitation.

Pour la précipitation, centrifuger les tubes 30 min à 4°C 10 000g. Retirer le surnageant. Recentrifuger rapidement pour éliminer les traces d'alcool. Laisser sécher le culot tubes ouverts sous la hotte. Reprendre le culot dans de l'eau + Superase-IN AMBION (1 U/µL final à partir de 20 U/µL). L'échantillon peut être conservé à -20°C.

5) Sélection des fragments de 28-30 nucléotides sur gel :

Migrer jusqu'à 200 µg de matériel par gel dénaturant d'acrylamide 15% avec 7M d'urée fin (20x20cm, espaceur de 0,75mm) ou entre 200 et 500 µg de matériel par gel épais (20x20cm, espaceur de 1,5mm).

Préparation des gels 15% acrylamide, 7M urée :

Pour 100 mL de gel : 42 g urée ; 2 mL TAE 50X ; 42,5 mL acrylamide 19:1 40% ; 66,7 µL TEMED ; 500 µL APS 10% ; env. 20 mL H₂O. Compter 35 mL par gel fin et 55 mL par gel épais.

Mettre les gels à pré-migrer pour les préchauffer à 45-55°C pendant au moins 30 min, avec le circuit d'eau qui maintient la température à 50°C.

Charger le gel en répartissant l'échantillon avec du bleu de charge ARN tout le long et de part et d'autre les oligonucléotides marqueurs de 28nt et 34nt. Faire migrer 30 min à 120V puis 5 à 6h à 200V.

Après migration, incuber les gels 30 min dans une solution de SYBR-Gold diluée au 10000ème (c'est un agent intercalant très sensible, idéal pour les petites quantités) : pour 100 mL de solution, diluer 10 µL de SYBR-Gold dans 100 mL de TAE 1X.

Placer le gel sous la lampe UV 300 nm et découper la bande située à 28 nt environ, marquée par l'oligo. Cette taille correspond aux fragments d'ARNm en cours de traduction protégés par les ribosomes et se trouvent bordés de part et d'autre par des bandes très intenses correspondant à de l'ARNr fragmenté également par la RNase I.

Placer les bandes de gel dans un eppendorf et les congeler rapidement à -20°C.

6) Pulvérisation du gel, extraction et précipitation :

Percer le fond d'un tube de 0,5 mL avec une aiguille de seringue 20G et le glisser dans un tube de 2 mL. Charger le petit tube avec le gel congelé et centrifuger 1 min à vitesse maximale. Le gel est pulvérisé à travers le trou et se trouve dans le grand tube (si il reste du gel dans le petit tube, le retourner et récupérer ce qui reste).

Resuspendre les débris de gel dans du tampon d'extraction (assez pour que du liquide circule en inversant le tube sur une roue). Placer la nuit sur une roue à 4°C.

Tampon d'extraction :

pour 50 mL : 300 mM NaOAc pH 5,5 ; 5 mL NaOAc 3M ; 1 mM EDTA ; 100 µL EDTA 0,5M ; 45 mL H₂O

Pour l'extraction de l'ARN du gel, transférer le liquide des débris de gel sur une colonne Spin-X cellulose acétate filter. Centrifuger 1 min vitesse maximale. Ajouter au volume récupéré 1 µL de glycogène 20mg/mL par tube puis 2 à 2,5 volume d'EtOH 100%. Congeler à -20°C pour la précipitation.

Pour la précipitation, centrifuger les ARNs extraits du gel à 4°C pendant 30 min, vitesse maximale. Retirer le surnageant à la pompe et bien laisser sécher le culot sous la hotte. Reprendre dans 20 µL d'eau + Superase-IN à 1U/µL final. Congeler à -20°C ou procéder à l'étape suivante.

7) Déplétion des ARNr contaminants

Parmi les ARNs récupérés, on retrouve une quantité importante d'ARNr issue de la digestion des ribosomes par la RNase1 (fragments d'ARNr accessibles). On cherche à s'en débarrasser le plus possible. Pour cela on utilise le kit Ribo-Zero Magnetic Gold Human/Mouse/Rat de chez Illumina, en suivant le protocole indiqué. On peut ensuite mesurer la quantité d'ARNs restants, avant de procéder à la création de la banque (réalisée par la plateforme du CNRS de Gif-sur-Yvette).

GTI-Seq (Global Translation Initiation Sequencing)

La préparation de chaque échantillon a nécessité 125 millions de cellules à hauteur de 4,5-5 millions de cellules par boîte de culture.

Le protocole de GTI-Seq est identique à celui du ribosome profiling classique, hormis pour l'étape d'extraction des polysomes. Dans le cas du GTI-Seq, les cellules ne sont pas traitées avec de la cycloheximide mais avec 2 µg/mL d'harringtonine à partir d'une solution mère à 2 mg/mL pendant 10 min à 37°C.

QTI-Seq (Quantitative Translation Initiation Sequencing)

La préparation de l'échantillon a nécessité 90 millions de cellules à hauteur de 4,5-5 millions de cellules par boîte de culture.

Le protocole de QTI-Seq diffère de celui du ribosome profiling uniquement pour l'étape d'extraction des polysomes. La molécule utilisée pour bloquer les ribosomes n'est pas la cycloheximide mais la lactimidomycine, qui bloque spécifiquement les ribosomes en initiation. Cette technique, contrairement au GTI-Seq, est quantitative.

Extraction de polysomes en QTI-Seq :

Par lot de 3 boîtes, aspirer le milieu et rincer chaque boîte avec 2 fois 5 mL de PBS. Placer les boîtes sur de la glace (à partir de cette étape, les échantillons doivent rester le plus possible au froid). Dans une boîte, ajouter 400 µL de tampon de lyse avec 10 µM de lactimidomycine. Détacher les cellules en grattant la boîte avec un policeman. Récupérer à la pipette le volume le plus important possible et le transférer dans une seconde boîte. De nouveau détacher les cellules à l'aide d'un policeman et transférer le liquide récupéré dans la troisième boîte. Recommencer et transférer le

volume final dans un tube eppendorf de 2 mL, dans la glace. Le volume ne doit pas dépasser 800 μ L afin d'avoir une concentration de lactimidomycine $> 5 \mu$ M.

Ajouter des billes de verre. Vortexer 5 min à 4°C. Centrifuger 10 min à 3000 rpm à 4 °C. Transférer le surnageant dans un tube eppendorf et faire une mesure de la concentration en ARN (OD à 260nm). Congeler rapidement les deux tubes dans l'azote liquide et congeler à -80°C.

Utilisation du système d'excision du marqueur

J'utilise le plasmide pDS4 qui porte le marqueur de sélection Leucine KILEU2 flanqués des séquences de reconnaissance de l'endonucléase Scel et de séquences homologues pour amplifier la cassette de délétion puis j'effectue la délétion du snoARN dans le génome par insertion de la cassette. Je sélectionne les clones qui portent la cassette KILEU2 insérée au locus du snoARN que je souhaite déléter sur milieu sans leucine et je vérifie l'insertion par PCR. Je transforme un clone sélectionné par le plasmide pUDC073 (URA+) qui porte l'endonucléase et j'induis son expression en plaçant le clone en milieu avec galactose. Au bout de quelques jours, je mets plusieurs sous-clones en culture dans du milieu avec du glucose et sans pression de sélection sur le plasmide pUDC073 afin de le perdre. Puis j'étale une centaine de clones sur une boîte 5-FOA afin de sélectionner uniquement ceux qui ont perdu le plasmide. Je sous-clone ensuite une quarantaine de clones, en parallèle, sur milieu -L (pour vérifier l'excision du marqueur), sur milieu -U (pour confirmer la perte du plasmide pUDC073) et YPG (pour garder les clones d'intérêt). Cinq clones -L-U sont testés par PCR afin de vérifier l'excision du marqueur de sélection KILEU2.

Références bibliographiques

Références bibliographiques

- 1 Warner, J. R., Vilardell, J. & Sohn, J. H. Economics of ribosome biosynthesis. *Cold Spring Harb Symp Quant Biol* **66**, 567-574 (2001).
- 2 Bentley, D. L. Coupling mRNA processing with transcription in time and space. *Nat Rev Genet* **15**, 163-175, doi:10.1038/nrg3662 (2014).
- 3 Jurado, A. R., Tan, D., Jiao, X., Kiledjian, M. & Tong, L. Structure and function of pre-mRNA 5'-end capping quality control and 3'-end processing. *Biochemistry* **53**, 1882-1898, doi:10.1021/bi401715v (2014).
- 4 Edery, I. & Sonenberg, N. Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc Natl Acad Sci U S A* **82**, 7590-7594 (1985).
- 5 Wurth, L. *et al.* Hypermethylated-capped selenoprotein mRNAs in mammals. *Nucleic Acids Res* **42**, 8663-8677, doi:10.1093/nar/gku580 (2014).
- 6 Matera, A. G., Terns, R. M. & Terns, M. P. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* **8**, 209-220, doi:10.1038/nrm2124 (2007).
- 7 Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev* **25**, 1770-1782, doi:10.1101/gad.17268411 (2011).
- 8 Kühn, U. *et al.* Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J Biol Chem* **284**, 22803-22814, doi:10.1074/jbc.M109.018226 (2009).
- 9 Lutz, C. S. & Moreira, A. Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression. *Wiley Interdiscip Rev RNA* **2**, 22-31, doi:10.1002/wrna.47 (2011).
- 10 Shi, Y. Alternative polyadenylation: new insights from global analyses. *RNA* **18**, 2105-2117, doi:10.1261/rna.035899.112 (2012).
- 11 Elkon, R., Ugalde, A. P. & Agami, R. Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet* **14**, 496-506, doi:10.1038/nrg3482 (2013).
- 12 Tian, B. & Manley, J. L. Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci* **38**, 312-320, doi:10.1016/j.tibs.2013.03.005 (2013).
- 13 Kahvejian, A., Roy, G. & Sonenberg, N. The mRNA closed-loop model: the function of PABP and PABP-interacting proteins in mRNA translation. *Cold Spring Harb Symp Quant Biol* **66**, 293-300 (2001).
- 14 Archer, S. K., Shirokikh, N. E., Hallwirth, C. V., Beilharz, T. H. & Preiss, T. Probing the closed-loop model of mRNA translation in living cells. *RNA Biol* **12**, 248-254, doi:10.1080/15476286.2015.1017242 (2015).
- 15 Will, C. L. & Lührmann, R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol* **3**, doi:10.1101/cshperspect.a003707 (2011).
- 16 Han, J., Xiong, J., Wang, D. & Fu, X. D. Pre-mRNA splicing: where and when in the nucleus. *Trends Cell Biol* **21**, 336-343, doi:10.1016/j.tcb.2011.03.003 (2011).
- 17 Lee, Y. & Rio, D. C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu Rev Biochem* **84**, 291-323, doi:10.1146/annurev-biochem-060614-034316 (2015).
- 18 Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**, 457-463, doi:10.1038/nature08909 (2010).
- 19 Li, S. & Mason, C. E. The pivotal regulatory landscape of RNA modifications. *Annu Rev Genomics Hum Genet* **15**, 127-150, doi:10.1146/annurev-genom-090413-025405 (2014).
- 20 Machnicka, M. A. *et al.* MODOMICS: a database of RNA modification pathways--2013 update. *Nucleic Acids Res* **41**, D262-267, doi:10.1093/nar/gks1007 (2013).
- 21 Hoernes, T. P., Hüttenhofer, A. & Erlacher, M. D. mRNA modifications: Dynamic regulators of gene expression? *RNA Biol* **13**, 760-765, doi:10.1080/15476286.2016.1203504 (2016).
- 22 Barciszewska, M. Z., Perrigue, P. M. & Barciszewski, J. tRNA--the golden standard in molecular biology. *Mol Biosyst* **12**, 12-17, doi:10.1039/c5mb00557d (2016).

23 Giegé, R. *et al.* Structure of transfer RNAs: similarity and variability. *Wiley Interdiscip Rev RNA* **3**, 37-61, doi:10.1002/wrna.103 (2012).

24 Zachau, H. G., Acs, G. & Lipmann, F. ISOLATION OF ADENOSINE AMINO ACID ESTERS FROM A RIBONUCLEASE DIGEST OF SOLUBLE, LIVER RIBONUCLEIC ACID. *Proc Natl Acad Sci U S A* **44**, 885-889 (1958).

25 Rajbhandary, U. L. *et al.* Studies on polynucleotides, Ixviii the primary structure of yeast phenylalanine transfer RNA. *Proc Natl Acad Sci U S A* **57**, 751-758 (1967).

26 Kim, S. H. *et al.* Three-dimensional structure of yeast phenylalanine transfer RNA: folding of the polynucleotide chain. *Science* **179**, 285-288 (1973).

27 Tuorto, F. & Lyko, F. Genome recoding by tRNA modifications. *Open Biol* **6**, doi:10.1098/rsob.160287 (2016).

28 El Yacoubi, B., Bailly, M. & de Crécy-Lagard, V. Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annu Rev Genet* **46**, 69-95, doi:10.1146/annurev-genet-110711-155641 (2012).

29 Guy, M. P. *et al.* Identification of the determinants of tRNA function and susceptibility to rapid tRNA decay by high-throughput in vivo analysis. *Genes Dev* **28**, 1721-1732, doi:10.1101/gad.245936.114 (2014).

30 Agris, P. F., Vendeix, F. A. & Graham, W. D. tRNA's wobble decoding of the genome: 40 years of modification. *J Mol Biol* **366**, 1-13, doi:10.1016/j.jmb.2006.11.046 (2007).

31 Machnicka, M. A., Olchowik, A., Grosjean, H. & Bujnicki, J. M. Distribution and frequencies of post-transcriptional modifications in tRNAs. *RNA Biol* **11**, 1619-1629, doi:10.4161/15476286.2014.992273 (2014).

32 Zhou, X. & Wang, E. Transfer RNA: a dancer between charging and mis-charging for protein biosynthesis. *Sci China Life Sci* **56**, 921-932, doi:10.1007/s11427-013-4542-9 (2013).

33 Perona, J. J. & Gruic-Sovulj, I. Synthetic and editing mechanisms of aminoacyl-tRNA synthetases. *Top Curr Chem* **344**, 1-41, doi:10.1007/128_2013_456 (2014).

34 Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920-930 (2000).

35 Sun, M. *et al.* Dynamical features of the Plasmodium falciparum ribosome during translation. *Nucleic Acids Res* **43**, 10515-10524, doi:10.1093/nar/gkv991 (2015).

36 Anger, A. M. *et al.* Structures of the human and Drosophila 80S ribosome. *Nature* **497**, 80-85, doi:10.1038/nature12104 (2013).

37 Selmer, M. *et al.* Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935-1942, doi:10.1126/science.1131127 (2006).

38 Klinge, S., Voigts-Hoffmann, F., Leibundgut, M. & Ban, N. Atomic structures of the eukaryotic ribosome. *Trends Biochem Sci* **37**, 189-198, doi:10.1016/j.tibs.2012.02.007 (2012).

39 Khatter, H., Myasnikov, A. G., Natchiar, S. K. & Klaholz, B. P. Structure of the human 80S ribosome. *Nature* **520**, 640-645, doi:10.1038/nature14427 (2015).

40 Sharma, S. & Lafontaine, D. L. J. 'View From A Bridge': A New Perspective on Eukaryotic rRNA Base Modification. *Trends Biochem Sci* **40**, 560-575, doi:10.1016/j.tibs.2015.07.008 (2015).

41 Lafontaine, D. L. & Tollervey, D. Birth of the snoRNPs: the evolution of the modification-guide snoRNAs. *Trends Biochem Sci* **23**, 383-388 (1998).

42 Fleischer, T. C., Weaver, C. M., McAfee, K. J., Jennings, J. L. & Link, A. J. Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes Dev* **20**, 1294-1307, doi:10.1101/gad.1422006 (2006).

43 Colón-Ramos, D. A. *et al.* Direct ribosomal binding by a cellular inhibitor of translation. *Nat Struct Mol Biol* **13**, 103-111, doi:10.1038/nsmb1052 (2006).

44 Fuchs, G., Diges, C., Kohlstaedt, L. A., Wehner, K. A. & Sarnow, P. Proteomic analysis of ribosomes: translational control of mRNA populations by glycogen synthase GYS1. *J Mol Biol* **410**, 118-130, doi:10.1016/j.jmb.2011.04.064 (2011).

45 Ceci, M. *et al.* Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* **426**, 579-584, doi:10.1038/nature02160 (2003).

46 Natchiar, S. K., Myasnikov, A. G., Kratzat, H., Hazemann, I. & Klaholz, B. P. Visualization of chemical modifications in the human 80S ribosome structure. *Nature* **551**, 472-477, doi:10.1038/nature24482 (2017).

47 Baum, S., Bittins, M., Frey, S. & Seedorf, M. Asc1p, a WD40-domain containing adaptor protein, is required for the interaction of the RNA-binding protein Scp160p with polysomes. *Biochem J* **380**, 823-830, doi:10.1042/BJ20031962 (2004).

48 Xue, S. & Barna, M. Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat Rev Mol Cell Biol* **13**, 355-369, doi:10.1038/nrm3359 (2012).

49 Moss, T., Langlois, F., Gagnon-Kugler, T. & Stefanovsky, V. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. *Cell Mol Life Sci* **64**, 29-49, doi:10.1007/s00018-006-6278-1 (2007).

50 de la Cruz, J., Karbstein, K. & Woolford, J. L. Functions of ribosomal proteins in assembly of eukaryotic ribosomes in vivo. *Annu Rev Biochem* **84**, 93-129, doi:10.1146/annurev-biochem-060614-033917 (2015).

51 Tafforeau, L. *et al.* The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol Cell* **51**, 539-551, doi:10.1016/j.molcel.2013.08.011 (2013).

52 Lestrade, L. & Weber, M. J. snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res* **34**, D158-162, doi:10.1093/nar/gkj002 (2006).

53 Thomson, E., Ferreira-Cerca, S. & Hurt, E. Eukaryotic ribosome biogenesis at a glance. *J Cell Sci* **126**, 4815-4821, doi:10.1242/jcs.111948 (2013).

54 Yu, Y., Ji, H., Doudna, J. A. & Leary, J. A. Mass spectrometric analysis of the human 40S ribosomal subunit: native and HCV IRES-bound complexes. *Protein Sci* **14**, 1438-1446, doi:10.1110/ps.041293005 (2005).

55 Odintsova, T. I. *et al.* Characterization and analysis of posttranslational modifications of the human large cytoplasmic ribosomal subunit proteins by mass spectrometry and Edman sequencing. *J Protein Chem* **22**, 249-258 (2003).

56 Schneider, D. A. RNA polymerase I activity is regulated at multiple steps in the transcription cycle: recent insights into factors that influence transcription elongation. *Gene* **493**, 176-184, doi:10.1016/j.gene.2011.08.006 (2012).

57 Matera, A. G., Frey, M. R., Margelot, K. & Wolin, S. L. A perinucleolar compartment contains several RNA polymerase III transcripts as well as the polypyrimidine tract-binding protein, hnRNP I. *J Cell Biol* **129**, 1181-1193 (1995).

58 Thiry, M. & Lafontaine, D. L. Birth of a nucleolus: the evolution of nucleolar compartments. *Trends Cell Biol* **15**, 194-199, doi:10.1016/j.tcb.2005.02.007 (2005).

59 Miller, O. L. The nucleolus, chromosomes, and visualization of genetic activity. *J Cell Biol* **91**, 15s-27s (1981).

60 Kos, M. & Tollervey, D. Yeast pre-rRNA processing and modification occur cotranscriptionally. *Mol Cell* **37**, 809-820, doi:10.1016/j.molcel.2010.02.024 (2010).

61 Mullineux, S. T. & Lafontaine, D. L. Mapping the cleavage sites on mammalian pre-rRNAs: where do we stand? *Biochimie* **94**, 1521-1532, doi:10.1016/j.biochi.2012.02.001 (2012).

62 Hennras, A. K., Plisson-Chastang, C., O'Donohue, M. F., Chakraborty, A. & Gleizes, P. E. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip Rev RNA* **6**, 225-242, doi:10.1002/wrna.1269 (2015).

63 Hadjolova, K. V., Nicoloso, M., Mazan, S., Hadjolov, A. A. & Bachellerie, J. P. Alternative pre-rRNA processing pathways in human cells and their alteration by cycloheximide inhibition of protein synthesis. *Eur J Biochem* **212**, 211-215 (1993).

64 Gadal, O. *et al.* Nuclear export of 60s ribosomal subunits depends on Xpo1p and requires a nuclear export sequence-containing factor, Nmd3p, that associates with the large subunit protein Rpl10p. *Mol Cell Biol* **21**, 3405-3415, doi:10.1128/MCB.21.10.3405-3415.2001 (2001).

65 Bange, G., Murat, G., Sinning, I., Hurt, E. & Kressler, D. New twist to nuclear import: When two travel together. *Commun Integr Biol* **6**, e24792, doi:10.4161/cib.24792 (2013).

66 Karbstein, K. Chaperoning ribosome assembly. *J Cell Biol* **189**, 11-12, doi:10.1083/jcb.201002102 (2010).

67 Karbstein, K. Inside the 40S ribosome assembly machinery. *Curr Opin Chem Biol* **15**, 657-663, doi:10.1016/j.cbpa.2011.07.023 (2011).

68 Mulder, A. M. *et al.* Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit. *Science* **330**, 673-677, doi:10.1126/science.1193220 (2010).

69 O'Donohue, M. F., Choesmel, V., Faubladier, M., Fichant, G. & Gleizes, P. E. Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. *J Cell Biol* **190**, 853-866, doi:10.1083/jcb.201005117 (2010).

70 Rodríguez-Mateos, M. *et al.* The amino terminal domain from Mrt4 protein can functionally replace the RNA binding domain of the ribosomal P0 protein. *Nucleic Acids Res* **37**, 3514-3521, doi:10.1093/nar/gkp209 (2009).

71 Strunk, B. S. *et al.* Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. *Science* **333**, 1449-1453, doi:10.1126/science.1208245 (2011).

72 Dunbar, D. A., Dragon, F., Lee, S. J. & Baserga, S. J. A nucleolar protein related to ribosomal protein L7 is required for an early step in large ribosomal subunit biogenesis. *Proc Natl Acad Sci U S A* **97**, 13027-13032, doi:10.1073/pnas.97.24.13027 (2000).

73 Narla, A. & Ebert, B. L. Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115**, 3196-3205, doi:10.1182/blood-2009-10-178129 (2010).

74 Fumagalli, S. & Thomas, G. The role of p53 in ribosomopathies. *Semin Hematol* **48**, 97-105, doi:10.1053/j.seminhematol.2011.02.004 (2011).

75 Bursać, S. *et al.* Mutual protection of ribosomal proteins L5 and L11 from degradation is essential for p53 activation upon ribosomal biogenesis stress. *Proc Natl Acad Sci U S A* **109**, 20467-20472, doi:10.1073/pnas.1218535109 (2012).

76 Yadav, G. V., Chakraborty, A., Uechi, T. & Kenmochi, N. Ribosomal protein deficiency causes Tp53-independent erythropoiesis failure in zebrafish. *Int J Biochem Cell Biol* **49**, 1-7, doi:10.1016/j.biocel.2014.01.006 (2014).

77 Reichow, S. L., Hamma, T., Ferré-D'Amaré, A. R. & Varani, G. The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res* **35**, 1452-1464, doi:10.1093/nar/gkl1172 (2007).

78 Mills, E. W. & Green, R. Ribosomopathies: There's strength in numbers. *Science* **358**, doi:10.1126/science.aan2755 (2017).

79 Kondrashov, N. *et al.* Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* **145**, 383-397, doi:10.1016/j.cell.2011.03.028 (2011).

80 Horos, R. *et al.* Ribosomal deficiencies in Diamond-Blackfan anemia impair translation of transcripts essential for differentiation of murine and human erythroblasts. *Blood* **119**, 262-272, doi:10.1182/blood-2011-06-358200 (2012).

81 Krogh, N. *et al.* Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res* **44**, 7884-7895, doi:10.1093/nar/gkw482 (2016).

82 Hüttenhofer, A. *et al.* RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J* **20**, 2943-2953, doi:10.1093/emboj/20.11.2943 (2001).

83 Ganot, P., Bortolin, M. L. & Kiss, T. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. *Cell* **89**, 799-809 (1997).

84 Henras, A. K., Plisson-Chastang, C., Humbert, O., Romeo, Y. & Henry, Y. Synthesis, Function, and Heterogeneity of snoRNA-Guided Posttranscriptional Nucleoside Modifications in Eukaryotic Ribosomal RNAs. *Enzymes* **41**, 169-213, doi:10.1016/bs.enz.2017.03.007 (2017).

85 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* **3**, 397-414, doi:10.1002/wrna.117 (2012).

86 Rodriguez-Corona, U., Sobol, M., Rodriguez-Zapata, L. C., Hozak, P. & Castano, E. Fibrillarin from Archaea to human. *Biol Cell* **107**, 159-174, doi:10.1111/boc.201400077 (2015).

87 van Nues, R. W. & Watkins, N. J. Unusual C'/D' motifs enable box C/D snoRNPs to modify multiple sites in the same rRNA target region. *Nucleic Acids Res* **45**, 2016-2028, doi:10.1093/nar/gkw842 (2017).

88 van Nues, R. W. *et al.* Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. *EMBO J* **30**, 2420-2430, doi:10.1038/emboj.2011.148 (2011).

89 Tessarz, P. *et al.* Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature* **505**, 564-568, doi:10.1038/nature12819 (2014).

90 Dieci, G., Preti, M. & Montanini, B. Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics* **94**, 83-88, doi:10.1016/j.ygeno.2009.05.002 (2009).

91 Bratkovič, T. & Rogelj, B. Biology and applications of small nucleolar RNAs. *Cell Mol Life Sci* **68**, 3843-3851, doi:10.1007/s00018-011-0762-y (2011).

92 Basu, A. *et al.* Requirement of rRNA methylation for 80S ribosome assembly on a cohort of cellular internal ribosome entry sites. *Mol Cell Biol* **31**, 4482-4499, doi:10.1128/MCB.05804-11 (2011).

93 Zebarjadian, Y., King, T., Fournier, M. J., Clarke, L. & Carbon, J. Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA. *Mol Cell Biol* **19**, 7461-7472 (1999).

94 Angrisani, A., Vicedomini, R., Turano, M. & Furia, M. Human dyskerin: beyond telomeres. *Biol Chem* **395**, 593-610, doi:10.1515/hsz-2013-0287 (2014).

95 Ruggero, D. *et al.* Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. *Science* **299**, 259-262, doi:10.1126/science.1079447 (2003).

96 Mochizuki, Y., He, J., Kulkarni, S., Bessler, M. & Mason, P. J. Mouse dyskerin mutations affect accumulation of telomerase RNA and small nucleolar RNA, telomerase activity, and ribosomal RNA processing. *Proc Natl Acad Sci U S A* **101**, 10756-10761, doi:10.1073/pnas.0402560101 (2004).

97 Yoon, A. *et al.* Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science* **312**, 902-906, doi:10.1126/science.1123835 (2006).

98 Jack, K. *et al.* rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol Cell* **44**, 660-666, doi:10.1016/j.molcel.2011.09.017 (2011).

99 Tollervey, D., Lehtonen, H., Jansen, R., Kern, H. & Hurt, E. C. Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell* **72**, 443-457 (1993).

100 Caboche, M. & Bachellerie, J. P. RNA methylation and control of eukaryotic RNA biosynthesis. Effects of cycloleucine, a specific inhibitor of methylation, on ribosomal RNA maturation. *Eur J Biochem* **74**, 19-29 (1977).

101 Marcel, V. *et al.* p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. *Cancer Cell* **24**, 318-330, doi:10.1016/j.ccr.2013.08.013 (2013).

102 Baudin-Baillieu, A. *et al.* Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids Res* **37**, 7665-7677, doi:10.1093/nar/gkp816 (2009).

103 Higa-Nakamine, S. *et al.* Loss of ribosomal RNA modification causes developmental defects in zebrafish. *Nucleic Acids Res* **40**, 391-398, doi:10.1093/nar/gkr700 (2012).

104 Liang, X. H., Liu, Q. & Fournier, M. J. Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA* **15**, 1716-1728, doi:10.1261/rna.1724409 (2009).

105 King, T. H., Liu, B., McCully, R. R. & Fournier, M. J. Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol Cell* **11**, 425-435 (2003).

106 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* **28**, 965-977, doi:10.1016/j.molcel.2007.10.012 (2007).

107 Badis, G., Fromont-Racine, M. & Jacquier, A. A snoRNA that guides the two most conserved pseudouridine modifications within rRNA confers a growth advantage in yeast. *RNA* **9**, 771-779 (2003).

108 Mannoor, K., Liao, J. & Jiang, F. Small nucleolar RNAs in cancer. *Biochim Biophys Acta* **1826**, 121-128, doi:10.1016/j.bbcan.2012.03.005 (2012).

109 Ruggero, D. & Pandolfi, P. P. Does the ribosome translate cancer? *Nat Rev Cancer* **3**, 179-192, doi:10.1038/nrc1015 (2003).

110 Su, H. *et al.* Elevated snoRNA biogenesis is essential in breast cancer. *Oncogene* **33**, 1348-1358, doi:10.1038/onc.2013.89 (2014).

111 Dong, X. Y. *et al.* SnoRNA U50 is a candidate tumor-suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. *Hum Mol Genet* **17**, 1031-1042, doi:10.1093/hmg/ddm375 (2008).

112 Dong, X. Y. *et al.* Implication of snoRNA U50 in human breast cancer. *J Genet Genomics* **36**, 447-454, doi:10.1016/S1673-8527(08)60134-4 (2009).

113 Pacilli, A., Ceccarelli, C., Treré, D. & Montanaro, L. SnoRNA U50 levels are regulated by cell proliferation and rRNA transcription. *Int J Mol Sci* **14**, 14923-14935, doi:10.3390/ijms140714923 (2013).

114 Koh, C. M. *et al.* Myc enforces overexpression of EZH2 in early prostatic neoplasia via transcriptional and post-transcriptional mechanisms. *Oncotarget* **2**, 669-683, doi:10.18632/oncotarget.327 (2011).

115 Miller, D. M., Thomas, S. D., Islam, A., Muench, D. & Sedoris, K. c-Myc and cancer metabolism. *Clin Cancer Res* **18**, 5546-5553, doi:10.1158/1078-0432.CCR-12-0977 (2012).

116 Melén, K. *et al.* Influenza A H3N2 subtype virus NS1 protein targets into the nucleus and binds primarily via its C-terminal NLS2/NoLS to nucleolin and fibrillarin. *Virol J* **9**, 167, doi:10.1186/1743-422X-9-167 (2012).

117 Yoo, D., Woottton, S. K., Li, G., Song, C. & Rowland, R. R. Colocalization and interaction of the porcine arterivirus nucleocapsid protein with the small nucleolar RNA-associated protein fibrillarin. *J Virol* **77**, 12173-12183 (2003).

118 Schuller, A. P. & Green, R. The ABC(E1)s of Ribosome Recycling and Reinitiation. *Mol Cell* **66**, 578-580, doi:10.1016/j.molcel.2017.05.017 (2017).

119 Haimov, O., Sinvani, H. & Dikstein, R. Cap-dependent, scanning-free translation initiation mechanisms. *Biochim Biophys Acta* **1849**, 1313-1318, doi:10.1016/j.bbagr.2015.09.006 (2015).

120 Dever, T. E., Kinzy, T. G. & Pavitt, G. D. Mechanism and Regulation of Protein Synthesis in *Saccharomyces cerevisiae*. *Genetics* **203**, 65-107, doi:10.1534/genetics.115.186221 (2016).

121 Aitken, C. E. & Lorsch, J. R. A mechanistic overview of translation initiation in eukaryotes. *Nat Struct Mol Biol* **19**, 568-576, doi:10.1038/nsmb.2303 (2012).

122 Wells, S. E., Hillner, P. E., Vale, R. D. & Sachs, A. B. Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* **2**, 135-140 (1998).

123 Park, E. H. *et al.* Multiple elements in the eIF4G1 N-terminus promote assembly of eIF4G1•PABP mRNPs in vivo. *EMBO J* **30**, 302-316, doi:10.1038/emboj.2010.312 (2011).

124 Cigan, A. M., Feng, L. & Donahue, T. F. tRNA^{i(met)} functions in directing the scanning ribosome to the start site of translation. *Science* **242**, 93-97 (1988).

125 Hinnebusch, A. G. Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiol Mol Biol Rev* **75**, 434-467, first page of table of contents, doi:10.1128/MMBR.00008-11 (2011).

126 Kozak, M. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res* **9**, 5233-5252 (1981).

127 Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218-223, doi:10.1126/science.1168978 (2009).

128 Ingolia, N. T., Lareau, L. F. & Weissman, J. S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**, 789-802, doi:10.1016/j.cell.2011.10.002 (2011).

129 Jackson, R. J., Hellen, C. U. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11**, 113-127, doi:10.1038/nrm2838 (2010).

130 Mailliot, J. & Martin, F. Viral internal ribosomal entry sites: four classes for one goal. *Wiley Interdiscip Rev RNA* **9**, doi:10.1002/wrna.1458 (2018).

131 Walters, B. & Thompson, S. R. Cap-Independent Translational Control of Carcinogenesis. *Front Oncol* **6**, 128, doi:10.3389/fonc.2016.00128 (2016).

132 Lacerda, R., Menezes, J. & Romão, L. More than just scanning: the importance of cap-independent mRNA translation initiation for cellular stress response and cancer. *Cell Mol Life Sci* **74**, 1659-1680, doi:10.1007/s00018-016-2428-2 (2017).

133 Weingarten-Gabbay, S. *et al.* Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science* **351**, doi:10.1126/science.aad4939 (2016).

134 King, H. A., Cobbold, L. C. & Willis, A. E. The role of IRES trans-acting factors in regulating translation initiation. *Biochem Soc Trans* **38**, 1581-1586, doi:10.1042/BST0381581 (2010).

135 Komar, A. A. & Hatzoglou, M. Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. *Cell Cycle* **10**, 229-240, doi:10.4161/cc.10.2.14472 (2011).

136 Leppek, K., Das, R. & Barna, M. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* **19**, 158-174, doi:10.1038/nrm.2017.103 (2018).

137 Mitchell, S. A., Spriggs, K. A., Coldwell, M. J., Jackson, R. J. & Willis, A. E. The Apaf-1 internal ribosome entry segment attains the correct structural conformation for function via interactions with PTB and unr. *Mol Cell* **11**, 757-771 (2003).

138 Sinvani, H. *et al.* Translational tolerance of mitochondrial genes to metabolic energy stress involves TISU and eIF1-eIF4GI cooperation in start codon selection. *Cell Metab* **21**, 479-492, doi:10.1016/j.cmet.2015.02.010 (2015).

139 Dever, T. E. & Green, R. The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb Perspect Biol* **4**, a013706, doi:10.1101/cshperspect.a013706 (2012).

140 Knight, R. D., Freeland, S. J. & Landweber, L. F. Rewiring the keyboard: evolvability of the genetic code. *Nat Rev Genet* **2**, 49-58, doi:10.1038/35047500 (2001).

141 Zhang, G. & Ignatova, Z. Folding at the birth of the nascent chain: coordinating translation with co-translational folding. *Curr Opin Struct Biol* **21**, 25-31, doi:10.1016/j.sbi.2010.10.008 (2011).

142 Saint-Léger, A. & Ribas de Pouplana, L. The importance of codon-anticodon interactions in translation elongation. *Biochimie* **114**, 72-79, doi:10.1016/j.biochi.2015.04.013 (2015).

143 Rodnina, M. V. & Wintermeyer, W. Protein Elongation, Co-translational Folding and Targeting. *J Mol Biol* **428**, 2165-2185, doi:10.1016/j.jmb.2016.03.022 (2016).

144 Jackson, R. J., Hellen, C. U. & Pestova, T. V. Termination and post-termination events in eukaryotic translation. *Adv Protein Chem Struct Biol* **86**, 45-93, doi:10.1016/B978-0-12-386497-0.00002-5 (2012).

145 Huard, F. P., Deane, C. M. & Wood, G. R. Modelling sequential protein folding under kinetic control. *Bioinformatics* **22**, e203-210, doi:10.1093/bioinformatics/btl248 (2006).

146 Thanaraj, T. A. & Argos, P. Protein secondary structural types are differentially coded on messenger RNA. *Protein Sci* **5**, 1973-1983, doi:10.1002/pro.5560051003 (1996).

147 Song, H. *et al.* The crystal structure of human eukaryotic release factor eRF1--mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell* **100**, 311-321 (2000).

148 Frolova, L. Y. *et al.* Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. *RNA* **5**, 1014-1020 (1999).

149 Merkulova, T. I., Frolova, L. Y., Lazar, M., Camonis, J. & Kisselov, L. L. C-terminal domains of human translation termination factors eRF1 and eRF3 mediate their in vivo interaction. *FEBS Lett* **443**, 41-47 (1999).

150 Frolova, L., Seit-Nebi, A. & Kisselov, L. Highly conserved NIKS tetrapeptide is functionally essential in eukaryotic translation termination factor eRF1. *RNA* **8**, 129-136 (2002).

151 Chavatte, L., Kervestin, S., Favre, A. & Jean-Jean, O. Stop codon selection in eukaryotic translation termination: comparison of the discriminating potential between human and ciliate eRF1s. *EMBO J* **22**, 1644-1653, doi:10.1093/emboj/cdg146 (2003).

152 Conard, S. E. *et al.* Identification of eRF1 residues that play critical and complementary roles in stop codon recognition. *RNA* **18**, 1210-1221, doi:10.1261/rna.031997.111 (2012).

153 Wong, L. E., Li, Y., Pillay, S., Frolova, L. & Pervushin, K. Selectivity of stop codon recognition in translation termination is modulated by multiple conformations of GTS loop in eRF1. *Nucleic Acids Res* **40**, 5751-5765, doi:10.1093/nar/gks192 (2012).

154 Hatin, I., Fabret, C., Rousset, J. P. & Namy, O. Molecular dissection of translation termination mechanism identifies two new critical regions in eRF1. *Nucleic Acids Res* **37**, 1789-1798, doi:10.1093/nar/gkp012 (2009).

155 Blanchet, S. *et al.* New insights into stop codon recognition by eRF1. *Nucleic Acids Res* **43**, 3298-3308, doi:10.1093/nar/gkv154 (2015).

156 Brown, A., Shao, S., Murray, J., Hegde, R. S. & Ramakrishnan, V. Structural basis for stop codon recognition in eukaryotes. *Nature* **524**, 493-496, doi:10.1038/nature14896 (2015).

157 Hoshino, S., Imai, M., Kobayashi, T., Uchida, N. & Katada, T. The eukaryotic polypeptide chain releasing factor (eRF3/GSPT) carrying the translation termination signal to the 3'-Poly(A) tail of mRNA. Direct association of erf3/GSPT with polyadenylate-binding protein. *J Biol Chem* **274**, 16677-16680 (1999).

158 Cosson, B. *et al.* Poly(A)-binding protein acts in translation termination via eukaryotic release factor 3 interaction and does not influence [PSI(+)] propagation. *Mol Cell Biol* **22**, 3301-3315 (2002).

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* **27**, 736-747, doi:10.1038/emboj.2008.17 (2008).

160 Alkalaeva, E. Z., Pisarev, A. V., Frolova, L. Y., Kisselov, L. L. & Pestova, T. V. In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. *Cell* **125**, 1125-1136, doi:10.1016/j.cell.2006.04.035 (2006).

161 Frolova, L. *et al.* Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. *RNA* **2**, 334-341 (1996).

162 Preis, A. *et al.* Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. *Cell Rep* **8**, 59-65, doi:10.1016/j.celrep.2014.04.058 (2014).

163 Heuer, A. *et al.* Structure of the 40S-ABCE1 post-splitting complex in ribosome recycling and translation initiation. *Nat Struct Mol Biol* **24**, 453-460, doi:10.1038/nsmb.3396 (2017).

164 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* **108**, E1392-1398, doi:10.1073/pnas.1113956108 (2011).

165 Pisarev, A. V. *et al.* The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Mol Cell* **37**, 196-210, doi:10.1016/j.molcel.2009.12.034 (2010).

166 Rabl, J., Leibundgut, M., Ataide, S. F., Haag, A. & Ban, N. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* **331**, 730-736, doi:10.1126/science.1198308 (2011).

167 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* **334**, 941-948, doi:10.1126/science.1211204 (2011).

168 Skabkin, M. A., Skabkina, O. V., Hellen, C. U. & Pestova, T. V. Reinitiation and other unconventional posttermination events during eukaryotic translation. *Mol Cell* **51**, 249-264, doi:10.1016/j.molcel.2013.05.026 (2013).

169 Gunišová, S., Hronová, V., Mohammad, M. P., Hinnebusch, A. G. & Valášek, L. S. Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. *FEMS Microbiol Rev* **42**, 165-192, doi:10.1093/femsre/fux059 (2018).

170 Luukkonen, B. G., Tan, W. & Schwartz, S. Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *J Virol* **69**, 4086-4094 (1995).

171 Zinoviev, A., Hellen, C. U. T. & Pestova, T. V. Multiple mechanisms of reinitiation on bicistronic calicivirus mRNAs. *Mol Cell* **57**, 1059-1073, doi:10.1016/j.molcel.2015.01.039 (2015).

172 Hershey, J. W., Sonenberg, N. & Mathews, M. B. Principles of translational control: an overview. *Cold Spring Harb Perspect Biol* **4**, doi:10.1101/cshperspect.a011528 (2012).

173 Schmidt, C., Beilsten-Edmands, V. & Robinson, C. V. Insights into Eukaryotic Translation Initiation from Mass Spectrometry of Macromolecular Protein Assemblies. *J Mol Biol* **428**, 344-356, doi:10.1016/j.jmb.2015.10.011 (2016).

174 Krishnamoorthy, T., Pavitt, G. D., Zhang, F., Dever, T. E. & Hinnebusch, A. G. Tight binding of the phosphorylated alpha subunit of initiation factor 2 (eIF2alpha) to the regulatory subunits of guanine nucleotide exchange factor eIF2B is required for inhibition of translation initiation. *Mol Cell Biol* **21**, 5018-5030, doi:10.1128/MCB.21.15.5018-5030.2001 (2001).

175 Spilka, R., Ernst, C., Mehta, A. K. & Haybaeck, J. Eukaryotic translation initiation factors in cancer development and progression. *Cancer Lett* **340**, 9-21, doi:10.1016/j.canlet.2013.06.019 (2013).

176 Ma, X. M. & Blenis, J. Molecular mechanisms of mTOR-mediated translational control. *Nat Rev Mol Cell Biol* **10**, 307-318, doi:10.1038/nrm2672 (2009).

177 Inoki, K., Zhu, T. & Guan, K. L. TSC2 mediates cellular energy response to control cell growth and survival. *Cell* **115**, 577-590 (2003).

178 Jacinto, E. & Hall, M. N. Tor signalling in bugs, brain and brawn. *Nat Rev Mol Cell Biol* **4**, 117-126, doi:10.1038/nrm1018 (2003).

179 Masui, K., Cavenee, W. K. & Mischel, P. S. mTORC2 in the center of cancer metabolic reprogramming. *Trends Endocrinol Metab* **25**, 364-373, doi:10.1016/j.tem.2014.04.002 (2014).

180 Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307**, 1098-1101, doi:10.1126/science.1106148 (2005).

181 Yang, Q. & Guan, K. L. Expanding mTOR signaling. *Cell Res* **17**, 666-681, doi:10.1038/cr.2007.64 (2007).

182 Pause, A. *et al.* Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature* **371**, 762-767, doi:10.1038/371762a0 (1994).

183 Wethmar, K. The regulatory potential of upstream open reading frames in eukaryotic gene expression. *Wiley Interdiscip Rev RNA* **5**, 765-778, doi:10.1002/wrna.1245 (2014).

184 Gandin, V. *et al.* mTORC1 and CK2 coordinate ternary and eIF4F complex assembly. *Nat Commun* **7**, 11127, doi:10.1038/ncomms11127 (2016).

185 Roux, P. P. & Topisirovic, I. Regulation of mRNA translation by signaling pathways. *Cold Spring Harb Perspect Biol* **4**, doi:10.1101/cshperspect.a012252 (2012).

186 Johnstone, T. G., Bazzini, A. A. & Giraldez, A. J. Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J* **35**, 706-723, doi:10.15252/embj.201592759 (2016).

187 Schleich, S. *et al.* DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. *Nature* **512**, 208-212, doi:10.1038/nature13401 (2014).

188 Schulz, J. *et al.* Loss-of-function uORF mutations in human malignancies. *Sci Rep* **8**, 2395, doi:10.1038/s41598-018-19201-8 (2018).

189 Law, G. L., Raney, A., Heusner, C. & Morris, D. R. Polyamine regulation of ribosome pausing at the upstream open reading frame of S-adenosylmethionine decarboxylase. *J Biol Chem* **276**, 38036-38043, doi:10.1074/jbc.M105944200 (2001).

190 Kozak, M. Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **299**, 1-34 (2002).

191 Ait Ghezala, H. *et al.* Translation termination efficiency modulates ATF4 response by regulating ATF4 mRNA translation at 5' short ORFs. *Nucleic Acids Res* **40**, 9557-9570, doi:10.1093/nar/gks762 (2012).

192 Brown, C. Y., Mize, G. J., Pineda, M., George, D. L. & Morris, D. R. Role of two upstream open reading frames in the translational control of oncogene mdm2. *Oncogene* **18**, 5631-5637, doi:10.1038/sj.onc.1202949 (1999).

193 Hinnebusch, A. G. Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol* **59**, 407-450, doi:10.1146/annurev.micro.59.031805.133833 (2005).

194 Pegg, A. E., Xiong, H., Feith, D. J. & Shantz, L. M. S-adenosylmethionine decarboxylase: structure, function and regulation by polyamines. *Biochem Soc Trans* **26**, 580-586 (1998).

195 Braunstein, S. *et al.* A hypoxia-controlled cap-dependent to cap-independent translation switch in breast cancer. *Mol Cell* **28**, 501-512, doi:10.1016/j.molcel.2007.10.019 (2007).

196 Ricci, E. P. *et al.* miRNA repression of translation in vitro takes place during 43S ribosomal scanning. *Nucleic Acids Res* **41**, 586-598, doi:10.1093/nar/gks1076 (2013).

197 Meyer, K. D. *et al.* 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell* **163**, 999-1010, doi:10.1016/j.cell.2015.10.012 (2015).

198 Zaher, H. S. & Green, R. Quality control by the ribosome following peptide bond formation. *Nature* **457**, 161-166, doi:10.1038/nature07582 (2009).

199 Lind, C. & Åqvist, J. Principles of start codon recognition in eukaryotic translation initiation. *Nucleic Acids Res* **44**, 8425-8432, doi:10.1093/nar/gkw534 (2016).

200 Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M. & Yusupova, G. A new understanding of the decoding principle on the ribosome. *Nature* **484**, 256-259, doi:10.1038/nature10913 (2012).

201 Noller, H. F., Lancaster, L., Zhou, J. & Mohan, S. The ribosome moves: RNA mechanics and translocation. *Nat Struct Mol Biol* **24**, 1021-1027, doi:10.1038/nsmb.3505 (2017).

202 Noller, H. F., Lancaster, L., Mohan, S. & Zhou, J. Ribosome structural dynamics in translocation: yet another functional role for ribosomal RNA. *Q Rev Biophys* **50**, e12, doi:10.1017/S0033583517000117 (2017).

203 Kramer, E. B. & Farabaugh, P. J. The frequency of translational misreading errors in *E. coli* is largely determined by tRNA competition. *RNA* **13**, 87-96, doi:10.1261/rna.294907 (2007).

204 Kramer, E. B., Vallabhaneni, H., Mayer, L. M. & Farabaugh, P. J. A comprehensive analysis of translational missense errors in the yeast *Saccharomyces cerevisiae*. *RNA* **16**, 1797-1808, doi:10.1261/rna.2201210 (2010).

205 Ketteler, R. On programmed ribosomal frameshifting: the alternative proteomes. *Front Genet* **3**, 242, doi:10.3389/fgene.2012.00242 (2012).

206 Firth, A. E. & Brierley, I. Non-canonical translation in RNA viruses. *J Gen Virol* **93**, 1385-1409, doi:10.1099/vir.0.042499-0 (2012).

207 Li, Y. *et al.* Transactivation of programmed ribosomal frameshifting by a viral protein. *Proc Natl Acad Sci U S A* **111**, E2172-2181, doi:10.1073/pnas.1321930111 (2014).

208 Belew, A. T. *et al.* Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway. *Nature* **512**, 265-269, doi:10.1038/nature13429 (2014).

209 Dunkle, J. A. & Dunham, C. M. Mechanisms of mRNA frame maintenance and its subversion during translation of the genetic code. *Biochimie* **114**, 90-96, doi:10.1016/j.biochi.2015.02.007 (2015).

210 Hassan, M. K., Kumar, D., Naik, M. & Dixit, M. The expression profile and prognostic significance of eukaryotic translation elongation factors in different cancers. *PLoS One* **13**, e0191377, doi:10.1371/journal.pone.0191377 (2018).

211 Goodarzi, H. *et al.* Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. *Cell* **165**, 1416-1427, doi:10.1016/j.cell.2016.05.046 (2016).

212 Blanchet, S. *et al.* Deciphering the reading of the genetic code by near-cognate tRNA. *Proc Natl Acad Sci U S A* **115**, 3018-3023, doi:10.1073/pnas.1715578115 (2018).

213 Namy, O., Hatin, I. & Rousset, J. P. Impact of the six nucleotides downstream of the stop codon on translation termination. *EMBO Rep* **2**, 787-793, doi:10.1093/embo-reports/kve176 (2001).

214 Schueren, F. & Thoms, S. Functional Translational Readthrough: A Systems Biology Perspective. *PLoS Genet* **12**, e1006196, doi:10.1371/journal.pgen.1006196 (2016).

215 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* **42**, 10061-10072, doi:10.1093/nar/gku663 (2014).

216 Brown, C. M., Stockwell, P. A., Trotman, C. N. & Tate, W. P. Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes. *Nucleic Acids Res* **18**, 6339-6345 (1990).

217 Loughran, G. *et al.* Evidence of efficient stop codon readthrough in four mammalian genes. *Nucleic Acids Res* **42**, 8928-8938, doi:10.1093/nar/gku608 (2014).

218 Tork, S., Hatin, I., Rousset, J. P. & Fabret, C. The major 5' determinant in stop codon readthrough involves two adjacent adenines. *Nucleic Acids Res* **32**, 415-421, doi:10.1093/nar/gkh201 (2004).

219 Dabrowski, M., Bukowy-Bierylo, Z. & Zietkiewicz, E. Translational readthrough potential of natural termination codons in eucaryotes--The impact of RNA sequence. *RNA Biol* **12**, 950-958, doi:10.1080/15476286.2015.1068497 (2015).

220 Karijolich, J. & Yu, Y. T. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* **474**, 395-398, doi:10.1038/nature10165 (2011).

221 Prokhorova, I. *et al.* Aminoglycoside interactions and impacts on the eukaryotic ribosome. *Proc Natl Acad Sci U S A* **114**, E10899-E10908, doi:10.1073/pnas.1715501114 (2017).

222 Bulteau, A. L. & Chavatte, L. Update on selenoprotein biosynthesis. *Antioxid Redox Signal* **23**, 775-794, doi:10.1089/ars.2015.6391 (2015).

223 Bidou, L., Allamand, V., Rousset, J. P. & Namy, O. Sense from nonsense: therapies for premature stop codon diseases. *Trends Mol Med* **18**, 679-688, doi:10.1016/j.molmed.2012.09.008 (2012).

224 Gaston, M. A., Jiang, R. & Krzycki, J. A. Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Curr Opin Microbiol* **14**, 342-349, doi:10.1016/j.mib.2011.04.001 (2011).

225 Théobald-Dietrich, A., Giegé, R. & Rudinger-Thirion, J. Evidence for the existence in mRNAs of a hairpin element responsible for ribosome dependent pyrrolysine insertion into proteins. *Biochimie* **87**, 813-817, doi:10.1016/j.biochi.2005.03.006 (2005).

226 Hay, N. & Sonenberg, N. Upstream and downstream of mTOR. *Genes Dev* **18**, 1926-1945, doi:10.1101/gad.1212704 (2004).

227 Martin, K. A. & Blenis, J. Coordinate regulation of translation by the PI 3-kinase and mTOR pathways. *Adv Cancer Res* **86**, 1-39 (2002).

228 Jastrzebski, K., Hannan, K. M., Tchoubrieva, E. B., Hannan, R. D. & Pearson, R. B. Coordinate regulation of ribosome biogenesis and function by the ribosomal protein S6 kinase, a key

mediator of mTOR function. *Growth Factors* **25**, 209-226, doi:10.1080/08977190701779101 (2007).

229 Rogers, G. W., Komar, A. A. & Merrick, W. C. eIF4A: the godfather of the DEAD box helicases. *Prog Nucleic Acid Res Mol Biol* **72**, 307-331 (2002).

230 Shahbazian, D. *et al.* The mTOR/PI3K and MAPK pathways converge on eIF4B to control its phosphorylation and activity. *EMBO J* **25**, 2781-2791, doi:10.1038/sj.emboj.7601166 (2006).

231 Yang, H. S. *et al.* The transformation suppressor Pdcd4 is a novel eukaryotic translation initiation factor 4A binding protein that inhibits translation. *Mol Cell Biol* **23**, 26-37 (2003).

232 Yang, H. S. *et al.* A novel function of the MA-3 domains in transformation and translation suppressor Pdcd4 is essential for its binding to eukaryotic translation initiation factor 4A. *Mol Cell Biol* **24**, 3894-3906 (2004).

233 Dorrello, N. V. *et al.* S6K1- and betaTRCP-mediated degradation of PDCD4 promotes protein translation and cell growth. *Science* **314**, 467-471, doi:10.1126/science.1130276 (2006).

234 Polymenis, M. & Schmidt, E. V. Coupling of cell division to cell growth by translational control of the G1 cyclin CLN3 in yeast. *Genes Dev* **11**, 2522-2531 (1997).

235 Kronja, I. & Orr-Weaver, T. L. Translational regulation of the cell cycle: when, where, how and why? *Philos Trans R Soc Lond B Biol Sci* **366**, 3638-3652, doi:10.1098/rstb.2011.0084 (2011).

236 Lai, M. C., Chang, W. C., Shieh, S. Y. & Tarn, W. Y. DDX3 regulates cell growth through translational control of cyclin E1. *Mol Cell Biol* **30**, 5444-5453, doi:10.1128/MCB.00560-10 (2010).

237 Lai, M. C., Lee, Y. H. & Tarn, W. Y. The DEAD-box RNA helicase DDX3 associates with export messenger ribonucleoproteins as well as tip-associated protein and participates in translational control. *Mol Biol Cell* **19**, 3847-3858, doi:10.1091/mbc.E07-12-1264 (2008).

238 Daga, R. R. & Jimenez, J. Translational control of the cdc25 cell cycle phosphatase: a molecular mechanism coupling mitosis to cell growth. *J Cell Sci* **112 Pt 18**, 3137-3146 (1999).

239 Yang, X. & Xu, T. Molecular mechanism of size control in development and human diseases. *Cell Res* **21**, 715-729, doi:10.1038/cr.2011.63 (2011).

240 Pyronnet, S. & Sonenberg, N. Cell-cycle-dependent translational control. *Curr Opin Genet Dev* **11**, 13-18 (2001).

241 Lykke-Andersen, J. & Bennett, E. J. Protecting the proteome: Eukaryotic cotranslational quality control pathways. *J Cell Biol* **204**, 467-476, doi:10.1083/jcb.201311103 (2014).

242 Buskirk, A. R. & Green, R. Ribosome pausing, arrest and rescue in bacteria and eukaryotes. *Philos Trans R Soc Lond B Biol Sci* **372**, doi:10.1098/rstb.2016.0183 (2017).

243 Shoemaker, C. J., Eyler, D. E. & Green, R. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science* **330**, 369-372, doi:10.1126/science.1192430 (2010).

244 Defenouillère, Q. & Fromont-Racine, M. The ribosome-bound quality control complex: from aberrant peptide clearance to proteostasis maintenance. *Curr Genet* **63**, 997-1005, doi:10.1007/s00294-017-0708-5 (2017).

245 Defenouillère, Q. *et al.* Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. *Proc Natl Acad Sci U S A* **110**, 5046-5051, doi:10.1073/pnas.1221724110 (2013).

246 Tsuboi, T. *et al.* Dom34:hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. *Mol Cell* **46**, 518-529, doi:10.1016/j.molcel.2012.03.013 (2012).

247 van Hoof, A., Frischmeyer, P. A., Dietz, H. C. & Parker, R. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**, 2262-2264, doi:10.1126/science.1067272 (2002).

248 Klauer, A. A. & van Hoof, A. Degradation of mRNAs that lack a stop codon: a decade of nonstop progress. *Wiley Interdiscip Rev RNA* **3**, 649-660, doi:10.1002/wrna.1124 (2012).

249 Kowalinski, E., Schuller, A., Green, R. & Conti, E. *Saccharomyces cerevisiae* Ski7 Is a GTP-Binding Protein Adopting the Characteristic Conformation of Active Translational GTPases. *Structure* **23**, 1336-1343, doi:10.1016/j.str.2015.04.018 (2015).

250 Losson, R. & Lacroute, F. Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proc Natl Acad Sci U S A* **76**, 5134-5137 (1979).

251 Hug, N., Longman, D. & Cáceres, J. F. Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res* **44**, 1483-1495, doi:10.1093/nar/gkw010 (2016).

252 Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V. & Izaurralde, E. A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *EMBO J* **26**, 1591-1601, doi:10.1038/sj.emboj.7601588 (2007).

253 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* **6**, e111, doi:10.1371/journal.pbio.0060111 (2008).

254 Barbosa, C., Peixeiro, I. & Romão, L. Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet* **9**, e1003529, doi:10.1371/journal.pgen.1003529 (2013).

255 Wollerton, M. C., Gooding, C., Wagner, E. J., Garcia-Blanco, M. A. & Smith, C. W. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol Cell* **13**, 91-100 (2004).

256 Ni, J. Z. *et al.* Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev* **21**, 708-718, doi:10.1101/gad.1525507 (2007).

257 Belew, A. T., Advani, V. M. & Dinman, J. D. Endogenous ribosomal frameshift signals operate as mRNA destabilizing elements through at least two molecular pathways in yeast. *Nucleic Acids Res* **39**, 2799-2808, doi:10.1093/nar/gkq1220 (2011).

258 Hogg, J. R. & Goff, S. P. Upf1 senses 3'UTR length to potentiate mRNA decay. *Cell* **143**, 379-389, doi:10.1016/j.cell.2010.10.005 (2010).

259 Eberle, A. B., Stalder, L., Mathys, H., Orozco, R. Z. & Mühlemann, O. Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. *PLoS Biol* **6**, e92, doi:10.1371/journal.pbio.0060092 (2008).

260 Bühl, M., Steiner, S., Mohn, F., Paillusson, A. & Mühlemann, O. EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. *Nat Struct Mol Biol* **13**, 462-464, doi:10.1038/nsmb1081 (2006).

261 Seyedali, A. & Berry, M. J. Nonsense-mediated decay factors are involved in the regulation of selenoprotein mRNA levels during selenium deficiency. *RNA* **20**, 1248-1256, doi:10.1261/rna.043463.113 (2014).

262 Celik, A., Baker, R., He, F. & Jacobson, A. High-resolution profiling of NMD targets in yeast reveals translational fidelity as a basis for substrate selection. *RNA* **23**, 735-748, doi:10.1261/rna.060541.116 (2017).

263 Medghalchi, S. M. *et al.* Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum Mol Genet* **10**, 99-105 (2001).

264 McIlwain, D. R. *et al.* Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. *Proc Natl Acad Sci U S A* **107**, 12186-12191, doi:10.1073/pnas.1007336107 (2010).

265 Nickless, A., Bailis, J. M. & You, Z. Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell Biosci* **7**, 26, doi:10.1186/s13578-017-0153-7 (2017).

266 Slavov, N., Semrau, S., Airoldi, E., Budnik, B. & van Oudenaarden, A. Differential Stoichiometry among Core Ribosomal Proteins. *Cell Rep* **13**, 865-873, doi:10.1016/j.celrep.2015.09.056 (2015).

267 Mauro, V. P. & Matsuda, D. Translation regulation by ribosomes: Increased complexity and expanded scope. *RNA Biol* **13**, 748-755, doi:10.1080/15476286.2015.1107701 (2016).

268 Sauert, M., Temmel, H. & Moll, I. Heterogeneity of the translational machinery: Variations on a common theme. *Biochimie* **114**, 39-47, doi:10.1016/j.biuchi.2014.12.011 (2015).

269 Deusser, E. & Wittmann, H. G. Ribosomal proteins: variation of the protein composition in *Escherichia coli* ribosomes as function of growth rate. *Nature* **238**, 269-270 (1972).

270 Ramagopal, S. & Ennis, H. L. Regulation of synthesis of cell-specific ribosomal proteins during differentiation of *Dictyostelium discoideum*. *Proc Natl Acad Sci U S A* **78**, 3083-3087 (1981).

271 Bortoluzzi, S., d'Alessi, F., Romualdi, C. & Danieli, G. A. Differential expression of genes coding for ribosomal proteins in different human tissues. *Bioinformatics* **17**, 1152-1157 (2001).

272 Shi, Z. *et al.* Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide. *Mol Cell* **67**, 71-83.e77, doi:10.1016/j.molcel.2017.05.021 (2017).

273 Kellis, M., Birren, B. W. & Lander, E. S. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**, 617-624, doi:10.1038/nature02424 (2004).

274 Cárdenas, D. *et al.* P1 and P2 protein heterodimer binding to the P0 protein of *Saccharomyces cerevisiae* is relatively non-specific and a source of ribosomal heterogeneity. *Nucleic Acids Res* **40**, 4520-4529, doi:10.1093/nar/gks036 (2012).

275 Barakat, A. *et al.* The organization of cytoplasmic ribosomal protein genes in the *Arabidopsis* genome. *Plant Physiol* **127**, 398-415 (2001).

276 Zsögön, A., Szakonyi, D., Shi, X. & Byrne, M. E. Ribosomal Protein RPL27a Promotes Female Gametophyte Development in a Dose-Dependent Manner. *Plant Physiol* **165**, 1133-1143, doi:10.1104/pp.114.241778 (2014).

277 Weijers, D. *et al.* An *Arabidopsis* Minute-like phenotype caused by a semi-dominant mutation in a RIBOSOMAL PROTEIN S5 gene. *Development* **128**, 4289-4299 (2001).

278 Marygold, S. J. *et al.* The ribosomal protein genes and Minute loci of *Drosophila melanogaster*. *Genome Biol* **8**, R216, doi:10.1186/gb-2007-8-10-r216 (2007).

279 Lopes, A. M. *et al.* The human RPS4 parologue on Yq11.223 encodes a structurally conserved ribosomal protein and is preferentially expressed during spermatogenesis. *BMC Mol Biol* **11**, 33, doi:10.1186/1471-2199-11-33 (2010).

280 Wong, Q. W. *et al.* RPL39L is an example of a recently evolved ribosomal protein paralog that shows highly specific tissue expression patterns and is upregulated in ESCs and HCC tumors. *RNA Biol* **11**, 33-41, doi:10.4161/rna.27427 (2014).

281 Sugihara, Y. *et al.* Proteomic analysis of rodent ribosomes revealed heterogeneity including ribosomal proteins L10-like, L22-like 1, and L39-like. *J Proteome Res* **9**, 1351-1366, doi:10.1021/pr9008964 (2010).

282 Mills, A. A., Mills, M. J., Gardiner, D. M., Bryant, S. V. & Stanbridge, E. J. Analysis of the pattern of QM expression during mouse development. *Differentiation* **64**, 161-171, doi:10.1046/j.1432-0436.1999.6430161.x (1999).

283 Green, H. *et al.* The ribosomal protein QM is expressed differentially during vertebrate endochondral bone development. *J Bone Miner Res* **15**, 1066-1075, doi:10.1359/jbmr.2000.15.6.1066 (2000).

284 Lee, S. W. *et al.* Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Proc Natl Acad Sci U S A* **99**, 5942-5947, doi:10.1073/pnas.082119899 (2002).

285 Lipson, R. S., Webb, K. J. & Clarke, S. G. Rmt1 catalyzes zinc-finger independent arginine methylation of ribosomal protein Rps2 in *Saccharomyces cerevisiae*. *Biochem Biophys Res Commun* **391**, 1658-1662, doi:10.1016/j.bbrc.2009.12.112 (2010).

286 Loenarz, C. *et al.* Hydroxylation of the eukaryotic ribosomal decoding center affects translational accuracy. *Proc Natl Acad Sci U S A* **111**, 4019-4024, doi:10.1073/pnas.1311750111 (2014).

287 Spence, J. *et al.* Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* **102**, 67-76 (2000).

288 Zeidan, Q., Wang, Z., De Maio, A. & Hart, G. W. O-GlcNAc cycling enzymes associate with the translational machinery and modify core ribosomal proteins. *Mol Biol Cell* **21**, 1922-1936, doi:10.1091/mbc.E09-11-0941 (2010).

289 Juszkiewicz, S. & Hegde, R. S. Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol Cell* **65**, 743-750.e744, doi:10.1016/j.molcel.2016.11.039 (2017).

290 Sundaramoorthy, E. *et al.* ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. *Mol Cell* **65**, 751-760.e754, doi:10.1016/j.molcel.2016.12.026 (2017).

291 Gonzalez, I. L. *et al.* Variation among human 28S ribosomal RNA genes. *Proc Natl Acad Sci U S A* **82**, 7666-7670 (1985).

292 Gonzalez, I. L., Sylvester, J. E. & Schmickel, R. D. Human 28S ribosomal RNA sequence heterogeneity. *Nucleic Acids Res* **16**, 10213-10224 (1988).

293 Parks, M. M. *et al.* Variant ribosomal RNA alleles are conserved and exhibit tissue-specific expression. *Sci Adv* **4**, eaao0665, doi:10.1126/sciadv.aao0665 (2018).

294 Kim, H. L., Song, W. S., Kim, K. & Lee, K. Characterization of heterogeneous LSU rRNA profiles in *Streptomyces coelicolor* under different growth stages and conditions. *Curr Microbiol* **57**, 537-541, doi:10.1007/s00284-008-9238-1 (2008).

295 Marchand, V., Blanloel-Oillo, F., Helm, M. & Motorin, Y. Illumina-based RiboMethSeq approach for mapping of 2'-O-Me residues in RNA. *Nucleic Acids Res* **44**, e135, doi:10.1093/nar/gkw547 (2016).

296 Yano, K. *et al.* Multiple rRNA operons are essential for efficient cell growth and sporulation as well as outgrowth in *Bacillus subtilis*. *Microbiology* **159**, 2225-2236, doi:10.1099/mic.0.067025-0 (2013).

297 van Spaendonk, R. M. *et al.* Functional equivalence of structurally distinct ribosomes in the malaria parasite, *Plasmodium berghei*. *J Biol Chem* **276**, 22638-22647, doi:10.1074/jbc.M101234200 (2001).

298 Taoka, M. *et al.* The complete chemical structure of *Saccharomyces cerevisiae* rRNA: partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. *Nucleic Acids Res* **44**, 8951-8961, doi:10.1093/nar/gkw564 (2016).

299 Castle, J. C. *et al.* Digital genome-wide ncRNA expression, including SnoRNAs, across 11 human tissues using polyA-neutral amplification. *PLoS One* **5**, e11779, doi:10.1371/journal.pone.0011779 (2010).

300 Nikolov, E. N., Dabeva, M. D. & Nikolov, T. K. Turnover of ribosomes in regenerating rat liver. *Int J Biochem* **15**, 1255-1260 (1983).

301 Birkedal, U. *et al.* Profiling of ribose methylations in RNA by high-throughput sequencing. *Angew Chem Int Ed Engl* **54**, 451-455, doi:10.1002/anie.201408362 (2015).

302 Erales, J. *et al.* Evidence for rRNA 2'-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes. *Proc Natl Acad Sci U S A* **114**, 12934-12939, doi:10.1073/pnas.1707674114 (2017).

303 Zheng, G. *et al.* ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell* **49**, 18-29, doi:10.1016/j.molcel.2012.10.015 (2013).

304 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* **13**, 30-34, doi:10.1038/nsmb1031 (2006).

305 Clark, M. B. *et al.* Mammalian gene PEG10 expresses two reading frames by high efficiency -1 frameshifting in embryonic-associated tissues. *J Biol Chem* **282**, 37359-37369, doi:10.1074/jbc.M705676200 (2007).

306 Saulquin, X. *et al.* +1 Frameshifting as a novel mechanism to generate a cryptic cytotoxic T lymphocyte epitope derived from human interleukin 10. *J Exp Med* **195**, 353-358 (2002).

307 Nicco, C. & Batteux, F. ROS Modulator Molecules with Therapeutic Potential in Cancers Treatments. *Molecules* **23**, doi:10.3390/molecules23010084 (2017).

308 Hasmim, M., Messai, Y., Noman, M. Z. & Chouaib, S. [Tumor hypoxia: a key player in the regulation of stromal and anti-tumor responses]. *Med Sci (Paris)* **30**, 422-428, doi:10.1051/medsci/20143004017 (2014).

309 Yan, W. *et al.* Transcriptional analysis of immune-related gene expression in p53-deficient mice with increased susceptibility to influenza A virus infection. *BMC Med Genomics* **8**, 52, doi:10.1186/s12920-015-0127-8 (2015).

310 Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. & Nomoto, A. Internal ribosome entry site within hepatitis C virus RNA. *J Virol* **66**, 1476-1483 (1992).

311 Herold, J. & Siddell, S. G. An 'elaborated' pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic Acids Res* **21**, 5838-5842 (1993).

312 Kusumaningtyas, E., Tan, W. S., Zamrod, Z., Eshaghi, M. & Yusoff, K. Existence of two forms of L protein of Newcastle disease virus isolates due to a compensatory mutation in Domain V. *Arch Virol* **149**, 1859-1865 (2004).

313 Jagger, B. W. *et al.* An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* **337**, 199-204, doi:10.1126/science.1222213 (2012).

314 Huang, M. T. Harringtonine, an inhibitor of initiation of protein biosynthesis. *Mol Pharmacol* **11**, 511-519 (1975).

315 Schneider-Poetsch, T. *et al.* Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat Chem Biol* **6**, 209-217, doi:10.1038/nchembio.304 (2010).

316 Lee, S., Liu, B., Huang, S. X., Shen, B. & Qian, S. B. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc Natl Acad Sci U S A* **109**, E2424-2432, doi:10.1073/pnas.1207846109 (2012).

317 Gao, X. *et al.* Quantitative profiling of initiating ribosomes in vivo. *Nat Methods* **12**, 147-153, doi:10.1038/nmeth.3208 (2015).

318 Solis-Escalante, D. *et al.* Efficient simultaneous excision of multiple selectable marker cassettes using I-SceI-induced double-strand DNA breaks in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **14**, 741-754, doi:10.1111/1567-1364.12162 (2014).

319 Li, Z. *et al.* Rational extension of the ribosome biogenesis pathway using network-guided genetics. *PLoS Biol* **7**, e1000213, doi:10.1371/journal.pbio.1000213 (2009).

320 Le, Q. G., Ishiwata-Kimata, Y., Kohno, K. & Kimata, Y. Cadmium impairs protein folding in the endoplasmic reticulum and induces the unfolded protein response. *FEMS Yeast Res* **16**, doi:10.1093/femsyr/fow049 (2016).

321 Korennykh, A. V. *et al.* The unfolded protein response signals through high-order assembly of Ire1. *Nature* **457**, 687-693, doi:10.1038/nature07661 (2009).

322 Rüegsegger, U., Leber, J. H. & Walter, P. Block of HAC1 mRNA translation by long-range base pairing is released by cytoplasmic splicing upon induction of the unfolded protein response. *Cell* **107**, 103-114 (2001).

323 Cox, J. S. & Walter, P. A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response. *Cell* **87**, 391-404 (1996).

324 Waldo, G. S., Standish, B. M., Berendzen, J. & Terwilliger, T. C. Rapid protein-folding assay using green fluorescent protein. *Nat Biotechnol* **17**, 691-695, doi:10.1038/10904 (1999).

325 Caine, J. *et al.* Alzheimer's Abeta fused to green fluorescent protein induces growth stress and a heat shock response. *FEMS Yeast Res* **7**, 1230-1236, doi:10.1111/j.1567-1364.2007.00285.x (2007).

326 Gregoire, S. & Kwon, I. A revisited folding reporter for quantitative assay of protein misfolding and aggregation in mammalian cells. *Biotechnol J* **7**, 1297-1307, doi:10.1002/biot.201200103 (2012).

327 Gerbasi, V. R., Weaver, C. M., Hill, S., Friedman, D. B. & Link, A. J. Yeast Asc1p and mammalian RACK1 are functionally orthologous core 40S ribosomal proteins that repress gene expression. *Mol Cell Biol* **24**, 8276-8287, doi:10.1128/MCB.24.18.8276-8287.2004 (2004).

328 Sengupta, J. *et al.* Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM. *Nat Struct Mol Biol* **11**, 957-962, doi:10.1038/nsmb822 (2004).

329 Coyle, S. M., Gilbert, W. V. & Doudna, J. A. Direct link between RACK1 function and localization at the ribosome *in vivo*. *Mol Cell Biol* **29**, 1626-1634, doi:10.1128/MCB.01718-08 (2009).

330 Thompson, M. K., Rojas-Duran, M. F., Gangaramani, P. & Gilbert, W. V. The ribosomal protein Asc1/RACK1 is required for efficient translation of short mRNAs. *Elife* **5**, doi:10.7554/elife.11154 (2016).

331 Kuroha, K. *et al.* Receptor for activated C kinase 1 stimulates nascent polypeptide-dependent translation arrest. *EMBO Rep* **11**, 956-961, doi:10.1038/embor.2010.169 (2010).

332 Chiabudini, M. *et al.* Release factor eRF3 mediates premature translation termination on polylysine-stalled ribosomes in *Saccharomyces cerevisiae*. *Mol Cell Biol* **34**, 4062-4076, doi:10.1128/MCB.00799-14 (2014).

333 Letzring, D. P., Wolf, A. S., Brule, C. E. & Grayhack, E. J. Translation of CGA codon repeats in yeast involves quality control components and ribosomal protein L1. *RNA* **19**, 1208-1217, doi:10.1261/rna.039446.113 (2013).

334 Wolf, A. S. & Grayhack, E. J. Asc1, homolog of human RACK1, prevents frameshifting in yeast by ribosomes stalled at CGA codon repeats. *RNA* **21**, 935-945, doi:10.1261/rna.049080.114 (2015).

335 Sitron, C. S., Park, J. H. & Brandman, O. Asc1, Hel2, and Slh1 couple translation arrest to nascent chain degradation. *RNA* **23**, 798-810, doi:10.1261/rna.060897.117 (2017).

336 Ikeuchi, K. & Inada, T. Ribosome-associated Asc1/RACK1 is required for endonucleolytic cleavage induced by stalled ribosome at the 3' end of nonstop mRNA. *Sci Rep* **6**, 28234, doi:10.1038/srep28234 (2016).

337 Limoncelli, K. A., Merrikh, C. N. & Moore, M. J. and. *RNA* **23**, 1946-1960, doi:10.1261/rna.061671.117 (2017).

338 Limoncelli, K. A., Merrikh, C. N. & Moore, M. J. Corrigendum:. *RNA* **24**, 620, doi:10.1261/rna.066159.118 (2018).

339 Rachfall, N. *et al.* RACK1/Asc1p, a ribosomal node in cellular signaling. *Mol Cell Proteomics* **12**, 87-105, doi:10.1074/mcp.M112.017277 (2013).

340 Kourba, T., Rutkai, E., Karásková, M. & Valášek, L. The eIF3c/NIP1 PCI domain interacts with RNA and RACK1/ASC1 and promotes assembly of translation preinitiation complexes. *Nucleic Acids Res* **40**, 2683-2699, doi:10.1093/nar/gkr1083 (2012).

341 Schmitt, K. *et al.* Asc1p/RACK1 Connects Ribosomes to Eukaryotic Phosphosignaling. *Mol Cell Biol* **37**, doi:10.1128/MCB.00279-16 (2017).

342 Opitz, N. *et al.* Capturing the Asc1p/. *Mol Cell Proteomics* **16**, 2199-2218, doi:10.1074/mcp.M116.066654 (2017).

343 Li, J. J. & Xie, D. RACK1, a versatile hub in cancer. *Oncogene* **34**, 1890-1898, doi:10.1038/onc.2014.127 (2015).

344 Volta, V. *et al.* RACK1 depletion in a mouse model causes lethality, pigmentation deficits and reduction in protein synthesis efficiency. *Cell Mol Life Sci* **70**, 1439-1450, doi:10.1007/s00018-012-1215-y (2013).

345 Valerius, O. *et al.* The *Saccharomyces* homolog of mammalian RACK1, Cpc2/Asc1p, is required for FLO11-dependent adhesive growth and dimorphism. *Mol Cell Proteomics* **6**, 1968-1979, doi:10.1074/mcp.M700184-MCP200 (2007).

346 Majzoub, K. *et al.* RACK1 controls IRES-mediated translation of viruses. *Cell* **159**, 1086-1095, doi:10.1016/j.cell.2014.10.041 (2014).

347 Valle, R. P., Drugeon, G., Devignes-Morch, M. D., Legocki, A. B. & Haenni, A. L. Codon context effect in virus translational readthrough. A study *in vitro* of the determinants of TMV and Mo-MuLV amber suppression. *FEBS Lett* **306**, 133-139 (1992).

348 Hilal, T. *et al.* Structural insights into ribosomal rescue by Dom34 and Hbs1 at near-atomic resolution. *Nat Commun* **7**, 13521, doi:10.1038/ncomms13521 (2016).

349 Ben-Shem, A., Jenner, L., Yusupova, G. & Yusupov, M. Crystal structure of the eukaryotic ribosome. *Science* **330**, 1203-1209, doi:10.1126/science.1194294 (2010).

350 Taylor, D. *et al.* Cryo-EM structure of the mammalian eukaryotic release factor eRF1-eRF3-associated termination complex. *Proc Natl Acad Sci U S A* **109**, 18413-18418, doi:10.1073/pnas.1216730109 (2012).

351 Beznosková, P. *et al.* Translation initiation factors eIF3 and HCR1 control translation termination and stop codon read-through in yeast cells. *PLoS Genet* **9**, e1003962, doi:10.1371/journal.pgen.1003962 (2013).

352 Chiu, W. L. *et al.* The C-terminal region of eukaryotic translation initiation factor 3a (eIF3a) promotes mRNA recruitment, scanning, and, together with eIF3j and the eIF3b RNA recognition motif, selection of AUG start codons. *Mol Cell Biol* **30**, 4415-4434, doi:10.1128/MCB.00280-10 (2010).

353 Cuchalová, L. *et al.* The RNA recognition motif of eukaryotic translation initiation factor 3g (eIF3g) is required for resumption of scanning of posttermination ribosomes for reinitiation on GCN4 and together with eIF3i stimulates linear scanning. *Mol Cell Biol* **30**, 4671-4686, doi:10.1128/MCB.00430-10 (2010).

354 Matsuo, Y. *et al.* Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nat Commun* **8**, 159, doi:10.1038/s41467-017-00188-1 (2017).

355 Giraud, S., Greco, A., Brink, M., Diaz, J. J. & Delafontaine, P. Translation initiation of the insulin-like growth factor I receptor mRNA is mediated by an internal ribosome entry site. *J Biol Chem* **276**, 5668-5675, doi:10.1074/jbc.M005928200 (2001).

Titre : Impact de la composition du ribosome sur la fidélité de la traduction

Mots clés : Ribosome, 2'-O-méthylation, ASC1, Traduction

Résumé : Les ribosomes, acteurs principaux de la synthèse protéique, sont constitués de protéines et d'ARNs. Ces dernières années la notion de "ribosome spécialisé" est apparue. Cela implique que les ribosomes sont hétérogènes dans leur composition protéique et les modifications chimiques des ARNr. Ces différentes populations de ribosome présenteraient des spécificités de traduction différentes. Au cours de ma thèse, je me suis intéressée à une modification chimique des ARNr particulière, les 2'-O-méthylation des riboses, à leur variabilité et à leur rôle dans la fidélité et la régulation de la traduction.

Pour réaliser cette étude, un modèle de cellules HeLa a été créé dans lequel la synthèse de la fibrillarine, la méthyltransférase responsable des 2'-O-méthylation, est inhibée par un shRNA (small hairpin RNA) intégré de façon stable dans le génome. L'étude de l'impact de la baisse de la fibrillarine sur les 2'-O-méthylation a permis de montrer que la diminution des méthylations des ARNr est globale mais varie selon la position. Ainsi, certaines positions sont plus sensibles que d'autres à la baisse du taux de fibrillarine.

J'ai tout d'abord étudié les effets de la baisse de la méthylation des ARNr sur la traduction globale, par la technique de ribosome profiling. Cette technique est fondée sur le séquençage à haut débit des fragments d'ARNm protégés par les ribosomes. J'ai ainsi pu montrer que 43 gènes candidats étaient différemment traduits en condition d'hypométhylation des ARNr. A partir de cette liste j'ai cherché des éléments fonctionnels et/ou moléculaires communs à plusieurs gènes candidats. J'ai par la suite montré que les taux de translecture et de décalage de cadre de lecture augmentaient quand les ARNr sont hypométhylés. La baisse de la méthylation des ARNr entraîne donc une baisse de la fidélité de la traduction.

Des études précédentes ont montré que l'initiation IRES-dépendante était impactée par la baisse des méthylations des ARNr. J'ai donc réalisé une étude globale sur l'initiation de la traduction en adaptant la technique de ribosome profiling de façon à identifier spécifiquement les ribosomes en cours d'initiation. J'ai ainsi révélé que 66 sites d'initiation étaient impactés par la baisse de la méthylation des ARNr.

Nous avons localisé les positions méthylées les plus impactées sur la structure 3D du ribosome. Ceci nous a permis de regrouper les modifications par région. Nous nous sommes intéressés à un groupe de méthylations conservées entre la levure et l'homme et situées au niveau du tunnel de sortie du peptide. J'ai déléte chez *S. cerevisiae* les snoARNs responsables de ces méthylations. J'ai ensuite cherché à démontrer si la perte de ces méthylations avait un impact sur la croissance cellulaire et la sensibilité à différents antibiotiques. J'ai aussi effectué des mesures de translecture et de décalage de cadre de lecture. L'ensemble de mes résultats a montré que la délétion conjointe de trois des quatre snoARNs impliqués dans les méthylations autour du tunnel de sortie du polypeptide n'a pas d'effet sur la fidélité de la traduction.

Au cours de cette étude chez la levure, j'ai révélé un effet inédit de la délétion du gène *ASC1* sur la translecture des codons stop. Asc1p est une protéine plateforme associée au ribosome, dont l'absence entraîne une diminution de la translecture du codon stop. Les mécanismes moléculaires impliqués demeurent actuellement inconnus.

Au cours de ma thèse, j'ai pu montrer par des approches globales et spécifiques que la baisse de



la méthylation des ARNr entraînait des variations spécifiques de l'expression protéique ainsi qu'une diminution spécifique de la fidélité de la traduction. Les mécanismes moléculaires impliqués sont toujours activement recherchés.

Title : Impact of ribosome composition onto translational fidelity

Keywords : Ribosome, 2'-O-methylation, ASC1, Translation

Abstract : Ribosomes are composed of proteins and RNAs. During these last years, the concept of « specialized ribosome » has been revived. This concept is based on the principle that ribosomes are heterogeneous in protein composition and rRNA chemical modifications. These different ribosomes populations would present different translational specificities. During my thesis, I was interested in a particular rRNA chemical modification, ribose 2'-O-methylation, its variability and its role in translational fidelity and regulation.

To make this study, a HeLa cell-line in which fibrillarin (the methyltransferase responsible for 2'-O-methylations) synthesis is inhibited by a shRNA (small hairpin RNA) stably integrated in the genome. The study of impact of fibrillarin decrease on 2'-O-methylations enabled us to show that rRNA methylation decrease is global but varies with the position. So, some positions are more sensitive than others positions to fibrillarin decrease.

First I studied rRNA methylation decrease effects on global translation, by ribosome profiling. This method is based on high-throughput sequencing of ribosome-protected mRNA fragments. By this way I revealed 43 candidate genes that are differentially translated in rRNA hypomethylated condition. From this list I searched functional and/or molecular elements common to several candidate genes. Then I showed that readthrough and frameshifting rates increase when rRNA is hypomethylated. So rRNA methylation decrease leads to translational fidelity decrease.

Previous studies have shown that IRES-dependant initiation is impacted by rRNA methylation decrease. Then I performed a global study of translation initiation by adapting ribosome profiling method to identify initiating ribosomes specifically. Therefore I revealed that 66 initiation sites are impacted by rRNA methylation decrease.

We localized the most impacted methylated positions on the 3D ribosome structure. It enabled us to group modifications by region. We focused our interest on one group of methylations conserved between yeast and human and localized around the peptide exit tunnel. I deleted snoRNAs responsible for these methylations in *S. cerevisiae*. Then I analysed if the loss of these methylations impacts the cell growth and the antibiotics sensitivity. I also make measures of readthrough and frameshifting. My results show that the targeted deletion of three out of four snoRNAs involved in the methylations around the peptide exit tunnel has no effect on translational fidelity.

During this study in yeast, I revealed an unprecedented effect of *ASC1* gene deletion on stop codon readthrough. Asc1p is a scaffold protein associated with the ribosome, whose absence causes a decrease of codon stop readthrough. Currently molecular mechanisms implicated remain unknown.

During my thesis, I showed with global and specific approaches that rRNA methylation decrease leads to specific variations of protein expression together with specific decrease of translational fidelity. Molecular mechanisms are still actively investigated.

