



**HAL**  
open science

## Suppression traductionnelle des codons stop chez les mammifères

Olivier Bugaud

► **To cite this version:**

Olivier Bugaud. Suppression traductionnelle des codons stop chez les mammifères. Biologie cellulaire. Université Paris Saclay (COMUE), 2016. Français. NNT : 2016SACLS222 . tel-02945040

**HAL Id: tel-02945040**

**<https://theses.hal.science/tel-02945040v1>**

Submitted on 22 Sep 2020

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

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

NNT : 2016SACLS222

THESE DE DOCTORAT  
DE  
L'UNIVERSITE PARIS-SACLAY  
PREPAREE A  
L'UNIVERSITE PARIS-SUD

ECOLE DOCTORALE N°577  
Structure et dynamique des Systèmes Vivants

Spécialité de doctorat : Sciences de la Vie et de la Santé

Dans l'équipe : Génomique, Structure & Traduction

Par

**Mr Olivier BUGAUD**

Suppression traductionnelle des codons stop  
chez les mammifères

**Thèse présentée et soutenue à Orsay, le 21 Septembre 2016 :**

**Composition du Jury :**

Mme. Monique Bolotin	Directrice de Recherche Emérite, Paris-Sud	Présidente
M. Jean-Jacques Diaz	Directeur de Recherche, CRCL	Rapporteur
M. Olivier Jean-Jean	Directeur de Recherche, CNRS	Rapporteur
Mme. Valérie Allamand	Directrice de Recherche, Inserm	Examinatrice
M. Olivier Namy	Directeur de Recherche, CNRS	Directeur de thèse
Mme. Laure Bidou	Maître de conférences, UPMC	Encadrante de thèse

## Remerciements

*Cette thèse constitue une riche expérience qui ne peut s'achever sans remercier les personnes qui m'ont encadré, aidé et soutenu au cours de ces quatre dernières années. Pourtant il existe une difficulté, celle de n'oublier personne. C'est pourquoi je tiens à remercier par avance tous ceux dont le nom n'apparaît pas dans cette page.*

*Je commencerai tout d'abord par remercier les membres du jury, les rapporteurs J-J. Diaz et O. Jean-Jean ainsi que les examinatrices V. Allamand et M. Bolotin, qui ont accepté de juger ce travail avec une spontanéité touchante. Vous avez toute ma reconnaissance pour avoir lu cet interminable manuscrit durant vos vacances et j'espère que vous aurez pris autant de plaisir à le lire que moi à l'écrire. Une mention spéciale à V. Allamand et M. Bolotin qui ont été mes tutrices de thèse et qui ont su par leur remarques et conseils faire avancer mes projets de recherche. Je tiens aussi à remercier l'école doctorale SDSV (anciennement GGC) et plus particulièrement Jean-Luc Pernodet et Sandrine Le Bihan pour leur disponibilité et leur gentillesse.*

*Un immense merci à Olivier, mon directeur de thèse et à Laure, mon encadrante qui ont guidé mes premiers pas dans le monde de la recherche en m'accueillant dans leur équipe il y a maintenant 6 ans. Vous m'aviez vendu à l'époque un sujet de rêve, et je n'ai pas été déçu... ;-) J'ai re-signé en thèse sans avoir connaissance des petites lignes louches de mon contrat doctoral dans lequel de nombreuses zones d'ombres résident encore et pour cela je ne vous en remercierai jamais assez. Laure, je n'ai jamais eu la chance de suivre tes cours, mais tu es une très bonne pédagogue et je te dois énormément, tant sur le plan technique que théorique. C'est aussi grâce à toi que j'ai eu la chance de travailler sur ces petites cellules trop mignonnes que nous avons consciencieusement maltraitées à coup de molécules en tout genre dès mon arrivée. Je n'oublierai pas non plus tous ces moments de fous rires où l'on aurait pu croire que nous avons fini par devenir les cobayes de nos propres expériences. Tu apportes une vraie joie de vivre à cette équipe et j'ai eu la chance d'être aux premières loges durant toutes ces années. Ton enthousiasme et ton soutien m'ont beaucoup apporté, tu as toujours été là lorsque j'avais besoin de toi, et ce malgré mes mille et une questions que je t'ai posées à genoux dans ton bureau. J'ai adoré discuter avec toi et débattre de tout et de rien. Ce fut un vrai plaisir de travailler à tes côtés, vivent les Bigou !! J'espère que tu ne m'en voudras pas, mais je ne trouve pas d'autres mots que ce simple, mais sincère « merci » pour t'exprimer toute ma reconnaissance et toute mon amitié. Olivier, la science te passionne et si je n'ai pas eu la chance de faire deux manip avec toi, tu m'as comblé à travers nos interminables discussions. Et même si tu as « presque » toujours raison, qu'est-ce que c'est stimulant d'échanger avec toi ! Tes connaissances encyclopédiques, ton sens logique et ta grande ouverture d'esprit m'ont toujours étonné et je te suis vraiment reconnaissant de les avoir partagés avec moi. Tu es*

*l'exemple du chercheur auquel j'aurai un jour souhaité ressembler et j'espère sincèrement pouvoir un jour arriver à ta hauteur, scientifiquement et humainement. J'ai également beaucoup apprécié ces moments de complicité, que ce soit sur le monde merveilleux d'Apple, en allant se manger de « bons » burgers, en stratéguisant pour tenter de contrer le terrifiant Dr T où nos soirées passées à repousser l'invasion Zerg. Tu es devenu bien plus que mon directeur de thèse et j'espère que cela perdurera.*

*J'aimerais également remercier toute l'équipe GST en commençant par Isa qui a su comprendre et partager avec moi de nombreux plans diaboliques. Je penserai à toi à chaque fois que je verrai les magnifiques fleurs d'amaryllis rouges qui pointeront le bout de leur nez chez nous et je suis sûr que tu sauras prendre soin de nos petits, toi qui a la main verte. Merci d'avoir toujours été là lorsque j'en ai eu besoin ainsi que pour tes précieux conseils que personne n'écoute. Tu as toujours une énergie débordante, bien plus que moi qui mange pourtant trois fois plus que toi à chaque repas, c'est louche... Cette énergie contribue largement au dynamisme de l'équipe et égaye chacune de nos journées. Un grand merci à Stéphane qui est toujours discret, mais son excessive gentillesse n'est pas passé inaperçu ;-)* Je te remercie de m'avoir initié au Northern que tu maîtrises parfaitement, comme les échecs et les énigmes d'ailleurs, et d'avoir partagé tes grandes connaissances en filmographie dont je retiendrai le mythique Rubber. Tu es plein de ressources, toujours de bonne volonté et j'ai toujours pu compter sur toi. Merci également à Agnès, le petit dragon, dont le sens de l'organisation peut parfois sembler un tout petit peu excessif, mais que l'on accepte facilement tellement tu sais te montrer chaleureuse et très maternelle. Je n'oublierai pas nos super barbecues que l'on a faits chez toi et tes mille et une raisons de faire des pots qui ont fait la joie de mon estomac. Céline, je n'ai jamais su comment tu faisais pour ne pas craquer devant toutes ces bonnes pâtisseries qui n'ont pas fait un pli au labo et encore moins pour manger du pain dur, jour après jour. Je te souhaite à toi comme à Laure et Agnès, beaucoup de courage pour l'enseignement. Un grand merci aussi aux deux derniers arrivants, Sandra et Pierre. Sandra, la dernière doctorante en date qui est également ma nouvelle voisine de bureau et qui a beaucoup de courage pour avoir tenté de prendre soin de moi en m'incitant à boire plus que 30cl d'eau par jour, une quête très difficile et qui aura porté ses fruits de temps en temps. Cette année passée à tes côtés sera passée bien trop rapidement et si je n'ai pas eu autant de temps que je le souhaitais pour échanger avec toi, je te remercie de m'avoir toujours soutenu pour manger tôt et pour nos multiples discussions sur des sujets très variés. Je te souhaite beaucoup de bonheur pour ces deux prochaines années et je suis sûr que tu feras une très bonne thèse. Pierre, je n'ai pas eu la chance de partager ton bureau, tu as préféré être près du « Chef » plutôt que des « Jeun's », un choix totalement irrationnel et un comble pour un bioinformaticien ;-) Tu as très vite su t'intégrer et ton énergie est un vrai bonheur pour ceux que tu côtoies. Je regrette également de ne pas avoir eu plus de temps pour échanger avec toi, mais j'espère que l'on aura d'autres occasions de se revoir. Je terminerai par Nelly, qui se fait toujours très discrète et qui joue pourtant un rôle prépondérant dans notre quotidien au laboratoire.

Tu es toujours prête à nous aider dès que tu le peux et c'est un vrai confort de pouvoir compter sur toi. Je n'oublierai pas non plus ton petit « bonjour » matinal en ramassant la vaisselle, accompagné d'un beau sourire, que tu ne manquerais pour rien au monde. J'ai eu beaucoup de chance d'avoir réalisé ma thèse dans d'aussi bonnes conditions, ce qui a largement contribué à mon épanouissement et vous m'avez tous énormément apporté. J'emporte avec moi cet héritage pour le faire vivre et peut-être un jour le transmettre à mon tour.

Je ne peux remercier l'équipe GST sans prendre ces quelques lignes pour remercier Jean-Pierre Rousset qui a fondé l'équipe GMT et que j'ai eu l'occasion de côtoyer aux réunions de laboratoire. Si je n'ai pas eu l'occasion de beaucoup échanger scientifiquement avec toi, je te remercie de nous avoir invités à partager ce très bon repas à Chavanon. Un très grand merci à Sandra B., ma première voisine de bureau avec qui j'ai partagé de très bons moments. Tu as apporté tellement de vie à cette équipe durant ta thèse, tu es toujours prête à rendre service et ton sourire en a émerveillé plus d'un, ce fut un vrai bonheur de passer ces 3 ans à tes côtés. Je te souhaite beaucoup de courage pour avoir emprunté ce long et rude chemin qui te permettra de devenir une très bonne chercheuse. Merci aussi à Mathieu et Rachel qui ont chacun su contribuer à la super ambiance de cette équipe. Je suis vraiment content que vous ayez tous les deux trouvé un travail malgré ce contexte pas toujours facile. Je vous souhaite à tous une très bonne continuation.

J'aimerais aussi remercier nos voisines de palier, Marie-Hélène et Magalie avec qui j'ai beaucoup aimé discuter que ce soit sur le chemin de la cantine ou dans notre aile commune. Vous m'avez toujours aidé lorsqu'il me manquait quelque chose ou lorsque j'avais besoin d'un conseil, et je vous en suis très reconnaissant.

Au cours de ma thèse, j'ai eu le privilège d'avoir collaboré avec deux formidables équipes, celle de J-C. Cintrat au CEA de Saclay et celle de N. Westbrook à l'Institut d'Optique. Vous m'avez accueilli très chaleureusement dans vos équipes respectives, vous m'y avez fait une place et je m'y suis tout de suite senti bien. Tout d'abord, j'aimerais remercier Jean-Christophe et Goulven sans qui le crible à grande échelle n'aurait pas été possible. Vous nous avez donné le temps et les moyens de mettre au point notre système, vous avez toujours été très disponible et travailler avec Goulven a été tout simplement génial ☺ Nous avons passé tellement de temps dans cette pièce à 16°C dans la pénombre, cela n'a pas toujours été facile, mais on s'en est sorti et je suis désolé de ne pas avoir réussi à en faire un Nature, mais peut-être qu'avec un peu de patience, tu finiras par l'avoir ! Toi et Jean-Christophe m'avez toujours soutenu tout au long de ce projet, même dans les moments les plus chaotiques où on se demandait si on allait vraiment finir par le valider ce facteur Z. La découverte de cette molécule 68 est le fruit d'une très belle collaboration, je sais que le projet est loin d'être terminé, mais je ne doute pas instant que vous arriviez à vos fins avec cette petite molécule ! Je terminerai par une petite note très importante à mes yeux, c'est également grâce à vous que j'ai pu goûter à toutes les délices de la cantine du

CEA, et lorsqu'on vient de Paris-Sud, c'est un vrai bonheur gustatif !!! Je profite également pour remercier Chantal et Sandra avec qui nous partageons nos repas, vous êtes toujours très souriantes et c'est aussi grâce à vous que ce séjour s'est super bien passé, alors merci !

Je remercie très sincèrement toute l'équipe Biophotonique, car c'est grâce à vous que j'ai découvert la molécule unique, une technique barbare qui fait rêver tant on a l'impression de pouvoir visualiser directement ce qu'il se passe au cours de la traduction. C'est néanmoins une technique extrêmement difficile à maîtriser et j'ai été très heureux de partager les bons moments comme les moins bons avec toi Nathalie. Je garderai de très bons souvenirs de ces journées passées dans le noir à regarder des petits points blancs disparaître de l'écran. Tu t'es impliquée à 200% sur ce projet, que ce soit en physique comme en biologie, tu as su les gérer les coups durs sans perdre patience, tu es toujours joyeuse et pleine de vie, et pour tout cela, je te remercie, car c'est vraiment grâce à toi que nous avons pu avancer et tenir le coup. Je remercie également Karen qui a initié ce projet ainsi que Nathalie W. qui a accepté de collaborer avec nous. Cela a été un gros travail de réflexion pour faire mûrir ce projet depuis sa création et vous avez fait au mieux, un grand merci pour ton investissement, ton accueil chaleureux et pour avoir eu la patience de m'expliquer ces équations barbares auquel j'étais plutôt allergique. Je remercie également Flavie avec qui j'ai partagé mes trajets, le bureau, de très nombreuses discussions et qui est désormais une très bonne amie sur qui je pourrai toujours compter. Échanger avec toi a été un vrai plaisir, tu es toujours à l'écoute et tu as une pêche d'enfer qui a égayé mes journées passées chez vous. J'espère très sincèrement que nous arriverons à maintenir cette amitié qui m'est chère. Je vous souhaite à toi et Nathalie bon courage pour la fin de votre thèse et merci pour tout. J'en profite aussi pour remercier les autres membres du groupe Biophotonique et principalement Aurore et Mitradeep avec qui j'ai eu la joie de partager nos repas et nos petites pauses à la cafétéria, et les doctorants qui nous ont précédés sur ce projet, Nicolas Fiszman et Hélène Chommy.

J'ai choisi de faire cette thèse en accordant 1/6 de mon temps à une activité doctorale que j'ai réalisé avec la merveilleuse équipe pédagogique du Deust et de la Licence Pro. de Bioindustries et Biotechnologies d'Orsay, eux-mêmes qui m'ont formé de la L1 à la L3. Je remercie donc Sylviane, Robert, Manue, Manu, Catherine, Guylaine, Armel et tous les autres membres de l'équipe sans oublier Gilles. Les missions qui m'ont été confiées n'ont pas été toujours faciles à réaliser, mais j'ai pris beaucoup de plaisir à travailler avec vous, à développer l'aspect communication et informatique de votre formation. Vous avez été les premiers à me soutenir lorsque j'ai souhaité poursuivre mes études et si aujourd'hui je suis là, c'est aussi grâce à vous, qui m'avez transmis votre passion pour les biotechnologies. Je vous serai donc éternellement reconnaissant, car j'ai passé une très bonne thèse, dans de très bonnes conditions, j'adore la recherche et j'ai été comme un poisson dans l'eau ces dernières années.

Bien sûr, je n'oublie pas nos super secrétaires, Muriel et Catherine qui sont juste adorables. Je ne compte plus les heures matinales passées à discuter dans vos bureaux. Vous faites un travail extra et dès que j'ai été dans le besoin, vous vous êtes mis en quatre pour m'aider. Je ne sais comment vous remercier... Je sais que vous n'attendez que cela, alors je vous souhaite une très bonne retraite que vous avez amplement méritée.

Je remercie aussi toutes les personnes que j'ai côtoyées à l'IGM puis à l'I2BC, notamment Murielle, qui nous a vus dans nos débuts avec Benoit. Tu ne perds jamais ton enthousiasme, tu es source de plein de bons conseils et tu es toujours à l'écoute. Un grand merci pour tout et surtout pour ta bonne humeur et ta bienveillance. Merci également à Karine, Anna, Sandrine et Mimi pour nos discussions et pour votre gentillesse. Merci à Ludivine, Agnès D., Christelle V., Bruno B., Jean-Luc P., Philippe B., Jacques O. et Luis A. pour vos précieux conseils et pour nos échanges.

Une pensée à tous les doctorants que j'ai côtoyés à l'IGM puis à l'I2BC et surtout à Benoit, un formidable ami qui m'est très cher et à qui je souhaite beaucoup de courage pour sa soutenance. Nous nous sommes suivis depuis la L1 et nous finirons ensemble cette année, cela aura été une sacrée aventure, pas toujours facile, mais riche en bons souvenirs. Tu es un très bon chercheur, tu mérites de croquer cette vie à pleine dent alors fonce et crois en toi !! Merci aussi à toi Libera qui a été un vrai moteur pour les doctorants de l'Institut, je te souhaite une très bonne continuation et j'espère que tu trouveras ton bonheur aux États-Unis. Je n'oublie pas non plus Drago avec qui nous avons eu de longues discussions, toutes plus délirantes les unes des autres, et merci à Audrey, Florence, Jerzy, Jorge et Matteo.

Je ne remercierai jamais assez mes amis, ma belle-famille et ma famille sans qui je ne serai pas celui que je suis avec une mention spéciale pour mes beaux-parents Germaine et Jean-Pierre, qui nous ont soutenus durant toute cette thèse, à Lise et François, mes parents exceptionnels, mes grands-parents Annette et Pierre, dont la curiosité, n'a d'égale que leur générosité, et mes deux frères extraordinaires, Juju et Coco à qui je souhaite tout le bonheur du monde.

Je finirai par la plus belle rencontre de ma vie, mon petit ange Aurélie, qui a émerveillé mon quotidien comme personne. Je ne compte plus les merveilleux moments passés ensemble depuis notre rencontre, notre soutien mutuel sans faille dans cette aventure qui fut jalonnée de doutes, d'excitations, de remises en question et de belles découvertes. Tu as été une source inépuisable de bonheur durant toute cette thèse et tes encouragements m'ont donné la force d'aller jusqu'au bout. Merci pour ta compréhension, pour ce sourire que tu m'offres chaque jour et pour ton amour qui me fait chaud au cœur. Il n'y a pas de mots pour te dire à quel point je te suis reconnaissant de tout ce que tu as fait pour moi. Je t'aime mon petit ange.

# Table des matières

## Introduction

<b>I. Préambule</b> .....	<b>10</b>
<b>II. La traduction chez les eucaryotes</b> .....	<b>10</b>
1. Le code génétique .....	11
2. Les acteurs de la traduction.....	12
A. L'ARN messager .....	12
B. L'ARN de transfert .....	13
C. Le ribosome .....	14
3. Les étapes de la traduction .....	15
A. L'initiation .....	16
B. L'élongation .....	18
a. Positionnement et accommodation .....	18
b. Formation de la liaison peptidique et translocation.....	19
C. La terminaison.....	19
a. Les facteurs de terminaison.....	19
b. Les mécanismes de la terminaison .....	22
D. Le recyclage du ribosome.....	24
a. La dissociation des sous-unités ribosomiques.....	24
b. Recyclage de l'ARNm, de l'ARNt et des sous-unités ribosomiques.....	25
E. La réinitiation .....	25
F. Les partenaires des facteurs de terminaison.....	26
a. La PABP .....	26
b. Les protéines Upf .....	26
c. L'ARN hélicase Dbp5 .....	27
4. La dégradation des ARNm : surveillance et contrôle de l'expression génique....	27
A. La dégradation cytoplasmique des ARNm .....	28
B. Le contrôle de l'expression génique .....	28
C. La surveillance de la qualité des ARNm.....	29
a. La voie NGD .....	30
b. La voie NSD.....	32
c. La voie NMD .....	33
<b>III. La fidélité de la traduction</b> .....	<b>37</b>
1. Le rôle du ribosome dans le maintien de la fidélité de la traduction .....	38
2. Les mécanismes de décodage .....	38
A. La sélection de l'ARNt .....	39

B.	Les régions impliquées dans le décodage .....	40
C.	Les mécanismes du décodage .....	41
a.	Contrôle de l'appariement codon-anticodon par les bases A1492-A1493... 41	
b.	Contrôle de l'appariement codon-anticodon via les contraintes exercées par le centre de décodage .....	41
3.	Le recodage.....	43
A.	Le décalage du cadre de lecture (Frameshift) .....	43
a.	Le décalage du cadre de lecture en -1 .....	44
b.	Le décalage du cadre de lecture en +1 .....	45
B.	Le passage du codon stop.....	46
a.	La translecture .....	46
b.	Incorporation de la sélénocystéine .....	47
c.	Incorporation de la pyrrolysine.....	48
C.	Le saut de ribosome .....	48
4.	Les facteurs connus pour modifier la fidélité de la terminaison de la traduction. 49	
A.	Les facteurs agissant en <i>cis</i> .....	49
a.	La nature du codon stop .....	49
b.	Le contexte nucléotidique.....	50
c.	Les structures secondaires.....	52
B.	Les facteurs agissant en <i>trans</i> .....	52
a.	Les ARNt supprimeurs.....	52
b.	Les pseudouridines .....	54
c.	Les antibiotiques .....	55
d.	Les aminoglycosides .....	56
e.	La négamycine .....	59
<b>IV.</b>	<b>Les maladies à codon stop .....</b>	<b>60</b>
1.	Les pathologies humaines concernées.....	60
A.	La myopathie de Duchenne.....	61
B.	La mucoviscidose.....	62
C.	Les cancers.....	63
a.	La cancérogenèse .....	63
b.	Le gène suppresseur de tumeur P53 .....	65
c.	Caractéristiques des cellules tumorales .....	66
d.	Cibles thérapeutiques .....	68
D.	Les autres maladies à codons stop étudiées .....	69
2.	Stratégies thérapeutiques .....	70
A.	Le saut d'exon.....	71
B.	Les ARNt supprimeurs .....	72

C.	La pseudouridylation .....	73
D.	L'inhibition de la voie NMD .....	74
3.	Suppression traductionnelle pharmacologique.....	76
A.	Les éléments de l'ARNm impactant la translecture .....	77
a.	La nature du codon stop et son contexte .....	77
b.	L'acide aminé incorporé .....	78
B.	Les molécules inductrices de translecture .....	79
a.	Les molécules ayant un effet avéré .....	79
b.	Les molécules ayant un effet supposé .....	87

## Objectifs

V.	Objectifs de thèse .....	92
----	--------------------------	----

## Résultats

### Résultats : Traduction en molécule unique..... 94

1.	Résumé de l'article : Kinetics of IRES-mediated eukaryotic using single molecule fluorescence microscopy .....	94
2.	Article : Kinetics of IRES-mediated eukaryotic using single molecule fluorescence microscopy.....	97
3.	Traduction in vitro sur billes magnétiques en mesure d'ensemble .....	98
4.	Acquisition de cinétiques de traduction complémentaires .....	98

### Résultats : Caractérisation de nouveaux inducteurs de translecture développés par design rationnel .....

1.	Résumé de l'article : Cancer cell apoptosis induced by translational suppression of p53 nonsense mutation .....	101
2.	Article : Cancer cell apoptosis induced by translational suppression of p53 nonsense mutation .....	103

### Résultats : Identification de nouvelles molécules inductrices de translecture par criblage haut débit..... 105

1.	Criblage de la chimiothèque Greenpharma .....	105
2.	Criblage de la Chimiothèque Nationale Essentielle (CNE) .....	106
3.	Criblage à haut débit de la Prestwick et d'une fraction de la Chembridge .....	108
A.	Construction du test biologique .....	108
B.	Contrôle qualité lors d'un crible à grande échelle .....	109
C.	Mise au point du dosage pour un crible à grande échelle .....	110
D.	Protocole de criblage .....	112
E.	Nature des banques criblées .....	112
F.	Correction des erreurs systématiques .....	113

G.	Analyse des résultats .....	113
H.	Crible secondaire .....	114
I.	Caractérisation des molécules hits.....	115
a.	Capacité à induire de la translecture.....	115
b.	Capacité à réexprimer une protéine de pleine taille.....	115
c.	Capacité à stabiliser le transcrit P53.....	116
d.	Synergie des molécules TLN avec un aminoglycoside .....	116
e.	Capacité de TLN468 à induire la réexpression d'une protéine p53 active .	117
f.	Capacité de la protéine p53 réexprimée à activer ses gènes cibles.....	117
g.	Toxicité cellulaire de TLN468.....	118
J.	Les analogues de TLN468 .....	118
a.	Recherche et criblage de 378 analogues.....	118
b.	Efficacité de translecture des analogues de TLN468.....	119
K.	Spécificité de la molécule TLN468 .....	119
L.	Effet de la nature du codon stop .....	120

## Discussion

<b>Discussion.....</b>	<b>122</b>
1. Les approches développées permettant l'identification de nouvelles molécules inductrices de translecture.....	122
2. Le NB124 et la TLN468 induisent la translecture des codons stop.....	124
3. Le NB124 et la TLN468 stabilisent l'ARNm et restaurent la synthèse d'une protéine de pleine taille.....	125
4. Le traitement par le NB124 et la TLN468 conduit à la synthèse d'une protéine fonctionnelle .....	127
5. La TLN468 est une molécule qui nécessite d'être optimisée.....	129
6. La protéine p53 pleine taille réexprimée par le NB124 est capable de limiter la prolifération cellulaire en induisant l'apoptose.....	130
7. La molécule unique, un outil d'analyse dynamique de la traduction eucaryote	131

## Conclusion et perspectives

<b>Conclusion - Perspectives .....</b>	<b>138</b>
<b>Références bibliographiques.....</b>	<b>144</b>

<b>Matériel et Méthodes</b> .....	<b>176</b>
<b>I. Le crible à grande échelle</b> .....	<b>176</b>
1. Protocole de réalisation du crible.....	176
2. Normalisation des résultats .....	179
<b>II. La molécule unique</b> .....	<b>179</b>
1. Transcription <i>in vitro</i> .....	179
2. Traduction sur billes magnétiques avec de la méthionine S <sup>35</sup> .....	180
<b>Annexes</b> .....	<b>184</b>
Annexe 1 : Nouvelle nomenclature des protéines ribosomales .....	184
Annexe 2 : Arbre décisionnel de l'analyse statistique et des résultats des cribles primaires et secondaires.....	186
Annexe 3 : Viabilité cellulaire des lignées NIH3T3, HDQ-P1 et H1299 après traitement par la molécule TLN468.....	187
Annexe 4 : Analyses statistiques de la contribution des ribosomes du RRL.....	188

## Liste des figures et des tableaux

### Introduction :

Figure 1 : Expression génique chez les eucaryotes.....	10
Figure 2 : Le code génétique .....	11
Figure 3 : Réassignation des codons selon les organismes .....	12
Figure 4 : Structure schématique d'un ARNm.....	13
Figure 5 : Structure secondaire et tertiaire de l'ARNt.....	13
Figure 6 : Evolution stratifiée du ribosome eucaryote. ....	14
Figure 7 : Structure cristallographique du ribosome 80S eucaryote.....	15
Figure 8 : Modèle de la voie classique d'initiation de la traduction eucaryote .....	16
Figure 9 : Modèle de l'initiation de la traduction médiée par les IRES.....	17
Figure 10 : Modèle du cycle d'élongation de la traduction eucaryote .....	18
Figure 11 : Formation de la liaison peptidique .....	19
Figure 12 : Comparaison de la position et de l'orientation des ARNt entre l'état classique et l'état hybride .....	19
Figure 13 : Structure cristallographique du facteur de terminaison eRF1 humain.....	20
Figure 14 : Structure cristallographique du domaine C du facteur de terminaison eRF3 <i>S.pombe</i> .....	21
Figure 15 : Structure cristallographique du complexe eRF1-eRF3 humain.....	22
Figure 16 : Interactions entre eRF1-eRF3.....	22
Figure 17 : Résidus du domaine N d'eRF1 impliqués dans la reconnaissance de chacune des trois bases du codon stop.....	23
Figure 18 : Structures des codons stop dans le centre de décodage.....	23
Figure 19 : Modèle de la terminaison de la traduction eucaryote .....	24
Figure 20 : Rôle du GGQ dans l'hydrolyse du peptide.....	24
Figure 21 : Structure cryo-EM du complexe de terminaison eRF1-eRF3c et pré-terminaison/recyclage eRF1-ABCE1 avec l'ARNt peptidyl.....	25
Figure 22 : Modèle de dissociation des sous-unités ribosomiques de Barthelme et son équipe.....	25
Figure 23 : Modèle de dissociation des sous-unités ribosomiques de Pestova et son équipe.....	26
Figure 24 : Modèle du recyclage de l'ARNm, de l'ARNt et des sous-unités ribosomiques .....	26
Figure 25 : Domaines de la protéine PABP1 .....	27
Figure 26 : Modèle de dégradation des ARNm normaux.....	28
Figure 27 : Modèles des mécanismes de dégradation des ARNm par le système d'ARN interférence.....	29
Figure 28 : Diversité des systèmes de contrôle qualité des ARNm eucaryotes.....	30
Figure 29 : Modèle partiel du mécanisme de la voie NGD (No-Go Decay) .....	31
Figure 30 : Modèle de prise en charge de la chaîne naissante par le complexe RQC ...	31
Figure 31 : Modèle du mécanisme de la voie NSD .....	32

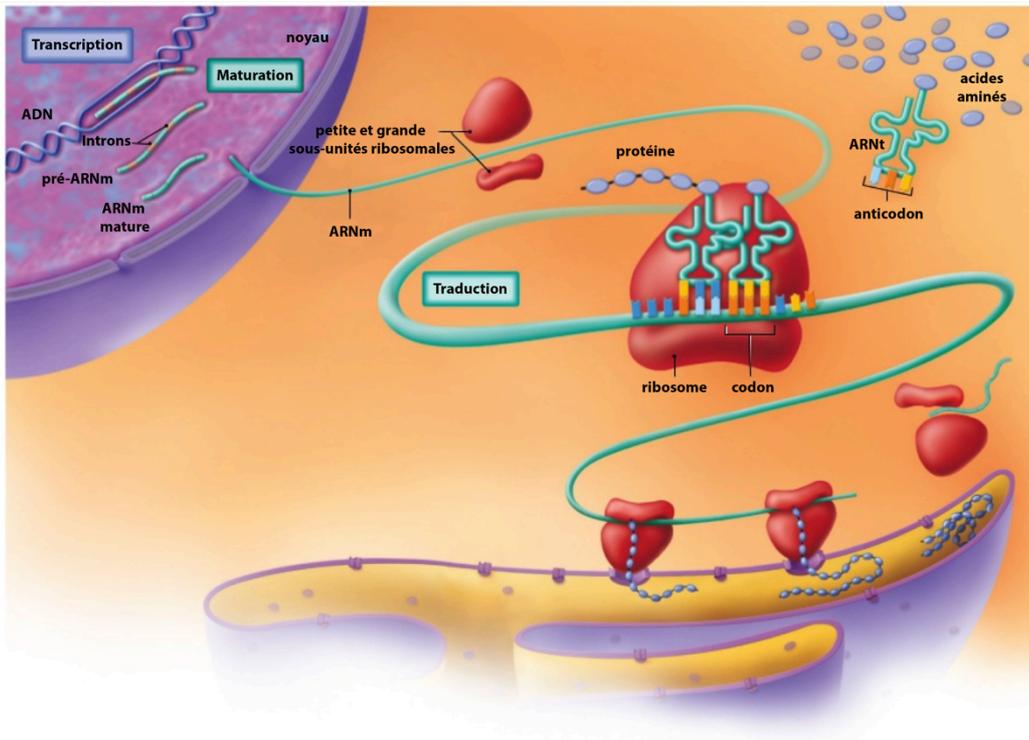
Figure 32 : Illustration schématique des domaines et des motifs des facteurs importants de la NMD chez les mammifères .....	33
Figure 33 : Modèle d'une voie NMD exclusive de la première traduction .....	34
Figure 34 : Modèle unifié de la voie NMD .....	35
Figure 35 : Centre de décodage et centre peptidyl-transférase du ribosome 70S .....	38
Figure 36 : Processus de sélection des ARNt.....	39
Figure 37 : Régions impliqués dans le décodage .....	40
Figure 38 : Interactions codon-anticodon dans le centre de décodage du ribosome 70S.....	40
Figure 39 : Conformation du site A dans la petite sous-unité 30S.....	41
Figure 40 : Conformation des bases conservées A1492, A1493 et G530 dans le site de décodage.....	41
Figure 41 : Mécanisme de l'infidélité traductionnelle .....	42
Figure 42 : Comparaison des appariements Watson-Crick, wobble et tautomérique rare .....	43
Figure 43 : Relation entre séquences nucléotidiques et protéines issues d'un décodage alternatif.....	43
Figure 44 : Modèle cinétique du frameshift -1 .....	44
Figure 45 : La translecture.....	46
Figure 46 : Synthèse et incorporation de la sélénocystéine chez les eucaryotes.....	47
Figure 47 : Site du recodage du gène 60 chez le Bactériophage T4.....	48
Figure 48 : Mécanisme du saut de ribosome .....	48
Figure 49 : Les facteurs modulant la fidélité de la terminaison de la traduction .....	49
Figure 50 : Structure secondaire des ARNt supprimeurs .....	52
Figure 51 : Répartition des pseudouridines et pseudouridines synthèses chez la levure.....	53
Figure 52 : Répartition des pseudouridines et pseudouridines synthèses chez l' <i>Homo sapiens</i> .....	54
Figure 53 : Schéma représentant l'isomérisation de l'uridine (U) en pseudouridine (Ψ).....	54
Figure 54 : Cible des antibiotiques au cours de la traduction bactérienne .....	55
Figure 55 : Site de fixation des antibiotiques sur la 30S .....	55
Figure 56 : Site de fixation des antibiotiques sur la 50S .....	56
Figure 57 : Structure des aminoglycosides .....	56
Figure 58 : Fixation des aminoglycosides dans le ribosome 70S.....	57
Figure 59 : Sites d'interactions de la paromomycine dans le centre de décodage .....	57
Figure 60 : Fixation de la paromomycine de centre de décodage bactérien.....	58
Figure 61 : Structure de la paromomycine dans le centre de décodage bactérien.....	58
Figure 62 : Structure d'aminoglycosides dans le centre de décodage eucaryote.....	59
Figure 63 : Fixation de la négamycine dans le ribosome 70S.....	60
Figure 64 : Schéma du complexe glycoprotéiques associé à la dystrophine .....	61
Figure 65 : Schéma de la structure du régulateur de la conductance transmembranaire de la fibrose kystique (CFTR) .....	62
Figure 66 : La cancérogénèse .....	63
Figure 67 : Structure tridimensionnelle de P53 sous forme de tetramères.....	64
Figure 68 : La protéine p53.....	65
Figure 69 : Signaux d'activation et réponses de p53.....	65

Figure 70 : Gènes cibles de la protéine p53 .....	66
Figure 71 : Traits caractéristiques des cellules tumorales .....	67
Figure 72 : Cibles thérapeutiques des caractéristiques du cancer.....	68
Figure 73 : Le saut d'exon .....	71
Figure 74 : Structure chimique de l'eteplirsén et drisapersén .....	72
Figure 75 : Les ARNt supresseurs .....	72
Figure 76 : La pseudouridylation comme supresseur de PTC .....	73
Figure 77 - Inhibition de la NMD .....	74
Figure 78 – Structure des molécules inhibitrices de la voie de dégradation NMD .....	75
Figure 79 : Principaux inducteurs de translecture de la famille des aminoglycosides ....	79
Figure 80 : Les trois groupes de dérivés d'aminoglycosides .....	82
Figure 81 : Molécules dérivées de la paromomycine (NB).....	83
Figure 82 : Structure de la négamycine et de l'un de ses analogues .....	85
Figure 83 : Molécules non-aminoglycosides dont l'effet sur la translecture est supposé .....	86

Tableau 1 : Influence de la nature du nucléotide +4 sur la translecture basale.....	50
Tableau 2 : Influence de la nature du nucléotide +4 sur la translecture induite .....	78
Tableau 3 : Synthèse des essais réalisés avec les aminoglycosides.....	80
Tableau 4 : Synthèse des essais cliniques réalisés avec les aminoglycosides.....	81
Tableau 5 : Synthèse des essais cliniques réalisés avec le PTC124 (Ataluren) .....	89

# ***Introduction***

***La sagesse commence dans l'émerveillement  
(Socrate)***



**Figure 1 : Expression génique chez les eucaryotes.**

Un gène présent sur la molécule d'ADN, et codant une protéine, est transcrit en un pré-ARNm qui sera mûré puis exporté dans le cytoplasme. L'ARNm est ensuite traduit par le ribosome pour aboutir à la synthèse d'une protéine. *Figure adaptée de The McGraw-Hill Education*

## I. Préambule

Au cours de ma thèse, je me suis intéressé à la suppression traductionnelle des codons stop chez les mammifères. Je débuterai cette introduction en présentant les règles du code génétique qui régissent la traduction, les acteurs de ce processus et les mécanismes moléculaires mis en jeu à chacune des étapes. Je détaillerai l'étape de terminaison de la traduction sur laquelle je me suis concentré durant mes recherches. Puis, je présenterai les mécanismes de surveillance et contrôle de l'expression génique via la régulation des ARNm. J'aborderai ensuite la fidélité de la traduction et les différents événements non conventionnels de traduction (recodage) dans lesquels les règles canoniques de traduction sont transitoirement contournées afin de réguler l'expression de certains gènes. J'exposerai ensuite les facteurs connus pour modifier la fidélité de la traduction. Je terminerai cette introduction en présentant les maladies dites à codon stop et les stratégies thérapeutiques développées pour restaurer l'expression des protéines impliquées dans ces maladies. L'introduction sera suivie des résultats obtenus sur le développement d'un nouvel outil d'étude dynamique de la traduction eucaryote et sur l'identification de nouvelles molécules inductrices de translecture. Je poursuivrai par une discussion de ces résultats avant de conclure et de présenter les perspectives de mes projets de recherche.

## II. La traduction chez les eucaryotes

La biosynthèse des protéines ou traduction est un processus complexe qui nécessite l'action contrôlée de nombreuses macromolécules. Ce processus essentiel à la vie de la cellule consomme plus de la moitié de l'énergie cellulaire et est par conséquent finement régulé <sup>1</sup>. L'information génétique nécessaire à la fabrication des protéines est codée dans la molécule d'ADN (AcideDésoxyriboNucléique), sous forme d'une succession de nucléotides. Chez les eucaryotes, les séquences d'ADN susceptibles d'être transcrites en ARN sont appelées gènes. Ces derniers sont généralement segmentés en régions codantes (exons) interrompues par des segments non codants (introns). Les gènes sont transcrits en pré-ARN messager (Acide Ribo-Nucléique) avant de subir plusieurs étapes de maturation, durant lesquelles les extrémités de l'ARN sont modifiées et les introns excisés (épissage), formant ainsi l'ARN messager (ARNm). L'ARNm est alors exporté dans le cytoplasme où il sera traduit en protéine (Figure 1).

		Deuxième position									
		U	C	A	G						
Première position	U	UUU ]	UCU ] UCC ] UCA ] UCG ]	UUA ] UUG ]	UGU ] UGC ] UGA ] UGG ]	U C A G					
		UUC ]					Serine (Ser)	Tyrosine (Tyr)	Cysteine (Cys)		
		UUA ]								Stop	Stop
		UUG ]									
	C	CUU ]	CCU ] CCC ] CCA ] CCG ]	CAU ] CAC ] CAA ] CAG ]	CGU ] CGC ] CGA ] CGG ]	U C A G					
		CUC ]					Leucine (Leu)	Histidine (His)	Arginine (Arg)		
		CUA ]								Glutamine (Gln)	
		CUG ]									
	A	AUU ]	ACU ] ACC ] ACA ] ACG ]	AAU ] AAC ] AAA ] AAG ]	AGU ] AGC ] AGA ] AGG ]	U C A G					
		AUC ]					Isoleucine (Ile)	Asparagine (Asn)	Serine (Ser)		
		AUA ]								Lysine (Lys)	Arginine (Arg)
		AUG ]									
	G	GUU ]	GCU ] GCC ] GCA ] GCG ]	GAU ] GAC ] GAA ] GAG ]	GGU ] GGC ] GGA ] GGG ]	U C A G					
		GUC ]					Valine (Val)	Aspartic acid (Asp)	Glycine (Gly)		
		GUA ]								Alanine (Ala)	
GUG ]		Glutamic acid (Glu)									

**Figure 2 : Le code génétique**

Tableau regroupant les correspondances entre codons et acides aminés. Le codon AUG initiateur de la traduction cytoplasmique est indiqué en vert et les trois codons de terminaison UAA, UAG et UGA sont indiqués en rouge.

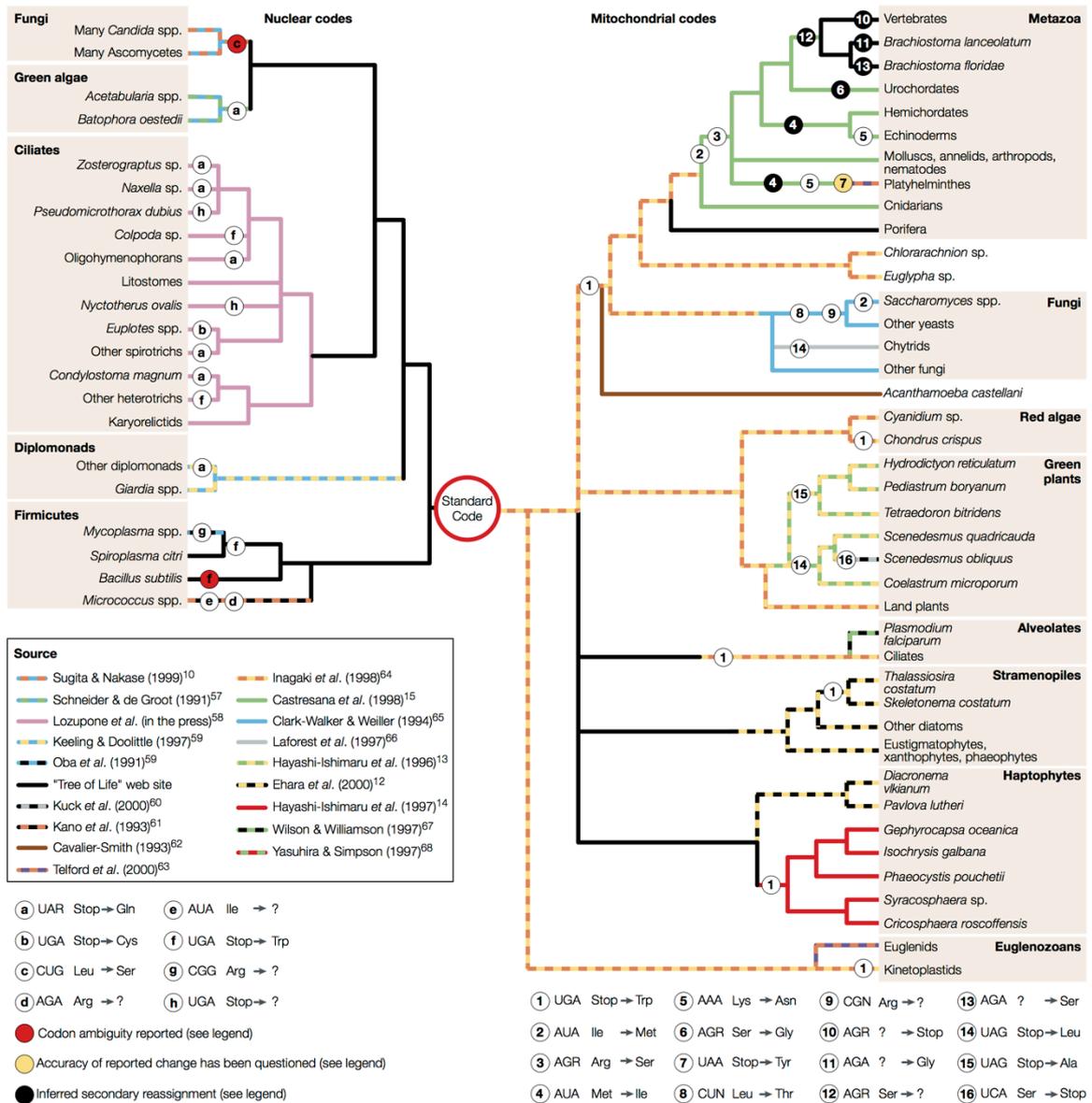
## 1. Le code génétique

Lors de la traduction, les quatre bases de l'ARNm (Adénine, Uracile, Guanine et Cytosine) sont traduites selon les règles du code génétique. Élucidé en 1961 par Francis Crick et ses collaborateurs, le code génétique correspond à un ensemble de règles établissant une correspondance entre un triplet de nucléotides de l'ARNm appelé codon et un acide aminé. Il fut ainsi établi que : - chaque séquence de nucléotides est lue à partir d'un point de départ spécifique - les triplets ne se chevauchent pas - il n'y a pas de séparateurs permettant d'identifier les « bons triplets » (bien que les introns furent découverts par la suite) - le code génétique est dégénéré <sup>2</sup>. En effet, le caractère dégénéré résulte de la disproportion entre les 64 codons possibles et les 20 acides aminés standards. Il faudra attendre les travaux de Nirenberg en 1965 pour connaître l'acide aminé associé à chaque codon. Il a ainsi été montré que les 61 codons « sens » sont chacun assignés à un acide-aminé et que les trois codons restants : ambre (UAG), ocre (UAA) et opale (UGA) sont dits « non-sens » ou « stop », car ils signalent l'arrêt de la traduction (Figure 2) <sup>3</sup>.

Hormis le codon d'initiation méthionine (AUG) et les codons stop, le tryptophane est l'unique acide aminé associé à un seul codon (UGG). Ainsi, chaque acide aminé correspond à plusieurs codons, mais cette distribution n'étant pas aléatoire, la majorité des codons spécifiant le même acide aminé ne diffèrent que par leur troisième base, dite « wobble » (flexible). Ceci à l'exception des acides aminés leucine, arginine et serine qui sont associés à six codons.

Le code génétique a longtemps été considéré comme un « accident figé » incapable d'évoluer <sup>4</sup>. La découverte d'un code génétique alternatif dans les mitochondries humaines a remis en cause son « universalité » <sup>5</sup>. La réassignation de certains codons fut largement étudiée chez les espèces dont le code génétique est différent du code canonique (Figure 3) <sup>4</sup>. Il apparaît ainsi que ces variations concernent la plupart du temps les codons stop. Certaines réassignations identiques semblent avoir eu lieu de manière indépendante dans différents taxons, ce qui laisserait penser que le code génétique poursuivrait son évolution dans certaines lignées.

En plus des 20 acides-aminés correspondant aux codons sens, il existe 2 acides aminés naturels rares : la sélénocystéine et la pyrrolysine, qui sont incorporés dans les protéines au niveau des codons UGA et UAG respectivement dans des cas particuliers de recodage. Cette notion sera abordée plus tard dans la partie III. *La fidélité de la traduction.*



**Figure 3 : Réassignation des codons selon les organismes**

L'arbre phylogénétique retrace l'évolution de la réassignation de différents codons dans les génomes nucléaires (à gauche) et mitochondriaux (à droite). *Figure issue de Knight et al., 2001* <sup>4</sup>

## 2. Les acteurs de la traduction

Bien que les principaux mécanismes de la traduction aient été découverts et élucidés grâce à l'étude des organismes procaryotes, le sujet de ce mémoire impose de se focaliser sur les caractéristiques de la traduction eucaryote, qui a connu des avancées significatives ces dernières années.

La traduction est réalisée grâce à l'action de plusieurs acteurs dont les principaux sont : l'ARN messager (ARNm), l'ARN de transfert (ARNt), le ribosome et de nombreux facteurs protéiques. Dans la cellule, environ 80% des ARN totaux sont des ARN ribosomiques, 15% des ARN de transfert et 5% des ARNm <sup>6</sup>.

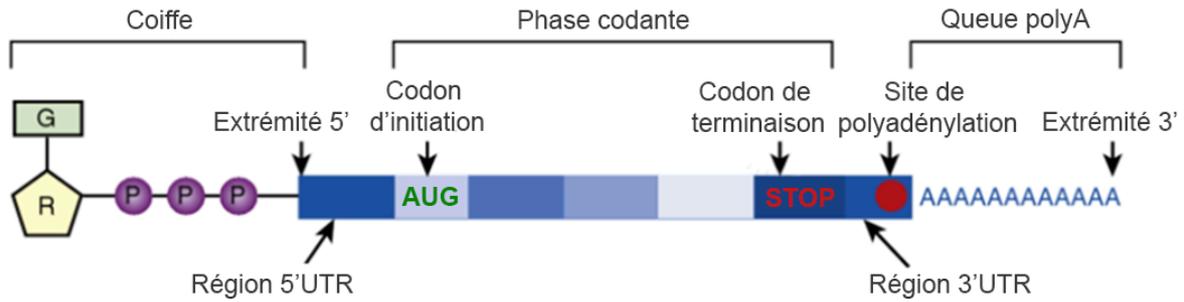
### A. L'ARN messager

Les ARNm sont transcrits sous la forme d'un pré-ARNm par l'ARN polymérase 2 (Pol II). Ils subissent ensuite plusieurs événements de maturation qui vont modeler le transcrit primaire en une forme lisible par la machinerie traductionnelle appelée mRNP pour messenger ribonucléoprotéique. Ce complexe contient l'ARNm et des protéines de liaisons à l'ARN (RBPs) qui jouent notamment un rôle majeur dans le contrôle post-transcriptionnel des ARN.

La biogenèse des ARNm est souvent décrite comme une suite séquentielle de modifications durant lesquels la transcription est suivie de l'ajout de la coiffe, puis de l'épissage et enfin de l'ajout de la queue poly(A). Ce schéma concorde avec la reconstitution de l'ensemble de ces réactions *in vitro*, indépendamment les unes des autres. Toutefois, dans les cellules, la transcription et la maturation sont le plus souvent simultanées <sup>7</sup>.

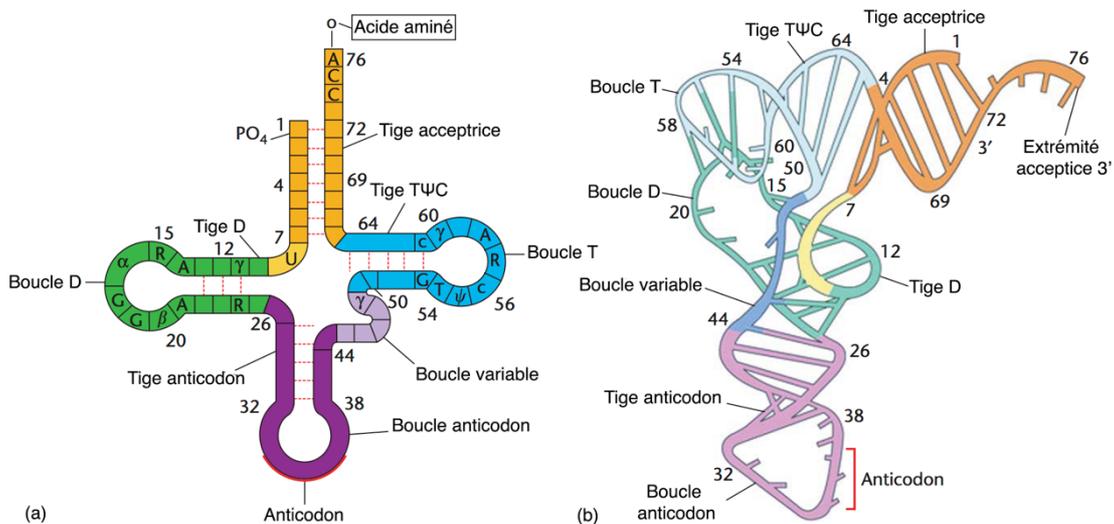
La première étape de la maturation consiste en la modification des extrémités du pré-ARNm. Tout d'abord, une coiffe (7-méthylguanosine triphosphate (m7G)) est ajoutée en 5', ce qui va permettre de protéger l'ARNm de la dégradation et de recruter le ribosome lors de l'initiation traductionnelle <sup>8,9</sup>.

Au cours de la transcription, les séquences non codantes appelées introns sont excisées et les exons sont réunis par ligation. Un gène humain contient en moyenne 8 introns, ce qui représente environ 90% de la taille totale de l'ARN pré-messager. Les introns ont une taille extrêmement variable, de quelques centaines de nucléotides à plusieurs centaines de milliers.



**Figure 4 : Structure schématique d'un ARNm**

L'ARNm mature comprend : une coiffe à l'extrémité 5', une région non-codante en 5' (5'UTR), une région codante commençant par un codon d'initiation et se finissant au premier codon de terminaison rencontré, une région non-codante en 3' (3'UTR) et enfin une queue poly A. *Figure adaptée de Clark and Pazdernik, 2016 – Chap2*<sup>10</sup>



**Figure 5 : Structure secondaire et tertiaire de l'ARNt**

(a) Diagramme de la structure secondaire d'un ARNt schématiquement représentée en forme de trèfle. Les bases universellement conservées sont indiquées par le numéro de leur position. (b) Structure tridimensionnelle en forme de L adoptée par l'ARNt lorsque la tige D se replie sur la tige TΨC. *Figure adaptée de Goldman, 2008*<sup>11</sup>

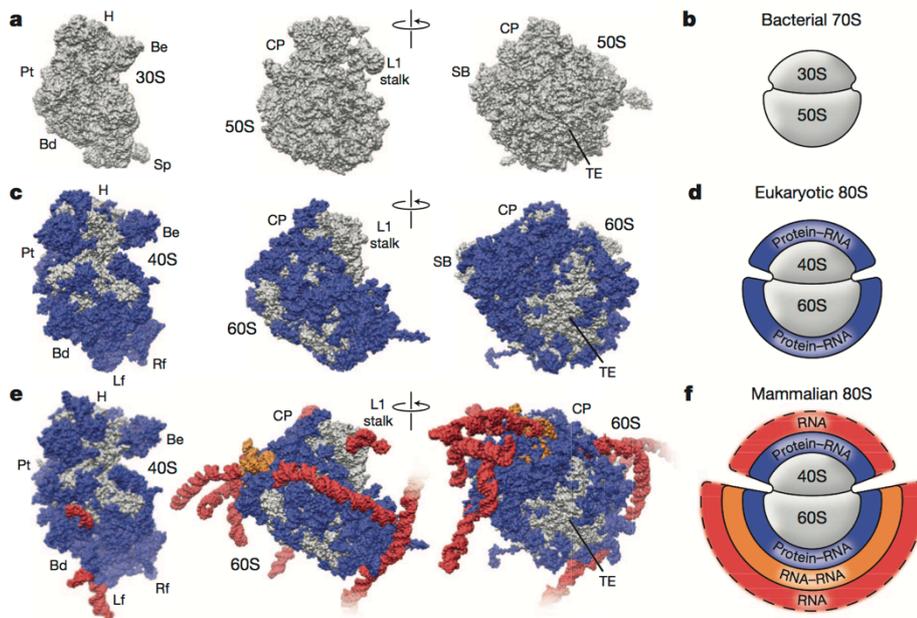
Enfin, l'extrémité 3' du pré-ARNm est modifiée. Ce processus de polyadénylation est réalisé en deux étapes : la coupure de la séquence d'ARNm en amont d'une séquence cible, puis l'ajout d'une queue poly(A) de 250 à 300 adénines chez les humains <sup>12</sup>. En cours de synthèse, certaines Poly(A) Binding Protein (PABP) sont également recrutées pour stimuler l'ajout d'adénines et protéger l'ARNm de la dégradation. Cette étape est nécessaire à l'export du transcrit mature dans le cytoplasme. Elle intervient également dans la stabilité du messenger et dans l'efficacité de la traduction <sup>13</sup>.

L'ARNm mûré comprend alors : une coiffe à l'extrémité 5' ; une région non codante en 5' (5'UTR) dont le rôle principal est de permettre l'initiation de la traduction ; une région codante commençant par un codon de démarrage et se finissant au premier codon stop rencontré ; une région non codante en 3' (3'UTR) ; une queue poly(A) (Figure 4). Nous verrons par la suite que ces régions ne sont pas aussi bien définies. En effet, il existe des régions codantes dans les régions non codantes en 5' et en 3'.

## B. L'ARN de transfert

Découverts en 1965 par R. Holley, les ARN de transfert (ARNt) sont une classe de molécules qui assurent la correspondance entre l'ARNm et la protéine. Ils servent d'adaptateurs entre les triplets de nucléotides correspondant aux codons et les acides aminés. Les gènes codant les ARNt eucaryotes sont transcrits par l'ARN polymérase III. Ces transcrits sont ensuite mûrés pour être complètement actifs. La majorité des ARNt contient entre 75 et 95 nucléotides <sup>11</sup>.

L'analyse de plusieurs séquences d'ARNt a conduit à la généralisation d'une représentation 2D s'apparentant à trèfle dont trois des quatre « feuilles » sont des tige-boucles (Figure 5a). La première correspond à la tige-boucle anticodon qui est impliquée dans la reconnaissance des bases 1, 2 et 3 du codon par interaction avec les bases 36, 35 et 34 de l'anticodon. Les deux autres sont la tige-boucle D et la tige-boucle T ou TΨC qui doivent leur nom à la présence de ribonucléotides modifiés, la dihydrouridine pour le bras D et la ribothymidine et la pseudouridine pour le bras TΨC. Le bras accepteur est formé par l'appariement des extrémités 5' et 3' de l'ARNt additionné d'un triplet de nucléotides CCA non appariés en 3', impliqué dans le chargement de l'acide aminé. Ce triplet a la particularité d'être ajouté au cours de la maturation des ARNt chez les bactéries Gram-positives et chez les eucaryotes <sup>14</sup>. Les ARNt possèdent également une boucle variable qui peut se composer de 4 à 21 nucléotides <sup>11</sup>.



**Figure 6 : Evolution stratifiée du ribosome eucaryote.**

Représentations en surface (**a,c,e**) et schématique (**b,d,f**) des ribosome de *T. thermophilus* (**a,b**), *S. cerevisiae* (**c,d**) et *H. sapiens* (**e,f**). La couche supplémentaire d'interactions protéine-ARN est représentée en bleu. Les deux couches supplémentaires d'interactions ARN-ARN et d'ARN seul chez le ribosome des mammifères sont représentées en orange et rouge respectivement. **TE** : tunnel de sortie, **CP**: protubérance centrale, **Pt** : plateforme, **Bd** : corps, **H** : tête. *Figure adaptée de Anger et al., 2013* <sup>15</sup>

Dans la structure tridimensionnelle, l'ARNt adopte une conformation en forme de L dans laquelle la tige D se replie sur la tige TΨC. La tige anticodon et la tige D forment alors le bras anticodon, la tige acceptrice et la tige TΨC constituent le bras accepteur (Figure 5b).

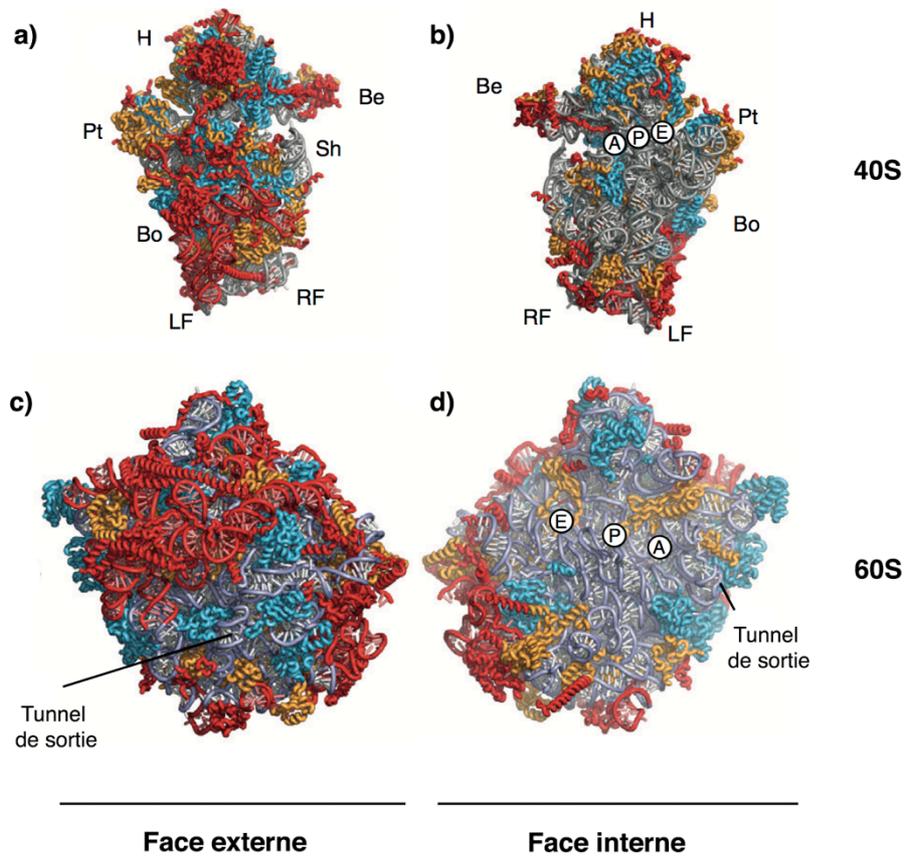
Alors qu'une seule enzyme est nécessaire à la synthèse de l'ensemble des ARNt, plus de 100 enzymes différentes sont impliquées dans la modification post-transcriptionnelle des ARNt <sup>16</sup>. Et bien que ces modifications soient le plus souvent impliquées dans l'efficacité de la traduction et dans l'aminocyclation des ARNt, le rôle exact de beaucoup d'entre elles est encore mal compris <sup>17</sup>. Comme pour l'ARNm, les étapes de maturation de l'ARNt ont lieu dans le noyau, avant le transport du transcrit dans le cytoplasme.

Après maturation, les acides aminés sont chargés sur l'extrémité 3' de l'ARNt par un groupe spécifique d'enzymes appelé aminoacyl-ARNt Synthetases (aaRS) pour former l'aminocycl-ARNt (aa-ARNt) au cours de la réaction d'aminocyclation. Cette réaction nécessite deux substrats : un groupe d'ARNt qui partagent un ensemble « d'éléments d'identités » et des acides aminés qui peuvent être distingués selon leur taille et par la présence de petites différences de propriétés des chaînes latérales <sup>18</sup>. L'aminocyclation constitue une étape cruciale de la traduction, car une fois que le chargement de l'acide aminé a eu lieu, il est irréversible. Grâce aux différents processus de contrôle, sur lequel je reviendrai ultérieurement, le taux d'erreur du chargement de l'acide aminé sur son ARNt est très faible, de  $10^{-4}$  à  $10^{-5}$  chez les procaryotes <sup>19</sup>.

### C. Le ribosome

Le ribosome humain est un complexe ribonucléoprotéique de 4,3 MDa composé d'ARN ribosomiques ayant une activité catalytique (ribozyme) et de protéines enchâssées dans l'ARNr. Bien que le cœur du ribosome soit conservé dans tout le règne du vivant, les ribosomes eucaryotes, et en particulier ceux des mammifères, ont une taille et une complexité significativement supérieures à celles des ribosomes procaryotes (Figure 6) <sup>15</sup>.

En décodant l'information contenue sur l'ARNm, le ribosome synthétise les protéines nécessaires à la cellule. Les eucaryotes possèdent un ribosome de 80S (en référence au coefficient de sédimentation) formé de deux sous-unités : la petite sous-unité 40S et la grande sous-unité 60S (<http://www.bangroup.ethz.ch>). A l'interface de ces deux sous-unités il y a 3 sites dans lesquels interagissent l'ARNm et les ARNt : le site A (Aminoacyl site) dans lequel entre l'aminocycl-ARNt, le site P (Peptidyl site) où l'acide



**Figure 7 : Structure cristallographique du ribosome 80S eucaryote**

Représentation de la face externe (a,c) et interne (b,d) de la petite (a,b) et grande (c,d) sous-unité du ribosome de *Tetrahymena thermophila* (PDB code 2xzm). Les protéines universellement conservées sont indiquées en bleu clair. Les protéines présentes chez les archées et eucaryotes en jaune et les protéines et éléments d'ARN exclusivement présents chez les eucaryotes sont en rouge. H : tête, Be : bec, Pt : plateforme, Sh : épaule, Bo : corps, LF : pied gauche, LR : pied droit, A-P-E : sites de positionnement de l'ARNt. *Figure adaptée de Klinge et al., 2012*<sup>20</sup>

aminé chargé sur l'ARNt est lié à la chaîne peptidique en cours de synthèse et enfin le site E (Exit site) dans lequel se trouve l'ARNt déacylé prêt à sortir du ribosome <sup>21</sup>.

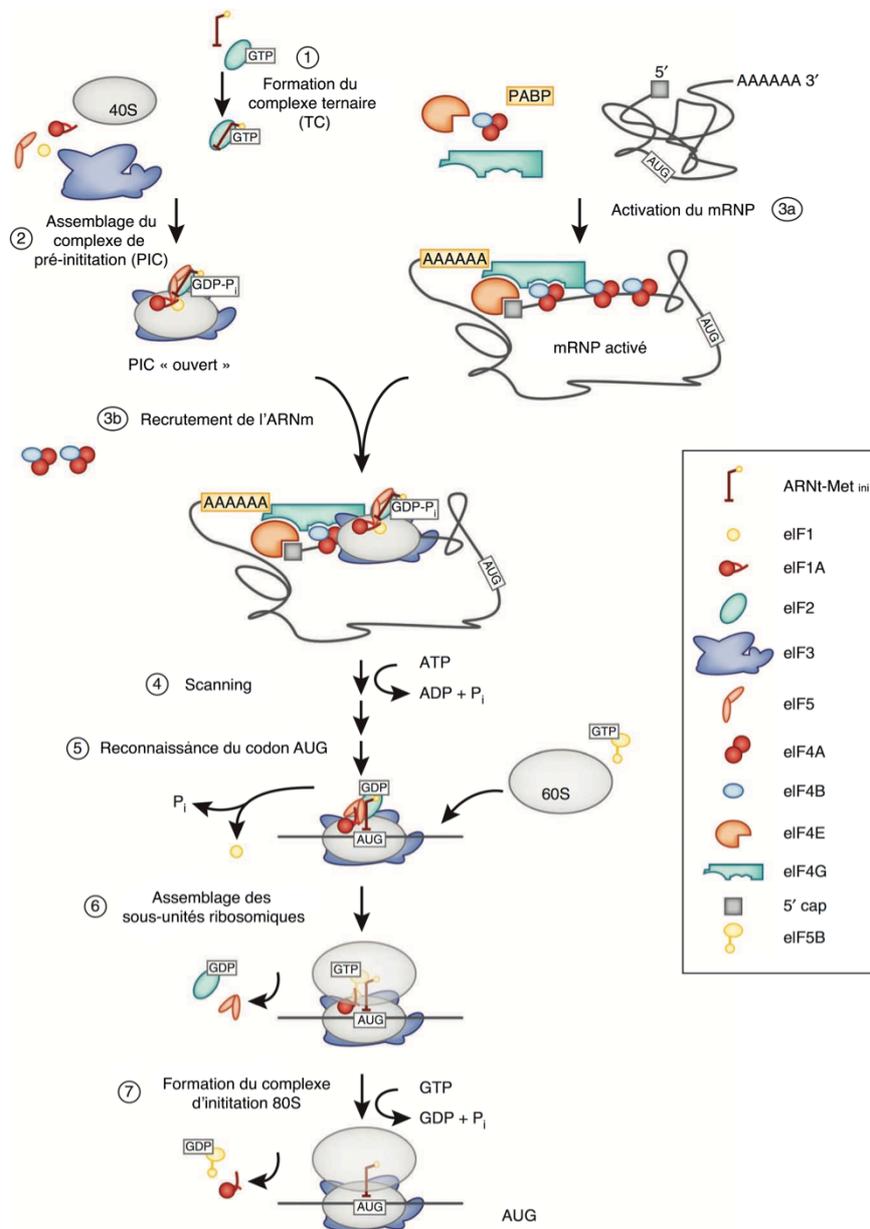
La petite sous-unité 40S est composée d'un ARNr 18S et de 33 protéines dont 14 sont universellement conservées. La structure générale comprend la tête, la plateforme et le corps. La tête possède un bec et le corps une épaule. Dans la partie inférieure du corps sont présentes deux régions désignées comme pied droit et pied gauche (Figure 7) <sup>20</sup>. Le point de contact entre la tête et le corps forme un sillon dans lequel l'ARNm se loge pour y être décodé par la boucle codon/ anticodon. La 40S est également la première des deux sous-unités à être recrutée lors du processus d'initiation de la traduction.

La grande sous-unité 60S est composée de l'ARNr 28S, 5S et 5,8S et de 47 protéines dont 18 sont universellement conservées. Elle contient le centre d'activité peptidyl-transférase <sup>22</sup> qui catalyse la liaison peptidique entre la chaîne en cours de synthèse portée par l'ARNt au site P et l'acide aminé porté par l'ARNt au site A.

La première structure du ribosome obtenue en haute résolution par cristallographie démontrait que la catalyse de la liaison peptidique était réalisée exclusivement par l'ARNr, permettant ainsi d'identifier le ribosome comme un ribozyme <sup>23</sup>. Bien qu'elles soient étudiées depuis plus de 40 ans, le rôle des protéines ribosomales est encore largement incompris <sup>24</sup>. Les avancées rapides de la cryo-microscopie électronique (cryo-EM) ont permis récemment d'obtenir la structure du ribosome eucaryote (80S) et ainsi d'identifier la position des 80 protéines ribosomales associées <sup>20</sup>. Certaines protéines sont universellement conservées, mais d'autres sont spécifiques aux bactéries, aux archaea et aux eucaryotes. Une nouvelle nomenclature a été établie afin d'éviter toute confusion dans le nom de ces protéines (Annexe 1) <sup>25</sup>. Bien qu'elles ne semblent pas participer directement à l'activité catalytique du ribosome, leur contribution est essentielle à la traduction. En plus de leur rôle protecteur de l'ARNr, beaucoup de protéines ribosomales sont essentielles à l'assemblage du ribosome (P1/P2), à l'activité peptidyl-transférase (uL3) et à celle du centre de décodage (uL11 - uS12), à l'appariement de ligands (uL29) et à la sortie de l'ARNm (uL1) <sup>26-30</sup>.

### 3. Les étapes de la traduction

La traduction est un processus au cours duquel les différents acteurs décrits précédemment interagissent pour permettre la synthèse d'une protéine. L'ARNm est traduit simultanément par plusieurs ribosomes, on parle alors de polysomes. Chaque ribosome accomplit un cycle de traduction qui se décompose en 4 étapes :



**Figure 8 : Modèle de la voie classique d'initiation de la traduction eucaryote**

L'initiation commence par la formation du complexe ternaire (TC) contenant eIF2-GTP et l'ARNt initiateur<sup>22</sup>. Le complexe ternaire est recruté sur la sous-unité 40S avec l'aide des eIF 1, 1A, 3 et 5 pour former le complexe de pré-initiation PIC<sup>22</sup>. Pendant ce temps, l'ARNm est préparé par les facteurs eIF4 et PABP pour former un mRNP activé (3a), qui est ensuite recruté sur le PIC (3b). Une fois lié à l'extrémité 5' de l'ARNm, le PIC scanne l'ARNm pour localiser le codon d'initiation (AUG)<sup>22</sup>. La reconnaissance du codon est suivie de la libération d'eIF1 et la conversion de eIF2-GTP en eIF2-GDP, ce qui stoppe le processus de scanning (5). eIF2-GDP et eIF5 se dissocient, ouvrant la voie à eIF5B qui va permettre le recrutement de la sous-unité 60S (6). La fixation de la 60S est suivie de l'hydrolyse du GTP par eIF5B et de la dissociation du facteur d'initiation pour former le complexe d'initiation 80S (IC)<sup>22</sup>. *Figure adaptée de Aitken et al, 2012*<sup>31</sup>

1. L'initiation au cours de laquelle la petite sous-unité ribosomique se lie à l'ARNm, scanne et se positionne sur le codon d'initiation. La 60S s'associe alors avec la 40S pour former le ribosome.

2. L'élongation correspond à la lecture de l'ARNm codon par codon et à la synthèse de la protéine via la catalyse des liaisons peptidiques entre les acides aminés portés par les ARNt.

3. La terminaison, qui se produit lorsqu'un codon stop entre au site A du ribosome, entraînant la libération de la chaîne polypeptidique.

4. Le recyclage, au cours duquel les sous-unités ribosomiques se dissocient et deviennent disponibles pour un nouveau cycle de traduction.

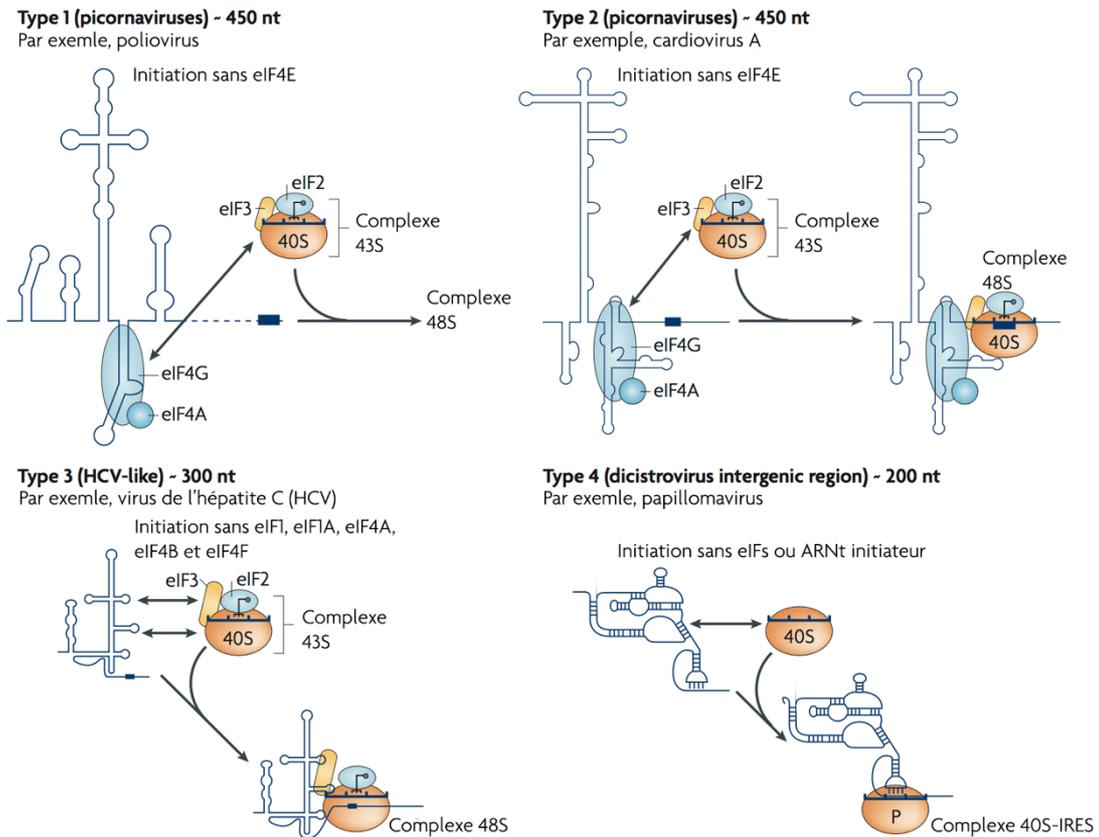
### A. L'initiation

Chez les eucaryotes, l'initiation de la traduction est un processus complexe et l'étape la plus étudiée du cycle de la traduction. Elle permet l'assemblage de la grande sous-unité (60S) et de la petite sous-unité (40S) en un ribosome (80S) fonctionnel contenant un ARNt-méthionine initiateur positionné sur le codon de départ (AUG) de l'ARNm, au site P du ribosome. Ceci est facilité et coordonné par au moins 13 facteurs d'initiations (eIFs)<sup>32</sup>.

L'initiation débute avec la formation d'un complexe ternaire (TC) composé de l'ARNt-Met, eIF2 et de GTP (Figure 8). Les facteurs eIF 1, 1A, 3 et 5 se lient à la sous-unité 40S libre et induisent une conformation « ouverte » de la 40S. Cela permet au complexe ternaire de s'y lier facilement et de former le complexe de préinitiation (PIC)<sup>33</sup>.

En parallèle, l'ARNm est préparé pour pouvoir être recruté par le complexe de préinitiation. Pour cela, la coiffe située en 5' est reconnue par eIF4E qui vient s'y fixer et ancre l'ARN hélicase eIF4A dans la région 5'UTR de l'ARNm. Cette hélicase, liée à eIF4B et eIF4G va permettre d'ouvrir les structures secondaires présentes dans cette région et facilite ainsi la liaison de l'ARNm au complexe de préinitiation. eIF4G rassemble également les extrémités 5' et 3' de l'ARNm par l'intermédiaire de ses interactions avec eIF4E et les PABP situées sur la queue poly(A). Cette boucle « fermée » interviendrait dans la liaison avec le PIC et permettrait également le couplage des étapes de terminaison de la traduction et de recyclage<sup>31</sup>.

Une fois chargé sur l'ARNm, le complexe de préinitiation « ouvert » scanne la région 5' UTR de l'ARNm base par base pour localiser le codon de départ. La présence de l'ARNt-Met initiateur au site P et des facteurs d'initiations eIF 1, 2 et 5 va permettre d'identifier par complémentarité le premier codon AUG. Sa reconnaissance dépend toutefois de plusieurs facteurs, notamment la présence d'une séquence consensus au



**Figure 9 : Modèle de l'initiation de la traduction médiée par les IRES**

Les sites d'entrée interne au ribosome (IRES) sont classés selon 4 types, en fonction de leur mécanisme d'initiation. Les virus de type 1 et 2 requièrent la fixation du complexe 43S sur eIF4G. Les virus de type 3 n'ont besoin que du complexe 43S qui va initier directement sur l'IRES. Alors que les virus de type 4 n'ont besoin ni de facteurs d'initiations ni d'ARNt initiateur. *Figure adaptée de Jackson et al., 2010* <sup>33</sup>

niveau du codon AUG (5'-(A/G)NNAUGG-3') et l'absence de structures secondaires à proximité.

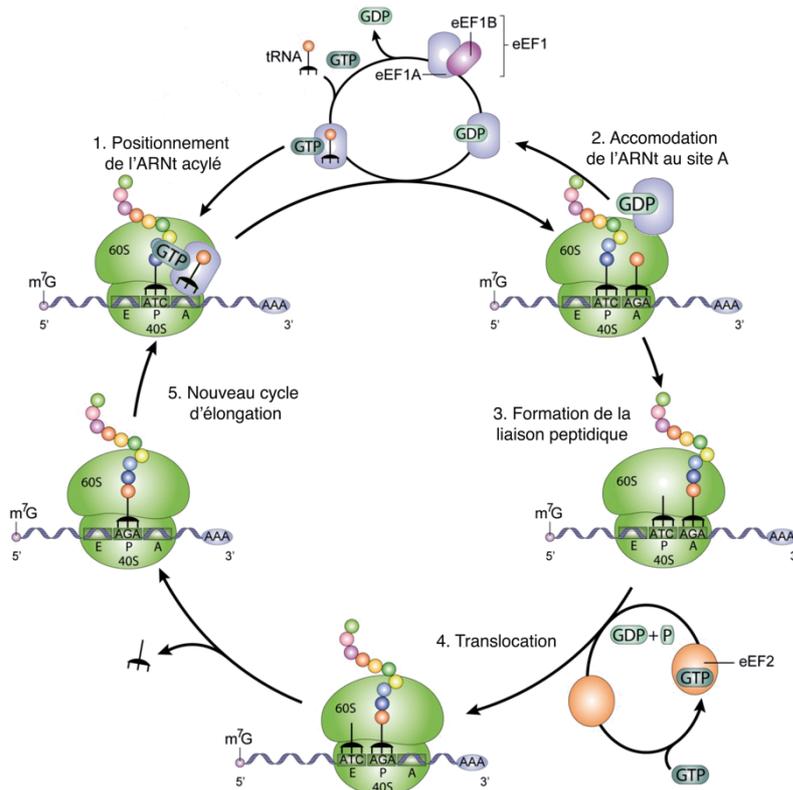
Une fois le codon d'initiation reconnu, eIF5 hydrolyse le GTP fixé à eIF2, ce qui provoque la dissociation d'eIF1 et la conformation « fermée » du complexe de préinitiation. Cette nouvelle conformation stabilise l'interaction de PIC avec l'ARNm et l'ARNt-Met initiateur et arrête le processus de scanning <sup>34</sup>.

A cette étape, eIF2-GDP et eIF5 sont dissociés. eIF6, un facteur anti-associatif lié à la grande sous-unité empêchant la 60S de se lier à la 40S serait détaché par le complexe protéique SBDS-eIF1 <sup>35,36</sup>. Une seconde GTPase liée à la grande sous-unité 60S, eIF5B, collabore avec eIF1A pour permettre l'assemblage des deux sous unités et former le ribosome 80S. eIF1A est alors relargué, ce qui libère la voie au ribosome pour entrer dans sa phase d'élongation.

Bien que la plupart des ARNm sont traduits via le mécanisme d'initiation canonique de la traduction, il existe d'autres mécanismes d'initiations dépendantes et indépendantes de la coiffe <sup>37</sup>.

C'est notamment le cas des ARNm ayant une courte région non traduite en 5'. Ces derniers ont été retrouvés chez les eucaryotes tels que *Saccharomyces cerevisiae* et chez les animaux. La traduction de ces ARNm requiert à la fois la présence de la coiffe et d'une séquence consensus d'initiation de la traduction de ces courts 5'UTR (TISU) (SAASAUGGCGGC) et se fait indépendamment de l'étape de scanning. Contrairement à l'initiation canonique, la formation du complexe 48S (PIC + eIF4F) serait immédiatement suivie du positionnement de la 60S sur le codon AUG et du détachement de eIF4F de la coiffe, pour permettre au codon AUG d'être correctement positionné au site P du ribosome. Cette initiation serait également facilitée par la présence de eIF1 et eIF4G1 <sup>38</sup>.

Il existe également un autre mécanisme d'initiation, cette fois indépendant de la coiffe, qui repose sur la présence de séquences particulières appelées IRES (Internal Ribosome Entry Site). Ces séquences forment une structure secondaire mimant l'ARNt initiateur. Les IRES sont de taille variable et permettent le recrutement du ribosome directement sur le codon d'initiation. Elles sont classées selon quatre types : les IRES de type 1 et 2 où l'initiation se fait indépendamment d'eIF4E, mais requiert la présence d'eIF4G, les IRES de type 3 où seuls eIF2 et eIF3 sont requis et les IRES de type 4 où l'initiation se fait indépendamment de tous les facteurs d'initiations et de l'ARNt initiateur (Figure 9). Elles ont été découvertes chez les virus <sup>39</sup>, qui les utilisent afin de détourner la machinerie cellulaire à leur profit. Il existe aussi des IRES cellulaires <sup>40</sup>, qui permettent par exemple la synthèse de protéines essentielles dans des cas de stress



**Figure 10 : Modèle du cycle d'élongation de la traduction eucaryote**

Le cycle d'élongation débute par la fixation de l'ARNt aminoacylé au site A/T du ribosome. Le facteur EF1A est alors relargué et l'ARNt s'accommode au site A. La liaison peptidique se forme entre la chaîne en cours de synthèse portée par l'ARNt au site P et l'acide aminé porté par l'ARNt au site A, avant l'étape de translocation des ARNt aux sites E et P respectivement. *Figure adaptée de Guerrero et al., 2015*<sup>41</sup>

cellulaires, comme l'apoptose ou dans certains cancers, où un grand nombre de facteurs canoniques d'initiation sont inactivés ou indisponibles<sup>42,43</sup>.

## B. L'élongation

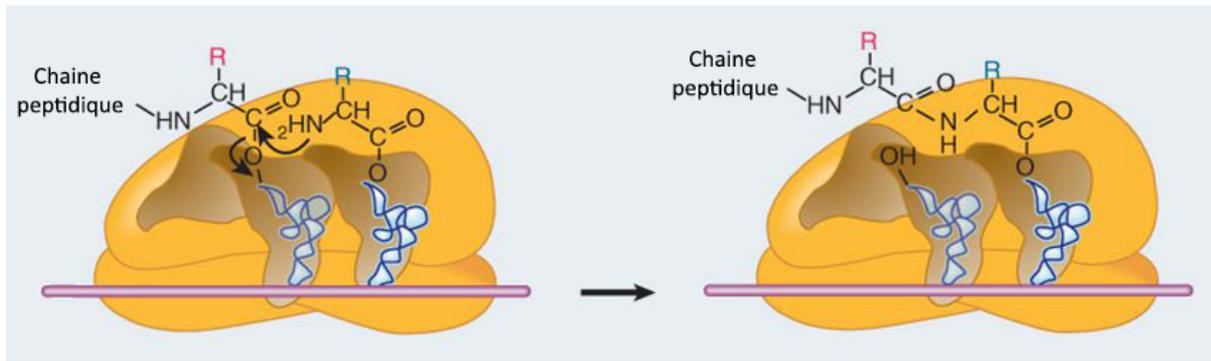
L'élongation consiste en la sélection des aminoacyl-ARNt correspondants aux codons de l'ARNm présents dans le site A du ribosome, puis la formation des liaisons peptidiques entre les acides aminés afin de synthétiser la protéine au cours de cycles successifs (Figure 10). Chaque cycle peut être découpé en 4 étapes principales :

- Le positionnement de l'ARNt aminoacylé en configuration A/T (A pour le site A et T pour EF-Tu)<sup>44</sup>.
- L'accommodation de l'ARNt au site A du ribosome
- La formation de la liaison entre le peptide et le nouvel acide aminé
- La translocation des ARNt du site A au site P, du site P au site E

### a. Positionnement et accommodation

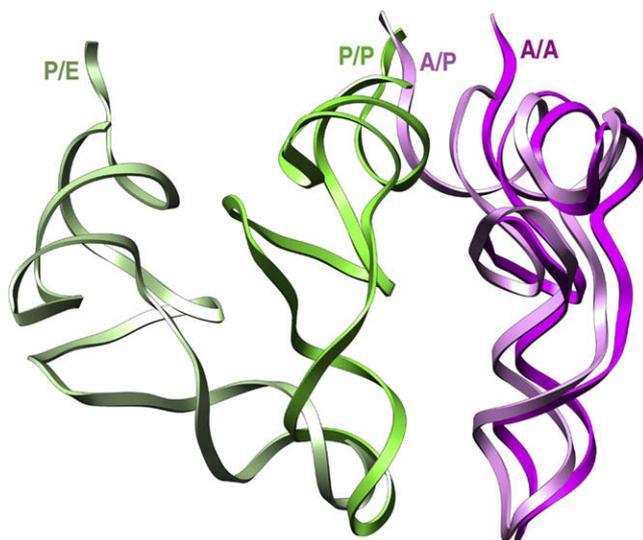
Au moment de débiter le premier cycle d'élongation, le complexe d'initiation contient l'ARNt-Met initiateur et le codon AUG libre au site P du ribosome. La prochaine étape de la traduction correspond à la formation du complexe ternaire avec la fixation du facteur eEF1A-GTP (eucaryote Elongation Factor 1A) sur l'ARNt aminoacylé. Ce complexe se positionne partiellement au site A du ribosome de telle sorte que la boucle anticodon de l'ARNt entre en contact avec le codon AUG. L'appariement codon/anticodon des 3 bases de l'ARNt est alors vérifié par le ribosome pour maintenir une fidélité optimale de la traduction. L'entrée d'un ARNt cognat et certains proche-cognat déclenche l'hydrolyse du GTP par eEF1A, entraînant son relargage et l'accommodation de l'ARNt aminoacylé au site A du ribosome. Le facteur eEF1A-GDP est recyclé en eEF1A-GTP grâce au facteur eEF1B<sup>45</sup>. La fixation de l'ARNt cognat au site A induit un changement de conformation de la petite sous-unité qui passe d'un état « ouvert » à un état « fermé ».

Les dernières études sur les règles de décodage montrent que la validation de l'appariement codon-anticodon par le ribosome serait déterminée par l'encombrement stérique et la géométrie globale de l'appariement plutôt que par le nombre de liaisons hydrogène mises en jeu<sup>46,47</sup>. Les mécanismes de discrimination des ARNt cognat par rapport aux ARNt proche-cognat ou non-cognat seront précisés et discutés dans le chapitre III. *Fidélité de la traduction.*



**Figure 11 : Formation de la liaison peptidique**

Lorsqu'un ARNt est accommodé au site A du ribosome, la chaîne peptidique portée par l'ARNt au site P est transférée sur l'acide aminé au site A. Ce transfert est réalisé via une attaque nucléophile du carbone de la liaison ester de l'ARNt peptidyl par le groupement NH<sub>2</sub> de l'acide aminé porté par l'ARNt au site A. *Figure adaptée de Lewin B., 2007*<sup>48</sup>.



**Figure 12 : Comparaison de la position et de l'orientation des ARNt entre l'état classique et l'état hybride**

Suite au transfert de la chaîne peptidique, les ARNt aux sites A et P qui sont dans l'état classique (A/A ; P/P) entrent dans l'état hybride (A/P ; P/E). *Figure issue de Agirrezabala et al. 2008*<sup>49</sup>.

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

Une fois l'accommodation réalisée, le transfert de la chaîne peptidique du site P sur l'acide aminé au site A est immédiat. Cette étape permet la formation de la liaison peptidique par le Centre Peptidyl Transférase <sup>22</sup> (Figure 11). L'ARNt au site A est à présent un ARNt peptidyl et l'ARNt au site P est déacylé. Ces nouvelles propriétés vont induire un mouvement d'encliquetage (ratcheting) des deux sous-unités ribosomiques et déclencher le mouvement des ARNt dans l'état hybride A/P et P/E (Figure 12), avec l'extrémité du bras accepteur des ARNt dans les sites E et P et la boucle anticodon restant respectivement dans les sites P et A <sup>49</sup>. La translocation de l'ARNt déacylé du site P vers le site E et du peptidyl-ARNt du site A vers le site P nécessite l'hydrolyse d'une molécule de GTP par le facteur d'élongation eEF2. Les changements conformationnels d'eEF2 induiraient alternativement le déblocage du ribosome permettant le déplacement des ARNt et de l'ARNm, puis le blocage des sous-unités dans l'état de post-translocation, lorsqu'eEF2 est relargué <sup>50</sup>. A cette étape, le site E est occupé par l'ARNt déacylé, le peptidyl-ARNt se trouve dans le site P et le site A devient libre pour accueillir un nouvel ARNt aminoacylé. Le cycle d'élongation se répète jusqu'à l'entrée d'un codon de terminaison dans le site A du ribosome.

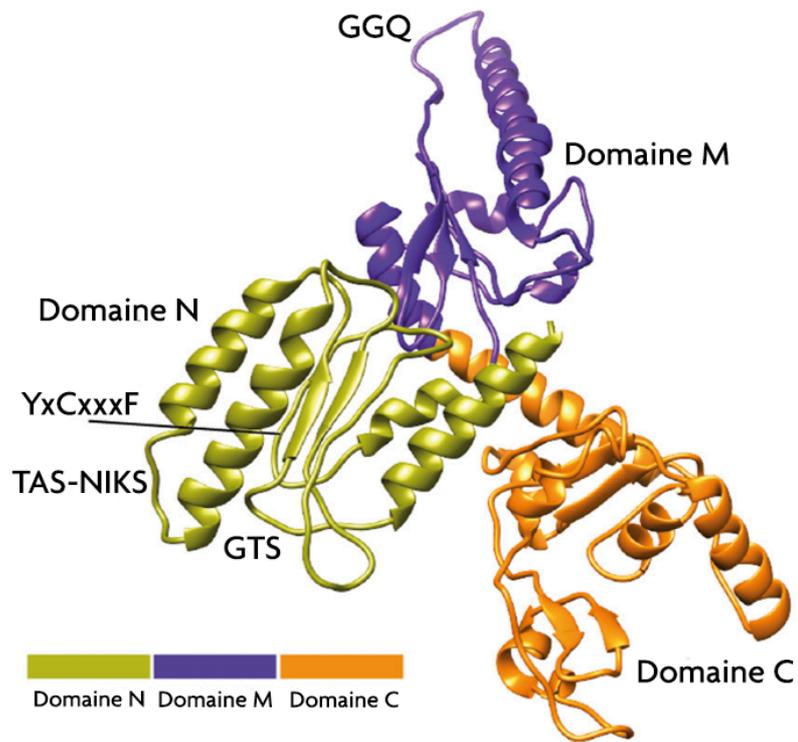
### **C. La terminaison**

La terminaison de la traduction intervient lorsque le ribosome rencontre un codon stop. Le code génétique standard contient trois codons de terminaison (UAA, UAG et UGA). Toutefois, certains organismes comme les algues vertes et les protozoaires ont un répertoire plus restreint. En effet, chez ces algues, seul le codon UGA permet de mettre fin à la traduction (UAA et UAG étant réaffectés en tant que codons glutamine). Tandis que chez les protozoaires, les codons UAA et UAG sont conservés en tant que codon stop, alors que UGA est réaffecté en codon tryptophane <sup>4</sup>. Le processus de terminaison consiste en la reconnaissance du codon stop au site A du ribosome par un complexe de terminaison formé d'un facteur de terminaison de classe I et d'un facteur de terminaison de classe II.

#### *a. Les facteurs de terminaison*

##### **★ Le facteur de terminaison de classe I eucaryote eRF1**

Les archées et les eucaryotes possèdent un seul facteur de classe I, respectivement appelés aRF1 et eRF1, qui reconnaissent les trois codons stop <sup>51</sup>. Ce facteur, codé par le



**Figure 13 : Structure cristallographique du facteur de terminaison eRF1 humain**

Le facteur de terminaison eRF1 est constitué de trois domaines distincts : le domaine N (jaune-vert), le domaine M (violet) et le domaine C <sup>52</sup>. Les motifs TAS-NIKS, GTS et YxCxxxF impliqués dans la reconnaissance du codon stop sont indiqués à l'extrémité du domaine N-terminal, le motif GGQ impliqué dans l'hydrolyse du peptide néo-synthétisé est indiqué à l'extrémité du domaine central. *Figure adaptée de Matheisl et al, 2015* <sup>53</sup>

gène *ETF1* est essentiel et très conservé dans le règne eucaryote, reflétant l'importance du mécanisme dans lequel il est impliqué <sup>54</sup>. La protéine eRF1 humaine possède 100% d'identité avec les protéines souris et lapin, et 68% avec *S. cerevisiae* (UniProtKB P12385). La protéine humaine est constituée de 436 acides aminés (UniProtKB P62495) <sup>55</sup> et adopte une structure tertiaire en forme de Y, mimant ainsi la conformation d'un ARNt. L'obtention de la première structure cristallographique du facteur eRF1 humain a permis d'identifier trois domaines distincts : le domaine N (N-terminal), le domaine M (central) et le domaine C (C-terminal) <sup>56</sup> (Figure 13).

- Le domaine 1 ou N (1-142) est analogue au bras anticodon de l'ARNt. Il est impliqué dans la reconnaissance des codons stop, notamment grâce au motif très conservé NIKS (Asn<sub>61</sub>, Ile<sub>62</sub> Lys<sub>63</sub>, Ser<sub>64</sub>) qui se situe dans une boucle aux extrémités de deux hélices <sup>57</sup>.

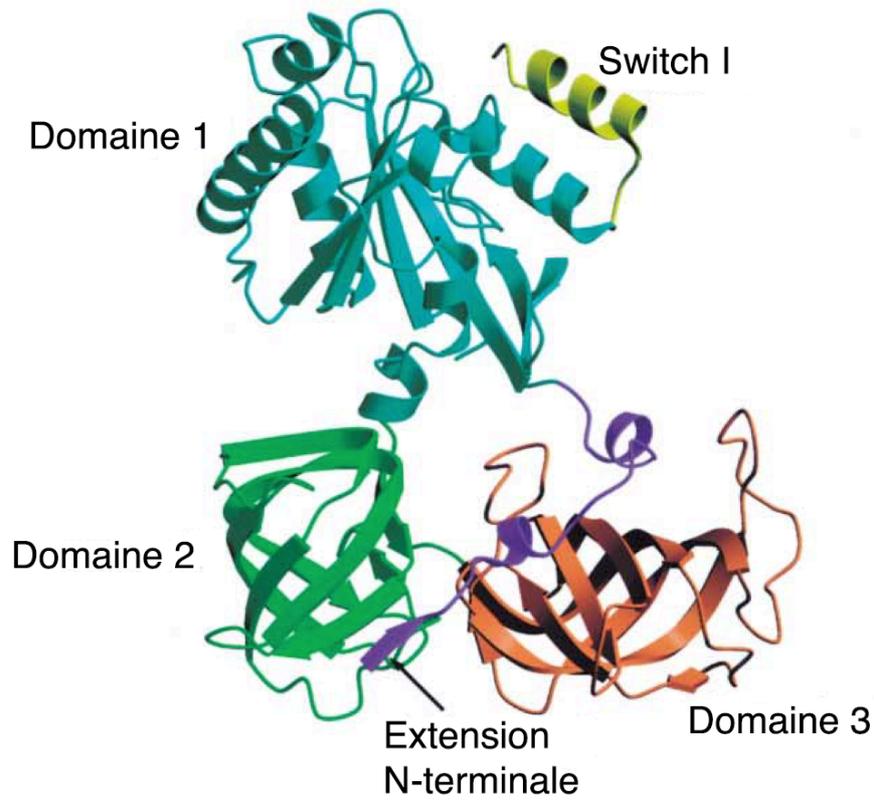
- Le domaine 2 ou M (143-272) présente des analogies avec le bras accepteur de l'ARNt. Il est impliqué dans l'hydrolyse de la liaison ester reliant le dernier ARNt peptidyl à la protéine synthétisée. Cette hydrolyse est déclenchée par le motif universellement conservé Glycine183-Glycine184-Glutamine185 (GGQ). Ce motif vient se placer au niveau du Centre Peptidyl-Transférerase (PTC, sous-unité 60S) et libère la protéine grâce après hydrolyse de la liaison ester. Le domaine M serait également capable de se lier directement à eRF3, ce qui stimulerait la fixation du GTP sur le complexe eRF1-eRF3 <sup>58</sup>.

- Le domaine 3 ou C (273-436) est impliqué dans des interactions protéiques avec les partenaires d'eRF1, en particulier avec le facteur de terminaison eRF3. <sup>57</sup>. Cette interaction serait nécessaire à l'activation de l'activité GTPase d'eRF3. Le domaine C comprend également un mini-domaine impliqué dans l'interaction avec l'ARNr 18S.

La protéine eRF1 possède des modifications post-traductionnelle sur chacun de ses domaines. Elle est hydroxylée au niveau de la lysine 63 du motif NIKS <sup>59</sup>, méthylée sur la glutamine du motif GGQ <sup>60</sup>, et phosphorylée au niveau des résidus sérine 421 et 432 du domaine C-terminal <sup>61</sup>. A l'inverse de l'hydroxylation et de la méthylation, la phosphorylation ne semble pas avoir d'impact significatif sur l'efficacité de terminaison.

### ★ Le facteur de terminaison de classe II eucaryote eRF3

Le facteur de classe II eRF3 est une GTPase qui forme un complexe avec le facteur eRF1 hors du ribosome. Il stimule ainsi l'activité d'eRF1 et intervient dans la libération de la chaîne polypeptidique. Aucun facteur de classe II n'a été identifié chez les archées, mais ce rôle serait joué par le facteur d'élongation EF1 $\alpha$  <sup>62</sup>. Le facteur de terminaison eRF3 est également appelé GSPT1 (G1 to S Phase Transition), car il a été initialement décrit comme étant essentiel pour la transition G1/S du cycle cellulaire.



**Figure 14 : Structure cristallographique du domaine C du facteur de terminaison eRF3 *S.pombe***

Structure des domaines 1 (cyan), 2 (vert) et 3<sup>52</sup>, d'eRF3 C et son extension N-terminale (magenta). Le domaine 1 possède l'activité GTPasique, le domaine 2 est impliqué dans la fixation de la PABP et le domaine 3 dans l'interaction avec eRF1. *Figure adaptée de Kong et al., 2004*<sup>63</sup>

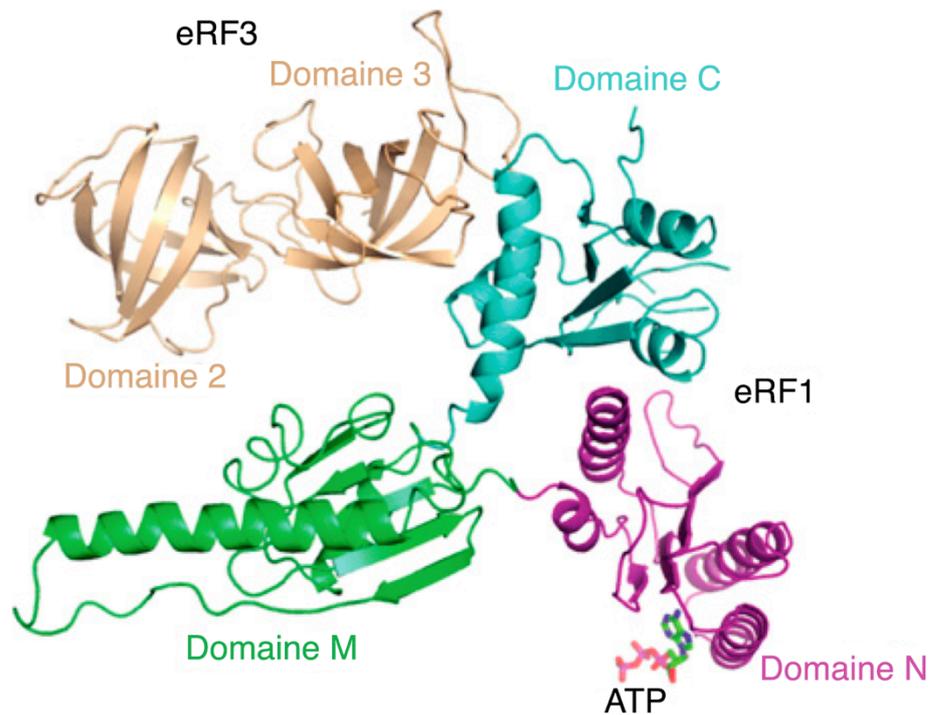
Chez les mammifères, il existe deux gènes d'eRF3 : eRF3a (GSPT1 – UniProtKB P15170) et eRF3b (GSPT2 – UniProtKB Q8IYD1). Les deux protéines humaines sont identiques à 94%. Entre la souris et l'Homme, les deux eRF3 sont très similaires (95% pour eRF3a (UniProtKB Q8R050) et 89% d'identité pour eRF3b (UniProtKB Q149F3). eRF3a représente la forme majoritaire d'eRF3 dans la majorité des tissus. eRF3b, qui peut compléter eRF3a dans son rôle de terminaison de la traduction, est particulièrement peu exprimé dans la plupart des tissus et ne semble pas jouer un rôle majeur dans la terminaison. Les protéines eRF3a et eRF3b diffèrent par leur domaine N-terminal. Dans le gène humain GSPT1, la séquence contient un codon GGC glycine, répété 7 à 12 fois, qui est absent du gène GSPT2. Le motif contenant 12 répétitions a été retrouvé chez certains patients atteints de cancers et semble diminuer l'affinité d'eRF3a 12-GGC pour PABP d'un facteur  $10^6$ .

Le facteur eRF3 se lie à eRF1 pour former un complexe stable dans le cytoplasme avant de réaliser la terminaison au sein du ribosome. La formation de ce complexe serait dépendante du niveau d'eRF3a qui serait régulé par le protéasome via la dégradation de la forme libre d'eRF3a<sup>65</sup>. La protéine eRF3 serait également formée de trois domaines distincts dont la taille varie selon les organismes<sup>63</sup>. Chez l'Homme, la structure complète d'eRF3 n'a pas encore été obtenue et seuls les domaines C (C-terminal) et N (N-terminal) d'eRF3 ont été cristallisés à ce jour<sup>66</sup>. Je vais donc me concentrer ici sur la description des trois domaines tels qu'ils sont décrits chez *S. cerevisiae* :

- Le domaine N est celui qui est le moins conservé entre les eucaryotes, il possède des séquences dans lesquelles un acide aminé est retrouvé de manière répétée, la glycine chez les mammifères, la glutamine chez les levures et la lysine chez les euplotes. Chez la levure *S. cerevisiae*, ces répétitions sont impliquées dans le phénomène d'agrégation de type prion conduisant au phénotype [PSI<sup>+</sup>]. Le domaine N n'est pas indispensable pour réaliser la terminaison de la traduction ni pour la viabilité<sup>67</sup>, il est néanmoins requis pour l'interaction avec la PABP<sup>68</sup>.

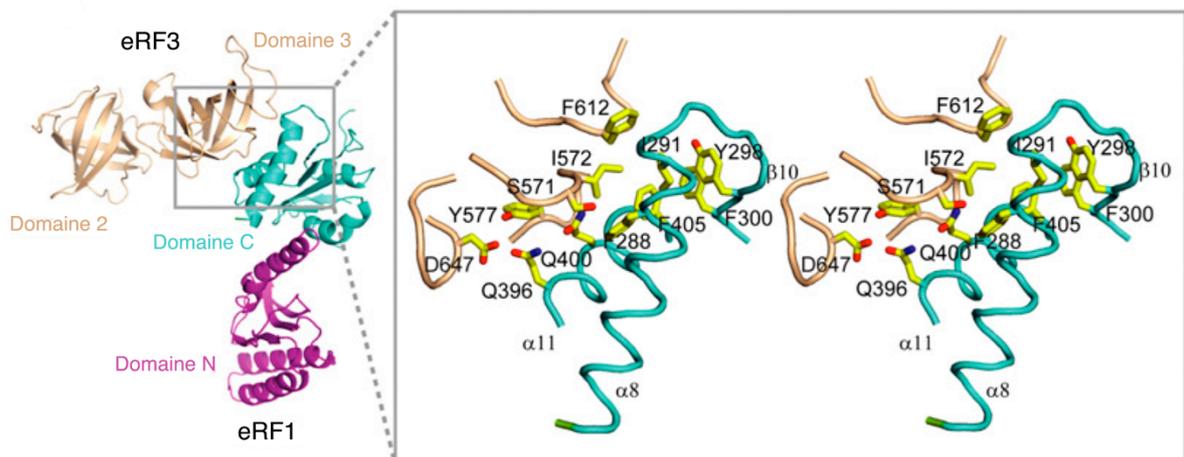
- Le domaine M est la région médiane entre le domaine N et le domaine C et dont le seul rôle connu à ce jour est de moduler la solubilité et le passage à l'état prion<sup>69</sup>.

- Le domaine C est le domaine le mieux conservé et présente de fortes similitudes avec le facteur d'élongation eucaryote eEF1A. Ce domaine est suffisant pour réaliser la terminaison de la traduction *in vitro* et *in vivo* chez *S. pombe* (Kong et al., 2004). Il se décompose en 3 sous domaines : Le domaine 1 ou G qui possède l'activité GTPasique, le domaine 2, impliqué dans la fixation de la PABP et le domaine 3, impliqué dans l'interaction avec eRF1 (Figure 14).



**Figure 15 : Structure cristallographique du complexe eRF1-eRF3 humain**

Structure des domaines N (magenta), M (vert) et C (cyan) d'eRF1 et des domaines 2 et 3 en C-terminal d'eRF3 (dorée). Les facteurs de terminaison interagissent via leur domaine C-terminal pour former le complexe de terminaison. *Figure adaptée de Cheng et al., 2009*<sup>70</sup>



**Figure 16 : Interactions entre eRF1-eRF3**

A gauche, le complexe d'eRF1 de la levure *Saccharomyces pombe* et des domaines 2 et 3 d'eRF3 humain. A droite, vue stéréo de l'interface entre eRF1 et eRF3. Les résidus en interaction sont indiqués en jaune en représentation bâtonnets et correspondent principalement à des acides aminés hydrophobes. *Figure issue de Cheng Z. et al., 2009*<sup>70</sup>

### ★ Les interactions entre eRF1 et eRF3

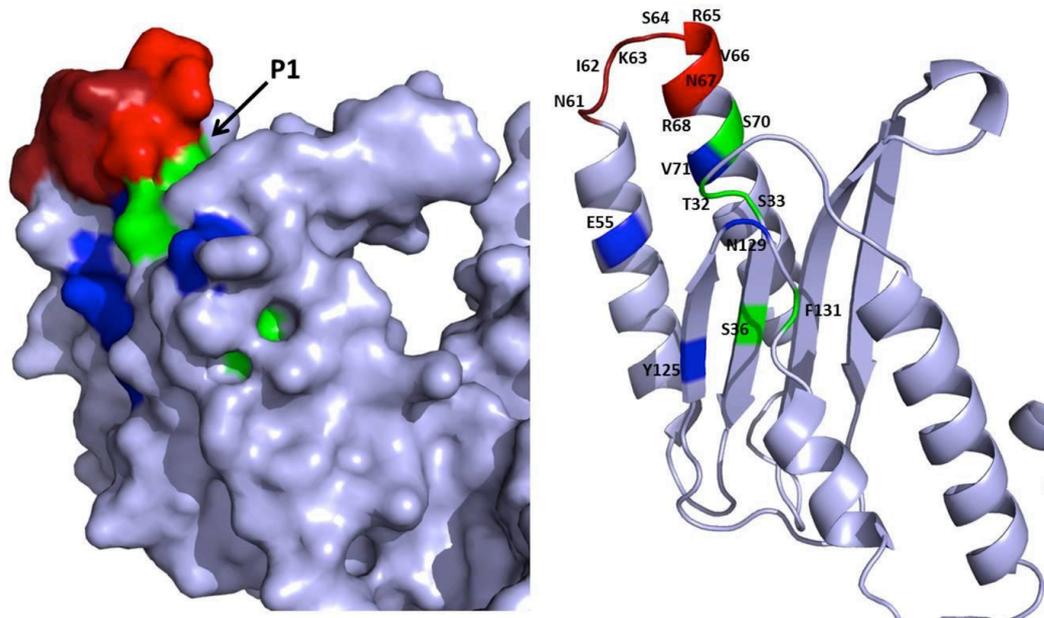
Les facteurs de terminaison eRF1 et eRF3 interagissent via leur domaine C-terminal pour former le complexe de terminaison (Figure 15). Lors de l'entrée de ce complexe dans le ribosome, eRF3 est lié au GTP, on parle de complexe ternaire eRF1/eRF3/GTP. Les domaines M et C d'eRF1 interagissent avec le domaine C-terminal d'eRF3. Cette interface est principalement composée d'interactions hydrophobes (Figure 16). Mais il existe également plusieurs interactions van der Waals entre le domaine C d'eRF1 et le domaine 3 d'eRF3 <sup>70</sup>. Plus récemment, la structure en cryo-EM du complexe de terminaison dans le ribosome mammifère a permis de confirmer ces résultats <sup>66</sup>. L'association d'eRF3 avec eRF1 stabilise la formation du complexe eRF1/eRF3/GTP, stimule l'activité GTPasique ribosome-dépendante et couple la reconnaissance des codons stop à l'hydrolyse de la liaison entre le peptide et l'ARNt dans le centre peptidyl-transférase <sup>71</sup>.

#### *b. Les mécanismes de la terminaison*

### ★ La reconnaissance du codon stop

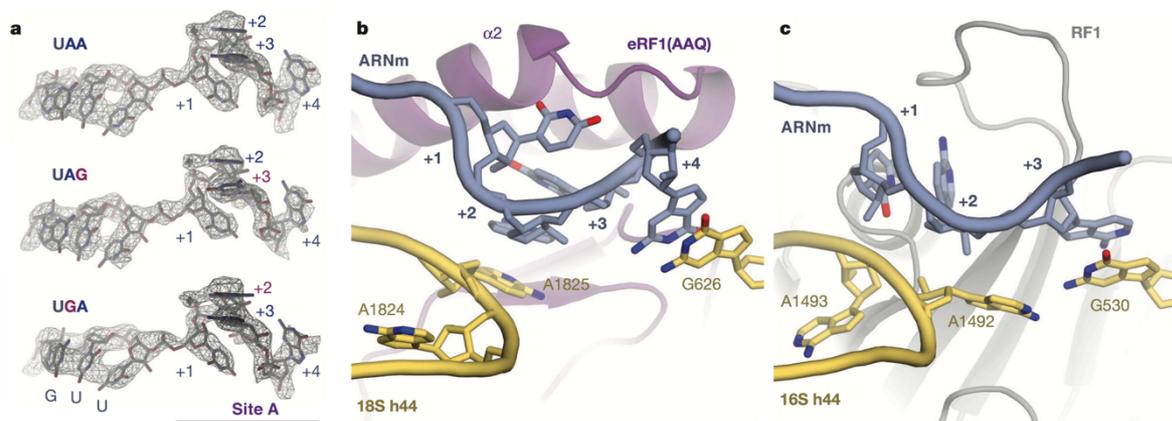
Le mécanisme de reconnaissance du codon stop se fait par l'intermédiaire du facteur de classe 1 eRF1 et en particulier par le domaine N-terminal de la protéine. De nombreuses études, en particulier de mutagenèse dirigée ou aléatoire, ont été menées afin de déterminer les résidus du domaine N-terminal d'eRF1 qui étaient impliqués dans la reconnaissance du codon stop. Certaines régions très conservées ont été identifiées (Figure 13) comme jouant un rôle essentiel dans l'efficacité de terminaison et de reconnaissance du codon stop <sup>72</sup>. Il s'agit des motifs GTS (31-33), TASNKS (58-64) et YxCxxxF (125-131) et de la poche P1 <sup>73</sup>, (numérotation pour la protéine humaine).

La mutagenèse de plusieurs résidus conservés dans la protéine eRF1 a permis de mettre en évidence que les résidus entre les positions 61 et 68, comprenant le NIKS, sont essentiels pour la reconnaissance des 3 codons stop et seraient donc impliqués dans le décodage de la base uridine en 1<sup>ère</sup> position. La lysine 63 (K63) et l'asparagine 61 (N61) joueraient un rôle crucial dans l'efficacité de terminaison en interagissant avec les résidus environnants. En effet, le signal de terminaison serait plus étendu que le codon de terminaison <sup>74</sup>. L'isoleucine 62 (I62) interviendrait dans le maintien de la conformation de la boucle NIKS et de sa chaîne latérale. Les autres résidus R65, N67 et N68 seraient nécessaires au bon positionnement d'eRF1 dans le site A du ribosome. Il a également été suggéré que les acides aminés S33 et S70 <sup>75</sup> participant à la reconnaissance du codon stop UGA, seraient impliqués dans la reconnaissance de la



**Figure 17 : Résidus du domaine N d'eRF1 impliqués dans la reconnaissance de chacune des trois bases du codon stop.**

Les acides aminés impliqués dans le décodage de la 1<sup>ère</sup> base U sont colorés en rouge (N61-R68), les résidus important pour la reconnaissance de la 2<sup>nd</sup>e base sont indiqués en vert (S33, S70, T32, S36, F131), ceux impliqués dans la reconnaissance de la 3<sup>ème</sup> base sont colorés en bleu (E55, Y125, N129). La cavité P1 découverte par notre équipe est indiquée par une flèche noire *Figure issue du mémoire de thèse de Sandra Blanchet, 2014*



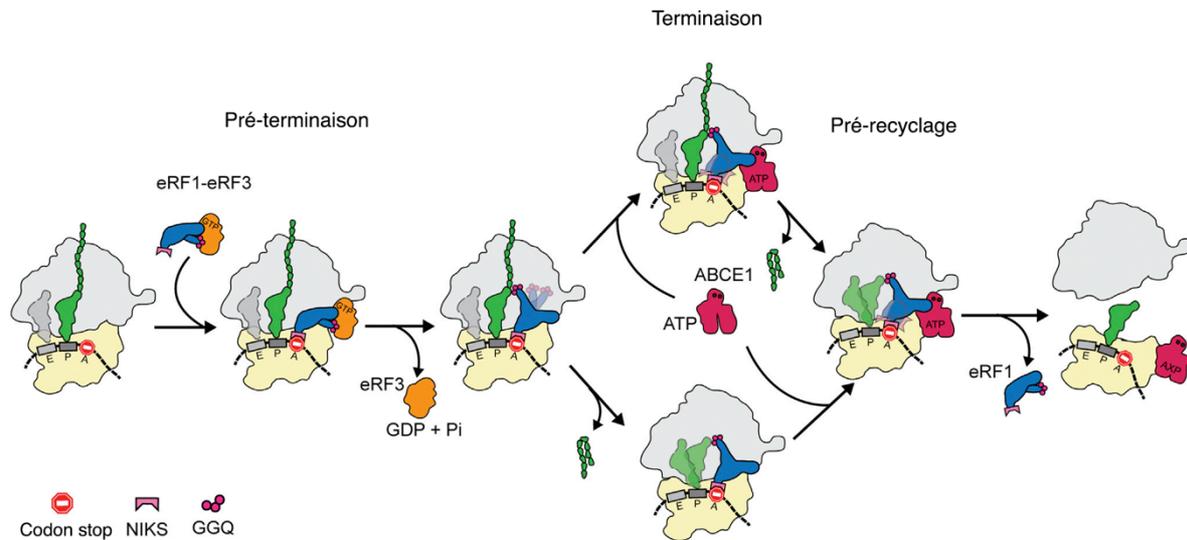
**Figure 18 : Structures des codons stop dans le centre de décodage**

**a** – Cartes des densités de microscopie électronique de l'ARNm pour chacun des trois complexes de terminaison obtenus pour les trois codons stop. La conformation compacte des positions +2 et +3 est retrouvée dans les trois conditions. Le codon Val (GUU) au site P et le codon stop (+1 à +3) et la base +4 au site A sont indiqués. **b** - eRF1 (AAQ) (violet) reconnaît quatre bases de l'ARNm (+1 à +4, couleur ardoise) dans le site A. Les bases +2 et +3 sont empilés et interagissent avec A1825 de l'ARNr, qui est dans l'état « flipped out ». La base en +4 interagit directement avec le résidu G626 de l'ARNr 18S. **c** - Chez les bactéries, RF1 (gris) reconnaît une configuration codon stop plus étendu où la base +3 interagit avec G530 (l'équivalent de G626) de l'ARNr 16S. *Figure adaptée de Brown et al., 2015* <sup>76</sup>

guanine en 2<sup>ème</sup> position. La thréonine en position 32 (incluse dans le motif GTS), la sérine en position 36 et la phénylalanine en position 131 (incluse dans le motif YxCxxxF) seraient impliquées dans la reconnaissance de l'adénine en 2<sup>ème</sup> position du codon. Concernant la 3<sup>ème</sup> position du codon stop, les résidus E55, Y125 (du motif YxCxxxF) et N129 seraient importants pour la reconnaissance de la guanine et la position V71 pour la reconnaissance de l'adénine <sup>54,77</sup> (Figure 17).

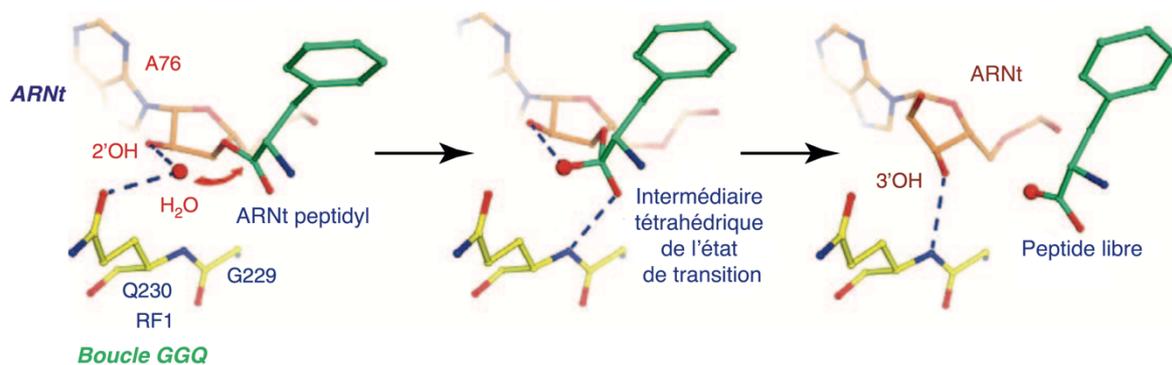
Les structures cryo-EM obtenues récemment par Brown et ses collaborateurs <sup>76</sup> ont confirmé ces résultats et ont montré que le motif NIKS imposait la présence d'une uridine sur la première position du codon présent au site A via des interactions de type Watson-Crick. Les purines (A/G) seraient défavorisées à cette position à cause de leur encombrement stérique et une cytidine ne remplirait pas les conditions nécessaires pour qu'il y ait formation de la liaison hydrogène. Les interactions du motif YxCxxxF et de Glu55 d'eRF1 avec les positions +2 et +3 discriminent les codons sens en autorisant seulement la présence de purines. Le motif GTS adopterait quant à lui deux conformations selon celle adoptée par YxCxxxF, ce qui lui permettrait de discriminer le nucléotide en 3<sup>ème</sup> position. Ainsi, la présence d'une adénosine en deuxième position induirait une conformation du motif GTS où T32 ferait face à la position +3 et pourrait interagir avec une guanine uniquement. Une guanosine en +2 impliquerait une autre conformation du motif GTS qui ne ferait plus face au codon stop. Enfin, les auteurs de ces travaux ont observé la présence de la base +4 dans le site A du ribosome en présence d'eRF1 avant hydrolyse de l'ARNt-peptidyl et ce quelque soit la nature du codon stop. L'entrée de cette base dans le centre de décodage serait possible grâce à la compaction de l'ARNm dans le site A. Ce nucléotide interagirait directement avec le résidu G626 (G530 chez les bactéries) de l'ARNr 18S (Figure 18). La présence d'une purine en +4 stabiliserait cette interaction, ce qui expliquerait la prévalence de purines en position +4 chez les eucaryotes.

Ces récents travaux ainsi que l'obtention de la structure eucaryote du complexe de terminaison <sup>53,76,78</sup> ont donc permis d'affiner le modèle de reconnaissance du codon stop en deux étapes. Le codon de terminaison serait reconnu par le motif NIKS d'eRF1, lors de son entrée au site A du ribosome, associé à eRF3. A cette étape, eRF1 serait quasiment en position A/T. Le(s) nucléotide(s) en position 1 ou 1-2 du codon stop serai(en)t alors « décodé(s) ». A ce stade, deux mécanismes sont envisagés : dans le premier cas, suite à l'hydrolyse du GTP, le facteur eRF3 est dissocié d'eRF1 et remplacé par ABCE1 sur le domaine C d'eRF1. Cette fixation s'accompagne de deux modifications de la conformation d'eRF1 : la première est située au niveau du domaine N, permettant au codon stop d'atteindre la poche P1 où le deuxième et troisième nucléotide du codon stop sont décodés <sup>77</sup> ; la seconde est située dans le domaine central, qui permet au motif GGQ de correctement se positionner dans le PTC et de



**Figure 19 : Modèle de la terminaison de la traduction eucaryote**

Lors de la terminaison de la traduction, le codon stop est situé dans le site A et est reconnu par le complexe ternaire eRF1-eRF3-GTP. Le facteur eRF3 se dissocie après hydrolyse du GTP et permet au domaine central de eRF1 de se loger dans le PTC. Un positionnement correct du motif GGQ situé dans le domaine central de eRF1 peut déjà permettre de libérer la protéine, ce qui conduit à un complexe de terminaison avec l'ARNt déacylé dans le site P ou dans un état hybride P/ E. Dans un autre modèle, la conformation active de eRF1 dans le complexe de pré-terminaison est stabilisée après la liaison de ABCE1. Cela stimule la libération du peptide tandis que le domaine N-terminal de eRF1 est délocalisé. Le décodage est ainsi découplé de la libération du peptide. *Figure adaptée de Preis et al, 2014* <sup>78</sup>



**Figure 20 : Rôle du GGQ dans l'hydrolyse du peptide**

Le processus d'hydrolyse peptidique impliquerait une molécule d'eau et les positions 2' et 3' de la base A76 du CCA à l'extrémité de l'ARNt au site P. Les résidus G229 et Q230 du motif GGQ du facteur RF1 sont représentés en jaune, le peptide est en vert, l'ARNt au site P est en orange. *Figure adaptée de Klaholz, 2011* <sup>79</sup>

libérer la protéine. Dans le second cas, ABCE1 n'est pas recruté après dissociation d'eRF3, mais après libération de la protéine, pour permettre la dissociation du complexe ribosomique (Figure 19).

### ★ La libération de la protéine néosynthétisée

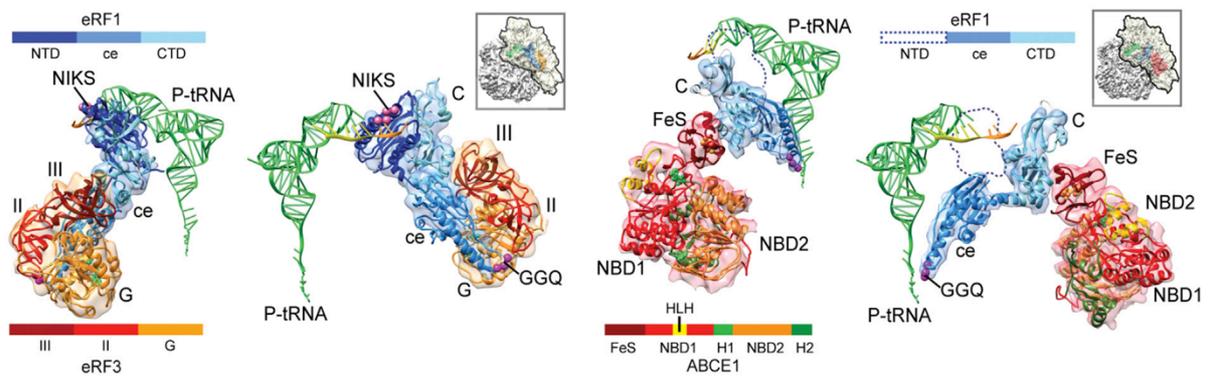
Une fois positionnée au site A du ribosome, le domaine M d'eRF1, par son interaction avec le domaine G d'eRF3, stimule son activité GTPase (Cheng et al., 2009). L'hydrolyse du GTP par eRF3 induit une modification de la conformation du domaine G de la protéine. Suite à ces changements, le domaine M d'eRF1 perdrait son interaction avec eRF3, ce qui permettrait le positionnement du motif GGQ dans le centre peptidyl-transférase<sup>22 22</sup>. Au cours de l'élongation, les bases de l'ARNr 28S protègent la liaison ester d'une attaque nucléophile. Dans le cas de la terminaison, le motif GGQ du facteur de terminaison eRF1 permet d'éloigner ces bases. Une molécule d'eau va alors hydrolyser la liaison ester et permettre la libération du polypeptide. La spécificité de la réaction d'hydrolyse est contrôlée par la glutamine du motif GGQ en excluant l'entrée dans le site actif des nucléophiles qui seraient plus grands qu'une molécule d'eau<sup>80</sup>. Le groupement amide de la chaîne principale de la glutamine 185 forme une liaison hydrogène avec le 3'-OH du dernier A (généralement en position 76) du motif CCA en 3' de l'ARNt au site P<sup>81</sup> (Figure 20). Suite à l'hydrolyse du GTP, le facteur eRF3 se dissocie et eRF1 est accommodé dans le site A du ribosome. Le facteur ABCE1 est recruté et va aider à la libération de la protéine, avant de procéder à la dissociation d'eRF1 et des sous-unités ribosomiques<sup>82</sup>.

## D. Le recyclage du ribosome

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

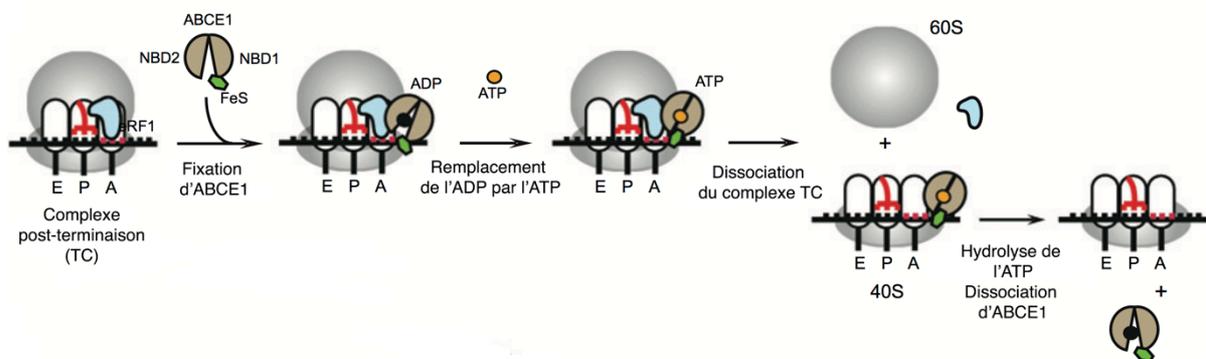
Chez les eucaryotes, la dissociation des sous-unités ribosomiques serait médiée par eRF1 et ABCE1, une protéine de la famille des ATP-Binding Cassette présente chez les eucaryotes et les archées. Les récents travaux de cryo-EM<sup>78,83</sup> ont permis de lever le voile sur la localisation d'ABCE1 dans le complexe de prérecyclage et de recyclage eucaryote. Il a ainsi été montré qu'ABCE1 était lié au domaine C d'eRF1 via son domaine fer-souffre (FeS). Le facteur eRF3 étant également lié au domaine C d'eRF1, la fixation d'ABCE1 nécessiterait le départ d'eRF3 (Figure 21).

Les protéines ABC possèdent deux domaines NBD (Nucleotide-Binding Domain) de fixation aux nucléotides qui leur permettent de convertir l'énergie chimique en mouvements mécaniques. L'activité des protéines ABC est basée sur leur capacité à



**Figure 21 : Structure cryo-EM du complexe de terminaison eRF1-eRF3c et pré-terminaison/recyclage eRF1-ABCE1 avec l'ARNt peptidyl.**

Modèles moléculaires eucaryotes du peptidyl ARNt, eRF1, eRF3c et ABCE1 sur le ribosome. A gauche, le complexe de terminaison comprenant l'ARNt (vert), eRF1 (bleu) et eRF3c (orange-rouge). Les trois domaines de eRF1 (N-terminal (NTD), central (ce) et C-terminal (CTD) et le domaine C-terminal de eRF3 (constitué des domaines 1 (G), 2 (II) et 3 (III)) sont représentés par une couleur distincte. A droite, le complexe de pré-terminaison/recyclage est constitué de l'ARNt (vert), eRF1 (bleu) et ABCE1 (orange-rouge). Seuls les domaines central et C-terminal de eRF1 sont visibles. Les domaines fer-soufre (FeS), domaines de fixations à l'ARN (NBD1 et NBD2) et les charnières (H1 et H2) sont visibles. Le motif NIKS (sphères roses) de eRF1 est positionné à proximité du codon stop<sup>52</sup>. Le domaine central de eRF1 contenant la boucle GGQ (sphères magenta) est situé contre eRF3 dans le complexe de terminaison et basculé vers le PTC dans le complexe de pré-terminaison/recyclage. *Figure issue de Preis et al., 2014*<sup>78</sup>



**Figure 22 : Modèle de dissociation des sous-unités ribosomiques de Barthelme et son équipe**

A la fin de la terminaison, ABCE1 est recruté et se lie à eRF1. La fixation d'ATP sur ABCE1 provoque un changement de conformation vers un état « fermé » ce qui déclenche la dissociation du ribosome et la libération de eRF1. Enfin, l'hydrolyse de l'ATP déclenche la dissociation de ABCE1 de la petite sous-unité. *Figure adaptée de Jackson et al, 2012*<sup>71</sup>

subir des changements conformationnels cycliques déterminés par les modifications des positions relatives des NBD, qui dépendent de l'état de liaison de la protéine. Comparativement aux autres protéines ABC, ABCE1 possède trois caractéristiques uniques : un motif hélice-boucle-hélice <sup>84</sup> dans le domaine NBD1 ; deux régions charnières (H1 et H2), et un domaine fer-souffre (FeS) <sup>85</sup> (Figure 21).

Le mécanisme de dissociation des sous-unités ribosomique est encore mal compris à ce jour. Les deux modèles défendus sont basés sur le changement de conformation d'ABCE1. Suite à l'hydrolyse du GTP par eRF3, ce dernier se dissocierait d'eRF1. ABCE1 serait alors recruté et interagirait avec eRF1 via son domaine fer-souffre. Dans le premier modèle, la fixation d'ATP sur ABCE1 provoquerait un changement de conformation vers un état « fermé ». Cela aurait pour conséquence de déplacer le domaine fer souffre et conduirait à la dissociation du ribosome et d'eRF1. L'hydrolyse de l'ATP serait nécessaire pour relarguer ABCE1 et probablement l'ARNt et l'ARNm grâce à l'aide d'autres facteurs <sup>86</sup> (Figure 22). Le deuxième modèle envisage une action moins directe d'ABCE1. Après fixation à eRF1, ABCE1 se lierait également à une molécule d'ATP qui ne permettrait pas de dissocier le ribosome. C'est l'hydrolyse de l'ATP qui provoquerait l'entrée d'ABCE1 dans un état « ouvert » et qui transmettrait à eRF1 via son domaine FeS le changement de conformation, entraînant une déstabilisation de l'ARNt déacylé et des ponts entre les sous-unités ribosomiques <sup>87</sup> (Figure 23). Dans les deux cas, la dissociation des ribosomes 80S par ABCE1 résulte en la libération de la sous-unité 60S et de la sous-unité 40S associée à l'ARNt déacylé et à l'ARNm.

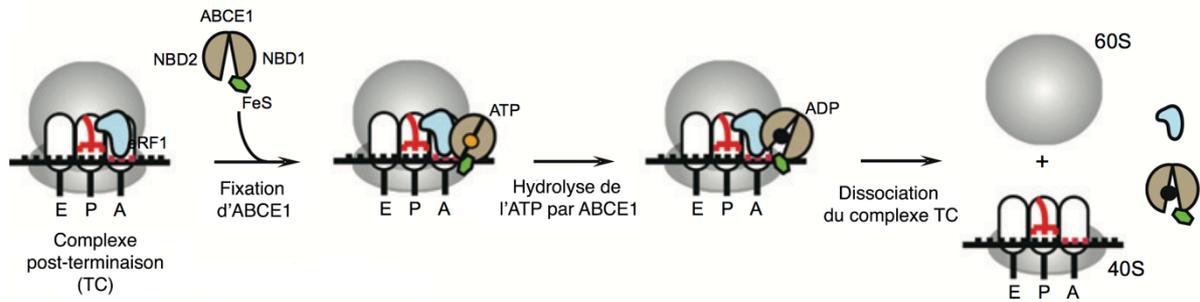
#### *b. Recyclage de l'ARNm, de l'ARNt et des sous-unités ribosomiques*

Après dissociation des sous-unités ribosomique, la sous-unité 60S est prise en charge par le facteur d'anti-association eIF6. L'ARNt déacylé au site P et l'ARNm restent liés à la 40S. La petite sous-unité est alors reconnue par les facteurs d'initiation eIF1, eIF3 et eIF1A ce qui permet la libération de l'ARNt déacylé du site P. L'ARNm est ensuite décroché grâce au recrutement d'eIF3j <sup>88</sup> (Figure 24).

Un autre intervenant, plus récemment identifié, la ligatine serait également capable de dissocier l'ARNt et l'ARNm de la sous-unité 40S par un mécanisme encore mal déterminé <sup>89</sup>.

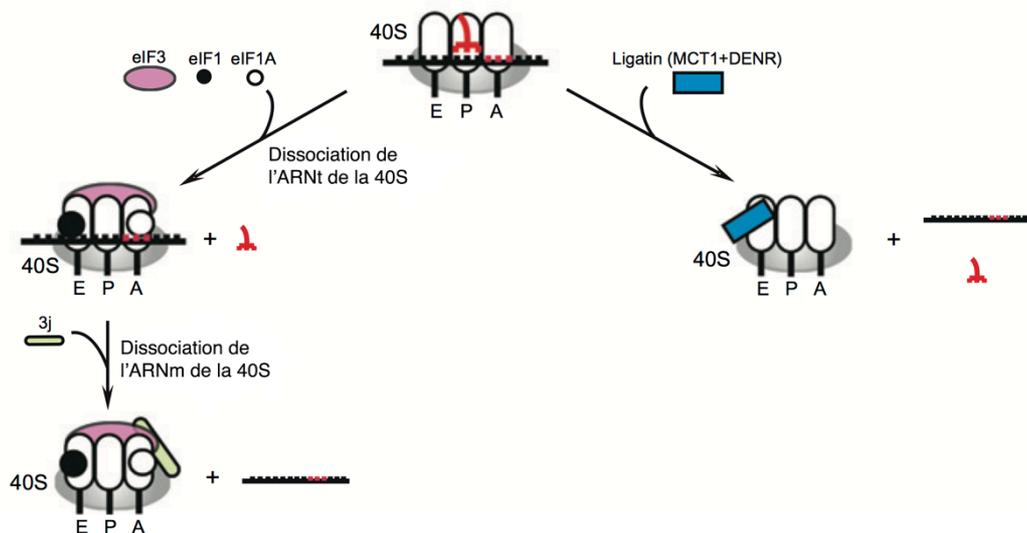
### **E. La réinitiation**

Dans certains cas, les complexes de post-terminaison ne sont pas complètement recyclés, l'ARNm reste associé à la petite sous-unité 40S et la terminaison est suivie par



**Figure 23 : Modèle de dissociation des sous-unités ribosomiques de Pestova et son équipe**

A la fin de la terminaison, ABCE1 est recruté et se lie à eRF1. L'hydrolyse de l'ATP par ABCE1 déclenche son entrée dans un état « ouvert ». Les changements de conformations associés sont transmis à eRF1 via le domaine fer-souffre (FeS) d'ABCE1, ce qui conduit à la dissociation des sous-unités ribosomiques. *Figure adaptée de Jackson et al, 2012*<sup>71</sup>



**Figure 24 : Modèle du recyclage de l'ARNm, de l'ARNt et des sous-unités ribosomiques**

Suite à la dissociation du ribosome, la sous-unité 60S est prise en charge par eIF6. Les facteurs eIF1, eIF1A et eIF3 se lient à la sous-unité 40S pour décrocher l'ARNt déacylé du site P. La fixation d'eIF3j permet le décrochage de l'ARNm. La ligatine constituée du domaine N-terminal de MCT1 et C-terminal de DENR serait également capable de dissocier l'ARNt, et l'ARNm de la sous-unité 40S. Le mécanisme impliqué n'est pas encore déterminé. *Figure adaptée de Jackson et al, 2012*<sup>71</sup>

une réinitiation<sup>33</sup>. En effet, il a été montré chez les mammifères que plus de 45% des ARNm contenaient au moins une upstream Open Reading Frame (uORF)<sup>90</sup> et qu'elles étaient fréquemment traduites<sup>91</sup>. La quasi-totalité des uORFs présentes dans les ARNm mammifères semble être permissive au rescanning et où à la réinitiation, ce qui suggère que ces processus n'impliquent pas de séquences spécifiques. Le mécanisme de réinitiation est encore mal connu, mais des travaux ont mis en évidence l'implication du complexe eIF4F et du complexe stable eIF4G-eIF3-40S<sup>92</sup>.

## F. Les partenaires des facteurs de terminaison

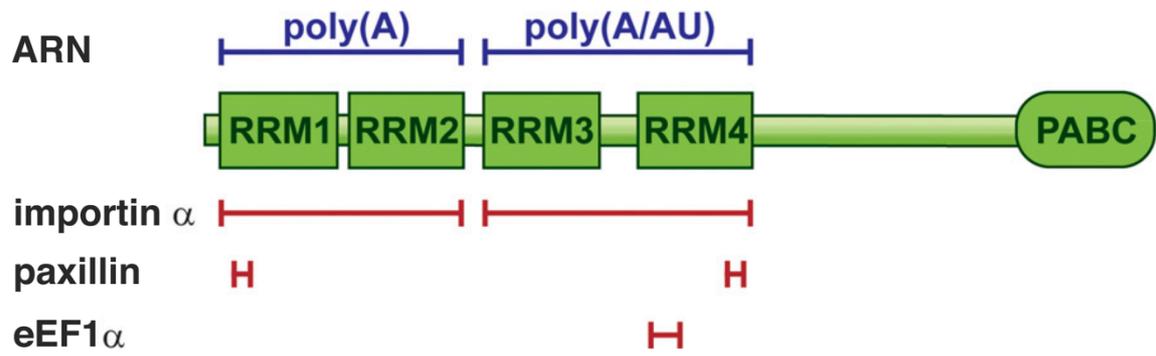
### a. La PABP

Les PABPs (Poly(A) Binding Proteins) sont des protéines multifonctionnelles qui se lient à l'ARN. Elles régulent de multiples aspects de la traduction et de la stabilité des ARNm. Bien qu'elles soient retrouvées principalement de manière diffuse dans le cytoplasme (PABPC1 - Poly A Binding Protein Cytoplasmic 1), les PABPs sont également localisées sous forme de foyers dans le cytoplasme et dans le noyau (PABPN - Poly A Binding Protein Nuclear)<sup>93</sup>.

La protéine PABP1 est constituée de 4 motifs de reconnaissances de l'ARN non identiques, séparés d'un domaine globulaire PABC par une région de liaison (Figure 25). Elle est considérée comme un régulateur central des ARNm cytoplasmiques, contrôlant leur dégradation et leur taux de traduction. En effet, cette protéine se fixe sur la queue poly(A) des ARNm et permet leur circularisation grâce à des interactions avec eIF4G situé sur la coiffe. Elle est également capable d'interagir avec le domaine N du facteur eRF3 via son domaine C-terminal<sup>94</sup>. Cette circularisation et l'interaction eRF3-PABP rapproche spatialement le ribosome en fin de traduction du début de l'ARNm, ce qui pourrait faciliter un nouveau cycle de traduction, car les sous-unités dissociées se retrouveraient à proximité du site d'initiation de la traduction. Cette interaction entre eRF3 et la PABP permettrait aussi de signaler que le codon de terminaison est positionné correctement sur l'ARNm et n'est donc pas un stop précoce susceptible de déclencher les systèmes de dégradation dédiés. La PABP ferait donc la connexion entre la terminaison de la traduction et la stabilité des ARNm.

### b. Les protéines Upf

Chez les mammifères, les facteurs de terminaison eRF1 et eRF3 interagissent avec les protéines Upf (Up-Frameshift) 1, 2 et 3 qui forment le complexe dit de surveillance et qui sont impliquées dans la voie de dégradation NMD (Nonsense



**Figure 25 : Domaines de la protéine PABP1**

PABP1 comporte quatre motifs de reconnaissances de l'ARN (RRM) différents, séparés d'un domaine globulaire (PABC) par une région charnière. Les domaines de liaison à l'ARN sont indiqués par des lignes bleues et les sites d'interactions avec des protéines par des crochets rouges. *Figure adaptée de Gray et al., 2015*<sup>93</sup>

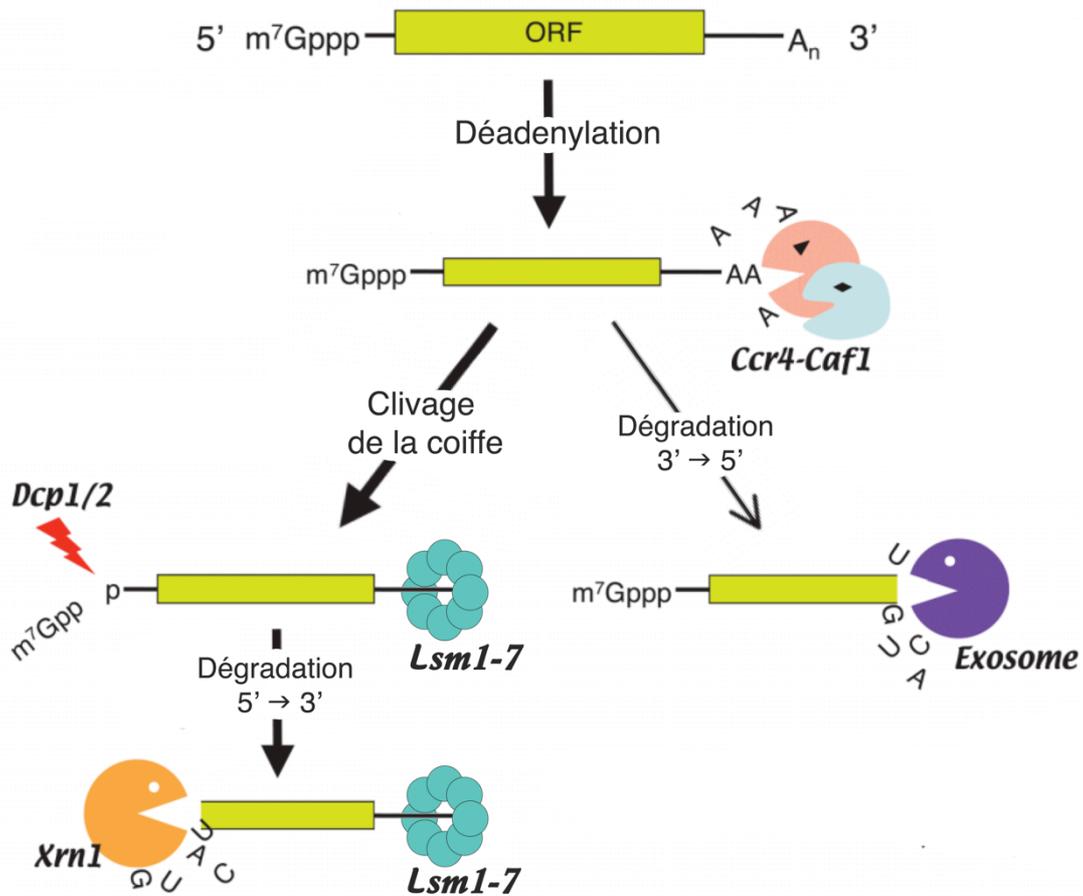
Mediated mRNA Decay). Ce complexe permet la reconnaissance des ARNm contenant un codon stop prématuré et l'enclenchement d'un processus de dégradation qui sera détaillée dans un chapitre suivant. Parmi elles, la protéine Upf1 interagit à la fois avec eRF1 et eRF3, les protéines Upf2 et Upf3 seulement avec eRF3. Indépendamment de son rôle dans la dégradation des ARNm, Upf1 aurait un rôle dans l'efficacité de terminaison de la traduction. Chez la levure, des expériences mesurant le taux de passage de codon stop ont montré qu'en absence d'Upf1, la terminaison était moins efficace, que l'ARNm soit soumis à la NMD ou non <sup>95</sup>. L'inverse a été observé dans les cellules de mammifères où UPF1 aurait un rôle dans l'inhibition de la terminaison de la traduction <sup>96</sup>. Cette différence s'expliquerait par le fait qu'UPF1 humain est homologue à Upf1p de levure, mais également à Mtt1p <sup>97</sup>. Or, Mtt1p est un facteur capable de fixer le domaine GTPase d'eRF3 et d'inhiber la terminaison de la traduction dans la levure. UPF1 humain aurait donc acquis au cours de l'évolution plutôt les caractéristiques de son homologue Mtt1p.

### c. L'ARN hélicase Dbp5

Chez la levure comme chez les mammifères, la protéine Dbp5 est une ARN hélicase de la famille des protéines DEAD-box, essentielles pour l'export des ARNm du noyau. Il a été suggéré qu'elle participerait aussi à la terminaison de la traduction, via une interaction avec le facteur eRF1 <sup>98</sup>. Les auteurs proposent un modèle dans lequel eRF1 serait recruté au site A du ribosome via Dbp5 avant son interaction avec eRF3 <sup>99</sup>. Néanmoins, il est aujourd'hui admis que le complexe eRF1/eRF3 se forme hors du ribosome et que ce complexe est recruté lors de l'entrée d'un codon stop au site A. Ces résultats ne sont pas compatibles avec le modèle d'action de Dbp5 proposé par les auteurs et le rôle potentiel de Dbp5 dans la terminaison reste à définir.

## 4. La dégradation des ARNm : surveillance et contrôle de l'expression génique

L'expression de l'information génétique est un processus essentiel dont la régulation permet aux cellules de s'adapter à de nouvelles conditions, de moduler leur croissance ou de se différencier. Chez les organismes eucaryotes, l'expression des gènes codant les protéines est un processus complexe qui se termine par la dégradation des ARNm. Bien que longtemps considérée comme une simple étape d'élimination, la dégradation des ARNm est également un moyen de contrôler l'expression génique. Ainsi, chez les mammifères, la demi-vie des ARNm varie de



**Figure 26 : Modèle de dégradation des ARNm normaux**

La majorité des ARNm sont dégradés par une voie dépendante de la déadénylation. Après dégradation de la queue poly(A) par Ccr4-Caf1, deux voies sont possibles. La première qui est majoritaire est la dégradation dans le sens 5'→3' par Xrn1. Cette dernière est réalisée après recrutement de Lsm1-7 qui améliore l'efficacité de clivage de la coiffe par Dcp1/2. La seconde voie est la dégradation dans le sens 3'→5' qui est réalisée par l'exosome. *Figure adaptée de Chen and Shyu, 2011*<sup>100</sup>

15 minutes à 24 heures, preuve que l'élimination des ARNm est un processus spécifique et régulé <sup>101</sup>. Elle est également une conséquence des mécanismes de contrôle-qualité mis en place afin de vérifier l'intégrité des transcrits <sup>102</sup>.

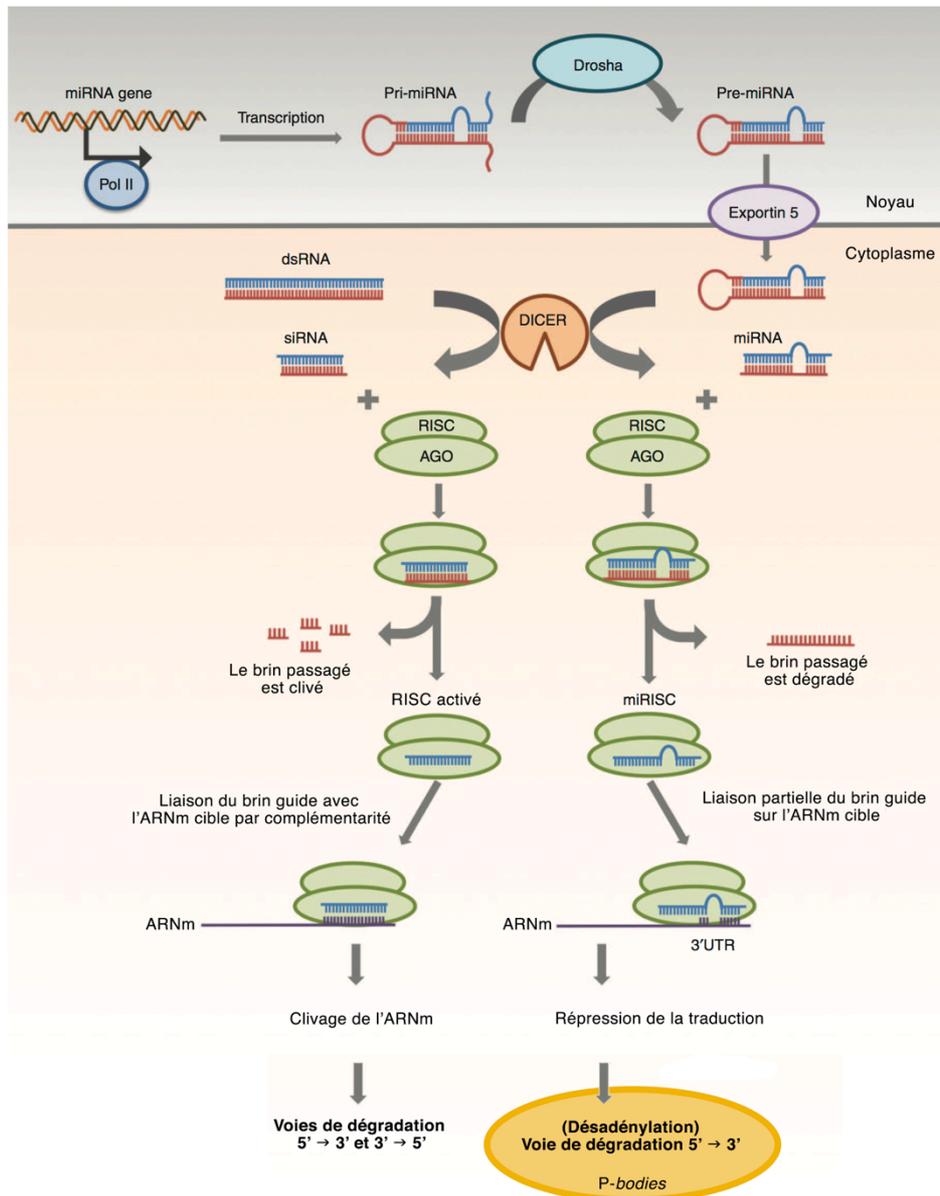
### A. La dégradation cytoplasmique des ARNm

Les ARNm sont des molécules majoritairement sous forme simple-brin, instables, et sont des cibles privilégiées pour de nombreuses ribonucléases. Dans la cellule, plusieurs facteurs contribuent à la protection des ARNm. Lors de sa synthèse, l'ARNm subit des modifications à ses extrémités : l'ajout d'une coiffe (m7G) en 5' et d'une queue poly(A) en 3'. Ces deux éléments, liés par des protéines cytoplasmiques spécifiques, les cap-binding protein (CBP) et Poly(A)-binding protein (PABP) le protègent de l'action des exonucléases. L'ARNm est également associé à d'autres protéines au sein de larges particules ribonucléoprotéiques (mRNP) qui le protègent de coupures endonucléolytiques non spécifiques. L'absence de protection de l'ARNm initie la dégradation de ces derniers.

Chez les eucaryotes, la majorité des ARNm cytoplasmiques sont dégradés par un processus de raccourcissement de la queue poly(A) appelé déadénylation. Toutefois, cette première étape peut être réversible. En effet, les transcrits portants des signaux intacts peuvent être readénylés par des protéines de liaison à l'ARNm tels que celles de la famille des CPEB (Cytoplasmic Polyadenylation Elements Binding proteins) dans les cellules germinales. Ces ARNm peuvent alors être de nouveau traduits <sup>103</sup>. La dégradation de la queue poly(A) est réalisée soit par le complexe Pan2/Pan3, soit par le complexe Ccr4/Not. Néanmoins, s'il reste moins de 10 adénines chez la levure ou moins de 30 à 60 adénines chez les mammifères, le processus de dégradation s'engage rapidement et devient irréversible, la PABP ne pouvant probablement plus se fixer pour protéger l'ARNm de la dégradation. Après déadénylation, le transcrit peut être dégradé soit dans le sens 5' vers 3' par l'exoribonucléase Xrn1, après clivage irréversible de la coiffe par le complexe Dcp1/Dcp2, soit de 3' vers 5' par l'exosome qui est un complexe d'exonucléases <sup>104</sup> (Figure 26).

### B. Le contrôle de l'expression génique

Le processus de renouvellement des ARNm est important, notamment dans la contrôle de l'expression génique à la fois en définissant un niveau basal d'expression et en apportant une réponse rapide à un changement environnemental. En l'absence



**Figure 27 : Modèles des mécanismes de dégradation des ARNm par le système d'ARN interférence**

**siRNA** : Des petits ARNs doubles brins sont modifiés par Dicer en siRNA qui sont ensuite chargés par le complexe RISC. AGO2, un composant de RISC, clive le brin passagé. Le brin guide, mène le complexe RISC actif jusqu'à l'ARNm cible qui est clivé et dégradé.

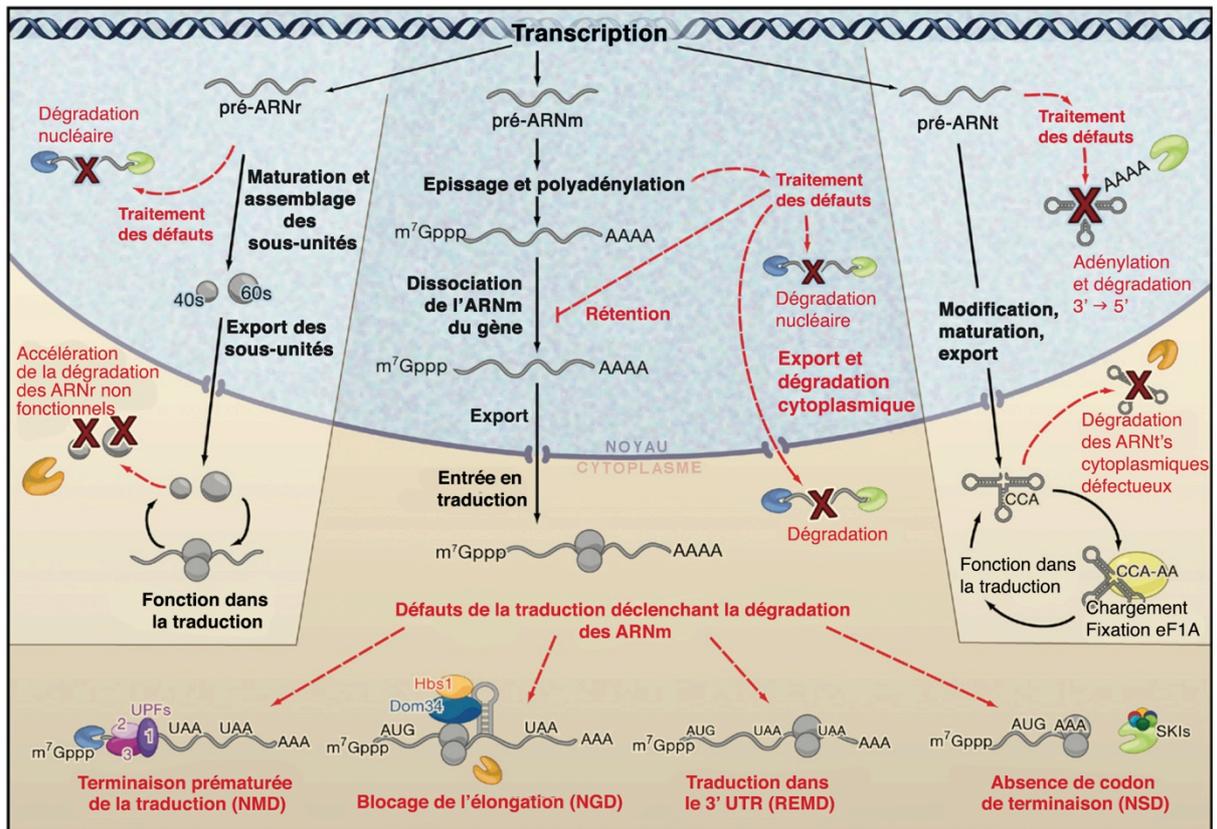
**miRNA** : Les transcrits miRNA sont clivés par Drosha pour former les pré-miRNA. Ces derniers sont transportés par l'Exportine 5 dans le cytoplasme où ils sont modifiés par Dicer en miRNA. Les miRNA sont ensuite chargés par le complexe RISC où le brin passagé est dégradé. Le complexe miRISC est guidé par le brin guide jusqu'à l'ARNm cible. Soit l'endonucléase Argonaute est activée et clive l'ARNm qui sera ensuite dégradé, soit elle n'est pas activée et il y a inhibition de la traduction, suivie éventuellement d'une dégradation. *Figure adaptée de Lam et al., 2015*<sup>105</sup>

d'une dégradation rapide des ARNm, un ajustement immédiat de la production de protéines par la seule régulation transcriptionnelle serait impossible <sup>102</sup>.

Des études globales ont montré que la régulation des quantités d'ARNm se faisait en grande partie par leur dégradation <sup>106</sup>. De nombreux exemples de dégradation régulée des ARNm ont permis de démontrer l'importance de ce phénomène dans le contrôle de la traduction. Cette régulation implique des protéines qui se fixent sur des séquences spécifiques situées dans les régions 5' UTR, 3' UTR et parfois dans la région codante de l'ARNm. Parmi elles, les protéines Puf <sup>107</sup> et les protéines se fixant sur les séquences ARE (AU-rich elements) chez les vertébrés <sup>108</sup> sont des exemples très bien décrits. Dans la majorité des cas, ces protéines induisent la dégradation des ARNm en accélérant le processus de déadénylation. Un autre type de régulation de la traduction par dégradation des ARNm est appelé ARN interférence (RNAi). Ces petits ARN sont des régulateurs importants de la traduction et de la dégradation des ARNm. De nombreux petits ARN ont été découverts ces dernières années, dont les plus connus sont les petits ARN interférents (siRNA) et les micro-ARN (miRNA) <sup>109,110</sup>. Ces RNAi peuvent éteindre l'expression d'un gène soit par clivage via une endonucléase, soit en favorisant la répression de la traduction ou en accélérant le processus polyadénylation. Dans le cas de la dégradation, ces petits ARN régulateurs viennent se fixer sur l'ARNm par complémentarité et induisent sa dégradation par l'intermédiaire du complexe RISC qui comprend l'endonucléase Argonaute <sup>111,112</sup> (Figure 27).

### C. La surveillance de la qualité des ARNm

Chez les eucaryotes, la synthèse d'ARNm fonctionnels nécessite plusieurs étapes de maturation (épissage, coiffage, polyadénylation, exportation, ...) qui peuvent être responsables de l'apparition d'erreurs. La traduction de ces ARNm défectueux conduirait à la synthèse de protéines aberrantes, ce qui pourrait être extrêmement délétère pour la cellule. Heureusement, ces ARNm aberrants sont rarement traduits dans les cellules eucaryotes, car ces dernières ont de nombreux systèmes de surveillance (ou QC pour « contrôle qualité ») des ARNm (Figure 28). Ces systèmes permettent de détecter et de dégrader rapidement ces ARNm ainsi que les protéines naissantes issues de ces ARNm. Le processus le plus largement décrit est la voie NMD (Nonsense-Mediated Decay), qui dégrade les ARNm contenant des codons stop prématurés <sup>22 113</sup>. D'autres voies de contrôle qualité spécialisées dans la dégradation rapide des ARNm responsables de blocage de la traduction ont été décrites plus récemment. La voie de dégradation NGD (No-Go Decay) et NSD (Non-Stop Decay)



**Figure 28 : Diversité des systèmes de contrôle qualité des ARNm eucaryotes**

Schéma représentant quelques uns des systèmes de contrôle qualité des ARN chez les eucaryotes. Les différents types d'ARN défectueux (ARN ribosomiques, ARN de transfert et ARN messagers) ont chacun leur propre système de dégradation. *Figure adaptée de Doma and Parker, 2006* <sup>114</sup>

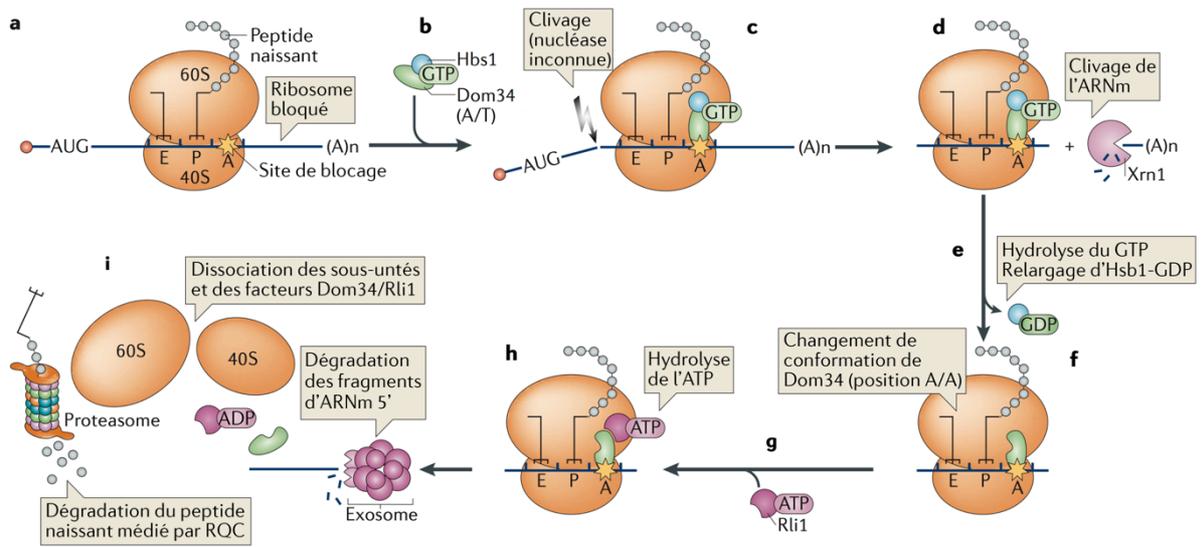
dégradent respectivement les ARNm induisant de forts blocages de la traduction <sup>114</sup> et les ARNm n'ayant pas de codons stop <sup>115</sup>. Ces mécanismes semblent être conservés au cours de l'évolution <sup>116</sup>.

#### a. La voie NGD

Le blocage des ribosomes en cours de traduction pose problème, car il réduit la population de ribosomes disponibles pour la traduction et il conduit à la synthèse d'une chaîne polypeptidique naissante (NC) aberrante et potentiellement toxique. Pour éviter cela, la cellule possède des facteurs protéiques capables de débloquent et dissocier ces ribosomes, de libérer l'ARNm pour permettre sa dégradation et de dissocier la chaîne naissante de l'ARNt. Cette voie de dégradation s'appelle la voie No-Go Decay <sup>117</sup>. Elle constitue probablement la voie de dégradation la moins bien connue en raison de sa découverte plus tardive. Le blocage des ribosomes sur l'ARN peut être dû à la présence d'une structure secondaire, comme un pseudonoeud, et dans une moindre mesure, à la présence de codons rares, d'un codon stop ou d'une dépurination de l'ARNm <sup>118</sup>. Bien que les protéines impliquées dans cette voie semblent être conservées chez les eucaryotes, la quasi-totalité des travaux ont été réalisés chez la levure, je détaillerai les mécanismes de la NGD chez cet organisme.

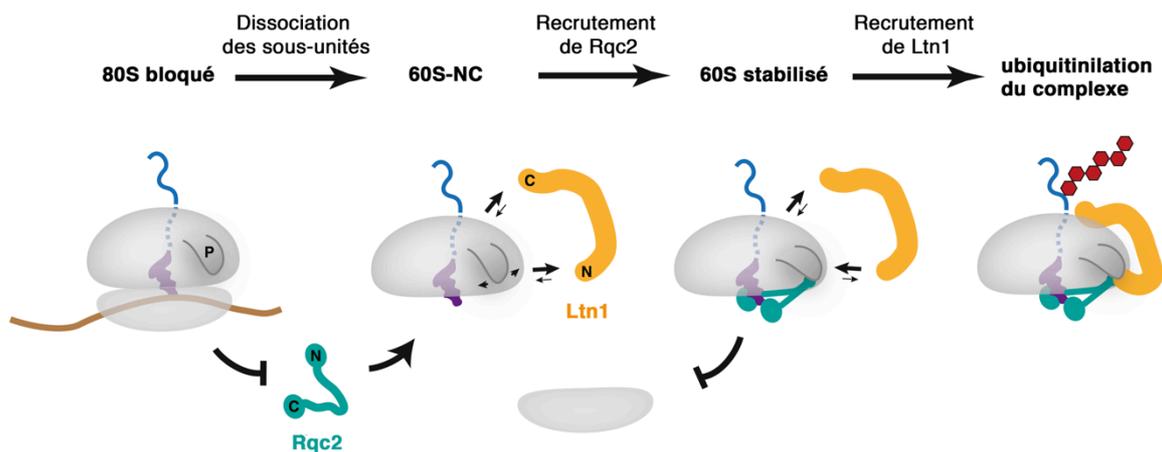
Les facteurs Dom34 (Pelota chez les mammifères) et Hbs1 jouent un rôle important dans le NGD. Ces protéines, dont la structure a été obtenue récemment <sup>83,119</sup>, sont respectivement des homologues d'eRF1 et d'eRF3. La protéine Dom34 est constituée de trois domaines, N-terminal, central et C-terminal. Toutefois, elle ne possède pas les motifs GGQ et NIKS caractéristiques de la fonction d'eRF1. La protéine Hbs1 est constituée de 2 domaines, un domaine N-terminal de taille variable non conservé et un domaine C-terminal conservé, qui est composé des 3 sous domaines G, 2 et 3.

Les récents travaux ont permis de proposer un modèle dans lequel le ribosome bloqué est reconnu par le complexe formé par Dom34 et Hbs1 sur lequel est fixé le GTP <sup>120</sup> (Figure 29). Dom34 est recruté en position A/T sur le ribosome avec Hbs1 fixé au GTP <sup>119</sup>. Comme pour le complexe eRF1-eRF3, la fixation de Dom34 sur Hbs1 augmente l'affinité de Hbs1 pour le GTP. Chez les mammifères, la dissociation par Pelota-Hbs1 requiert également la présence d'ABCE1 <sup>121</sup>, alors que chez la levure, la protéine Rli1 (homologue d'ABCE1) n'est pas nécessaire, mais accélère la dissociation du ribosome par Dom34-Hbs1. L'arrêt de la traduction induit également un clivage endonucléolytique de l'ARNm en amont du ribosome bloqué. Il a été montré que ce clivage n'était pas déclenché par Dom34-Hbs1, puisqu'il a lieu même en l'absence du complexe, mais celui-ci stimulerait la coupure par une endonucléase encore non



**Figure 29 : Modèle partiel du mécanisme de la voie NGD (No-Go Decay)**

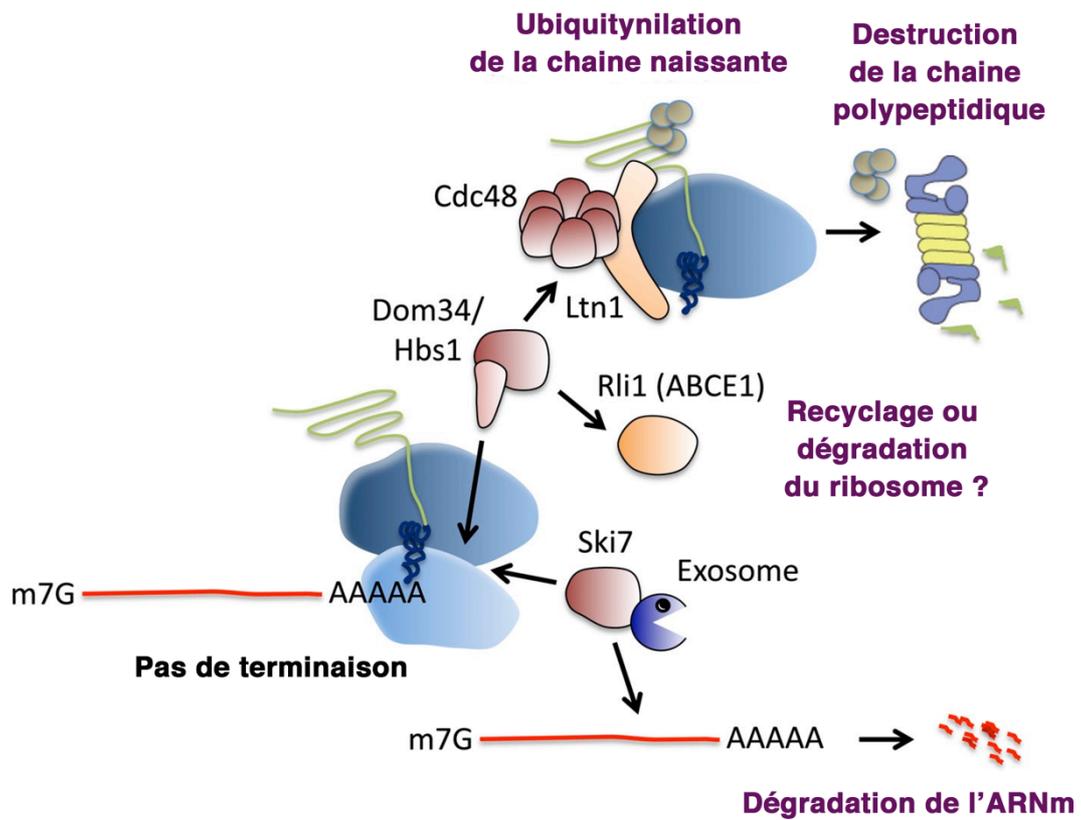
Lorsque le ribosome est bloqué en cours de traduction, la voie NGS se met en place. **a** - Le ribosome est bloqué sur l'ARNm. **b** - Les facteurs Dom34/Hbs1 sont recrutés. Dom34 se positionne dans l'état A/T. Hbs1, fixé à Dom34, est chargé avec une molécule de GTP. **c** - Parallèlement et indépendamment du recrutement de Dom34/Hbs1, il y a clivage endonucléolytique de l'ARNm en amont du ribosome, qui est ensuite dégradé. **d** - Il y a également clivage de l'ARNm en aval du site du blocage et dégradation de ce fragment par Xrn1. **e** - Le GTP porté par Hbs1 est hydrolysé, ce qui provoque la dissociation d'Hbs1. **f** - L'hydrolyse du GTP induit également un changement de conformation de Dom34 qui s'accommode dans le site A. **g** - Rli1 associé à une molécule d'ATP est recruté et se fixe sur Dom34. **h** - L'hydrolyse de l'ATP modifie globalement la conformation de Rli1 qui se dissocie de Dom34. **i** - Dom34 change très probablement de conformation aussi, ce qui provoque la dissociation des sous-unités ribosomique, de l'ARNm, des facteurs Dom34 et Rpli1. La chaîne naissante est prise en charge par le complexe RQC (Ribosome-associated Quality Control) et l'ARNm est dégradé par l'exosome. *Figure adaptée de Graille and Séraphin, 2012*<sup>116</sup>



**Figure 30 : Modèle de prise en charge de la chaîne naissante par le complexe RQC**

Lorsque le ribosome est bloqué, la voie NGD est activée et les sous-unités ribosomiques sont dissociées. La chaîne polypeptidique (NC) liée à l'ARNt au site P se retrouve bloquée avec la 60S. Le complexe RQC, constitué des facteurs Ltn1, Rqc1 et Rqc2 intervient dans la libération de la NC. Tout d'abord, Rqc2 se fixe au niveau de l'interface 40S-60S et stabilise le complexe. Ltn1 est alors recruté et va ubiquitiner la chaîne naissante à dégrader. *Figure adaptée de Shao et al., 2015*<sup>122</sup>

identifiée <sup>123</sup>. Une coupure a également lieu en aval du site de blocage. Ce fragment d'ARNm 3' est dégradé par l'exoribonucléase Xrn1. Suite au clivage de l'ARNm, le GTP est hydrolysé, Dom34 change de conformation et se positionne au site A et Hbs1 se dissocie du ribosome. La protéine Rli1 chargée d'une molécule d'ATP est recrutée à la position de Hbs1 sur Dom34, comme eRF3 sur eRF1. Rli1 hydrolyse l'ATP, ce qui provoque un changement global de conformation de Rli1 puis de Dom34, probablement proche de celui d'ABCE1 lors de l'étape de recyclage de la traduction <sup>116</sup>. Cette étape permet de dissocier les sous-unités ribosomiques et les facteurs Dom34 et Rli1 de l'ARNm qui est dégradé par l'exosome. Le complexe RQC (Ribosome-associated Quality Control), conservé chez les eucaryotes, est composé des protéines Ltn1 (E3 ubiquitine ligase), Rqc1 et Rqc2. Il intervient dans la libération du peptide naissant <sup>124</sup>, mais par un mécanisme qui commence seulement à être identifié (Figure 30). Tout d'abord, Rqc2 reconnaît la chaîne naissante bloquée sur la sous-unité 60S et viendrait se fixer sur les régions d'interactions 60S-40S libérées pour faciliter la fixation de Ltn1 qui va ubiquitinyler la chaîne naissante à dégrader <sup>122</sup>. Ensuite, Cdc48, une AAA ATPase, et ses cofacteurs Ufd1 et Npl4 se lieraient au complexe RQC via Rqc1. Le rôle de Cdc48 et de ses cofacteurs n'est pas encore bien défini, mais ils pourraient faciliter l'export de la chaîne naissante jusqu'au protéasome <sup>124</sup>. Récemment, Shen et ses collaborateurs ont découvert que Rqc2 pouvait non seulement recruter les ARNt Alanine et Thréonine au niveau du site A de la 60S, mais aussi diriger l'élongation en C-terminal des chaînes polypeptidiques bloquées indépendamment de la sous-unité 40S <sup>125</sup>. Ces queues CAT (Carboxy-terminal Ala Thr) n'auraient pas de séquence définie et leur taille est hétérogène. Les auteurs ont observé sur gel SDS-PAGE un signal diffus pouvant dépasser de 5 kDa la taille de la séquence contrôle. La signification biologique de ce processus n'est pas claire et il n'a été observé que dans des cellules déficientes en Ltn1 ou Rqc1. Ce mécanisme ne serait toutefois pas essentiel pour la dégradation de la chaîne naissante (NC) médiée par Ltn1 <sup>126</sup>. Dernièrement, il a été montré que les NC-CAT seraient « marquées » pour former des agrégats. Toutefois, ces agrégats n'ont pas été caractérisés et les données obtenues les rapprocheraient d'amyloïdes qui interagiraient avec la protéine Sis1. Cette protéine pourrait fonctionner comme un facteur de reconnaissance des amyloïdes et activer le désassemblage ou la fragmentation aléatoire de la chaîne naissante via les protéines chaperonnes Hsp104 et Ssa. Bien que le devenir de ces agrégats soit encore inconnu, ils permettraient de séquestrer les NC aberrantes et leur éviter ainsi d'interférer avec les activités cellulaires ou de médier leur élimination via une dégradation vacuolaire <sup>126</sup>.



**Figure 31 : Modèle du mécanisme de la voie NSD**

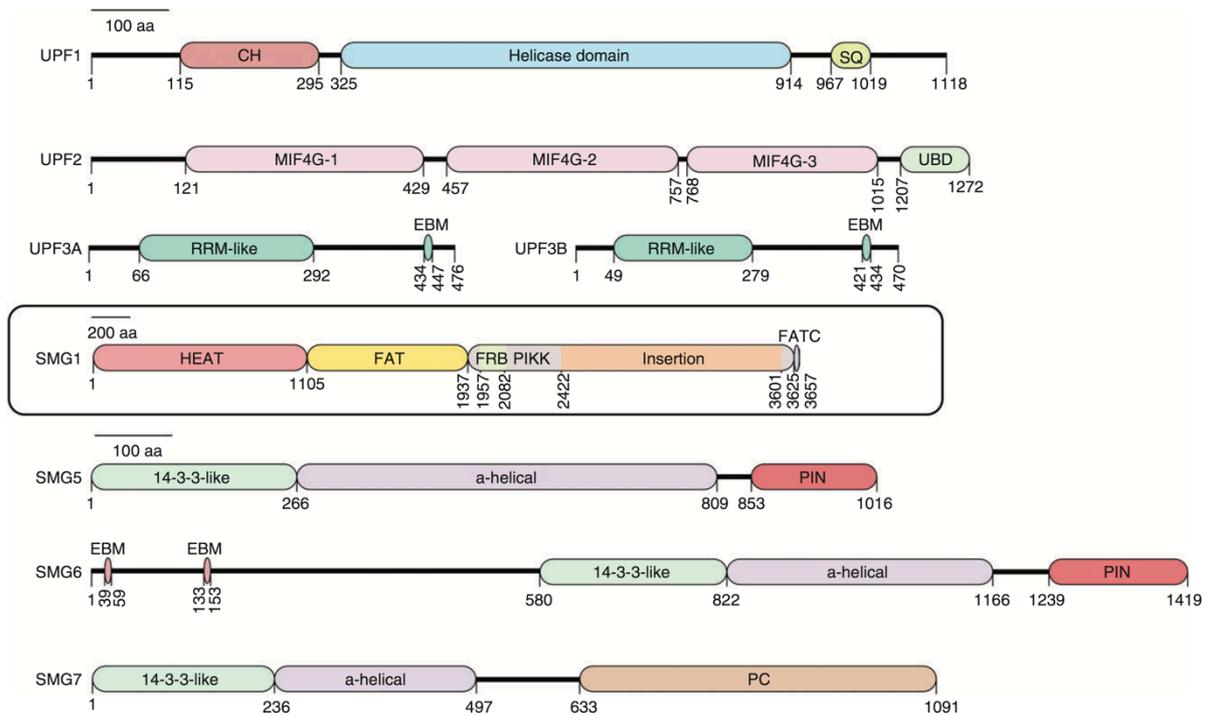
Lorsque le ribosome atteint la fin d'un ARNm sans rencontrer de codon stop, ce dernier se retrouve bloqué sur la queue Poly(A). Ces ribosomes sont détectés par la voie NGD (Dom34/Hbs1) et le facteur Ski7. L'ARNm est dégradé par l'exosome et le ribosome est dissocié par le complexe Dom34-Hbs1 puis Dom34-Rli1. L'extraction et la destruction de la chaîne naissante est probablement médiée par les facteurs de la voie RQC et Cdc48. *Figure adaptée de Lykke-Andersen and Bennett, 2014* <sup>127</sup>

*b. La voie NSD*

La voie Non-Stop Decay (NSD) cible les ARNm qui sont dépourvus de codons stop. De tels transcrits peuvent être générés suite à une rupture de l'ARNm, une polyadénylation prématurée après un arrêt précoce de la transcription, une traduction du codon stop (translecture) ou simplement l'absence d'un codon stop en phase. La voie NSD possède de fortes similitudes avec la voie NGD. Découverte par Parker et Van Hoof chez la levure, elle n'est encore pas complètement comprise à ce jour <sup>128</sup>. Découverte par Frischmeyer et Dietz <sup>129,130</sup> dans les cellules de mammifères cette voie est restée longtemps sujette à controverse et de nombreux travaux ont mis en cause son existence <sup>131-133</sup>. Il faudra attendre les travaux de Saito et ses collaborateurs pour lever le doute et démontrer l'efficacité de cette voie chez les mammifères <sup>134</sup>.

En l'absence de codon stop, le ribosome va traduire jusque dans la queue poly(A) provoquant l'apparition d'une suite de lysines. L'incorporation dans la protéine de 10 acides aminés lysine suffit à réprimer l'élongation en bloquant le ribosome dans la queue poly(A) lorsque l'ARNt lysine est peu abondant dans la cellule. Cela s'expliquerait par la charge positive portée par les lysines qui interagiraient par des liaisons électrostatiques avec la paroi du tunnel de sortie du ribosome qui est chargée négativement <sup>135,136</sup>. Il a également été montré que RACK1 était requis pour l'arrêt de la traduction par une suite d'acides aminés basiques, et qu'il était important pour la détection d'évènements d'élongations aberrantes <sup>137</sup>. Cependant, son mode d'action n'a pas encore été éclairci.

Le blocage des ribosomes dans la queue poly(A) correspond à un cas de NGD, c'est donc le complexe Dom34-Hbs1 qui intervient (Figure 31). L'ARNm est clivé en amont du ribosome par un clivage endonucléolytique. <sup>138</sup>. Le fragment 3' sur lequel le ribosome est bloqué avec un site A vide devient un substrat pour la voie NSD. Chez les mammifères, la dégradation des ARNm messagers dépourvus de stop par l'exosome nécessite l'intervention du complexe SKI, formé des protéines Ski2, Ski3 et Ski8, et de l'adaptateur Hbs1 <sup>134</sup> (et Ski7 chez la levure <sup>129</sup>). Le facteur Hbs1 interagirait simultanément avec l'exosome et le complexe SKI, via son domaine N-terminal, mais ceci reste à vérifier. Chez la levure, l'extrémité C-terminale de Ski7, qui a une structure similaire au domaine GTPase du facteur d'élongation EF1-A et d'eRF3, stimulerait la dégradation de l'ARNm de manière poly(A)-dépendante via l'exosome. Le ribosome serait dissocié grâce à Rli1 après l'hydrolyse de son ATP. L'extraction et la destruction de la chaîne naissante seraient médiées par Ltn1 et probablement par la voie RQC <sup>127</sup>. Bien que le mécanisme de la NSD semble quasiment identique à celui de la NGD et particulièrement chez les mammifères où la protéine Ski7 est remplacée par Hbs1, davantage de travaux sont nécessaires afin d'évaluer en détail ce mécanisme.



**Figure 32 : Illustration schématique des domaines et des motifs des facteurs importants de la NMD chez les mammifères**

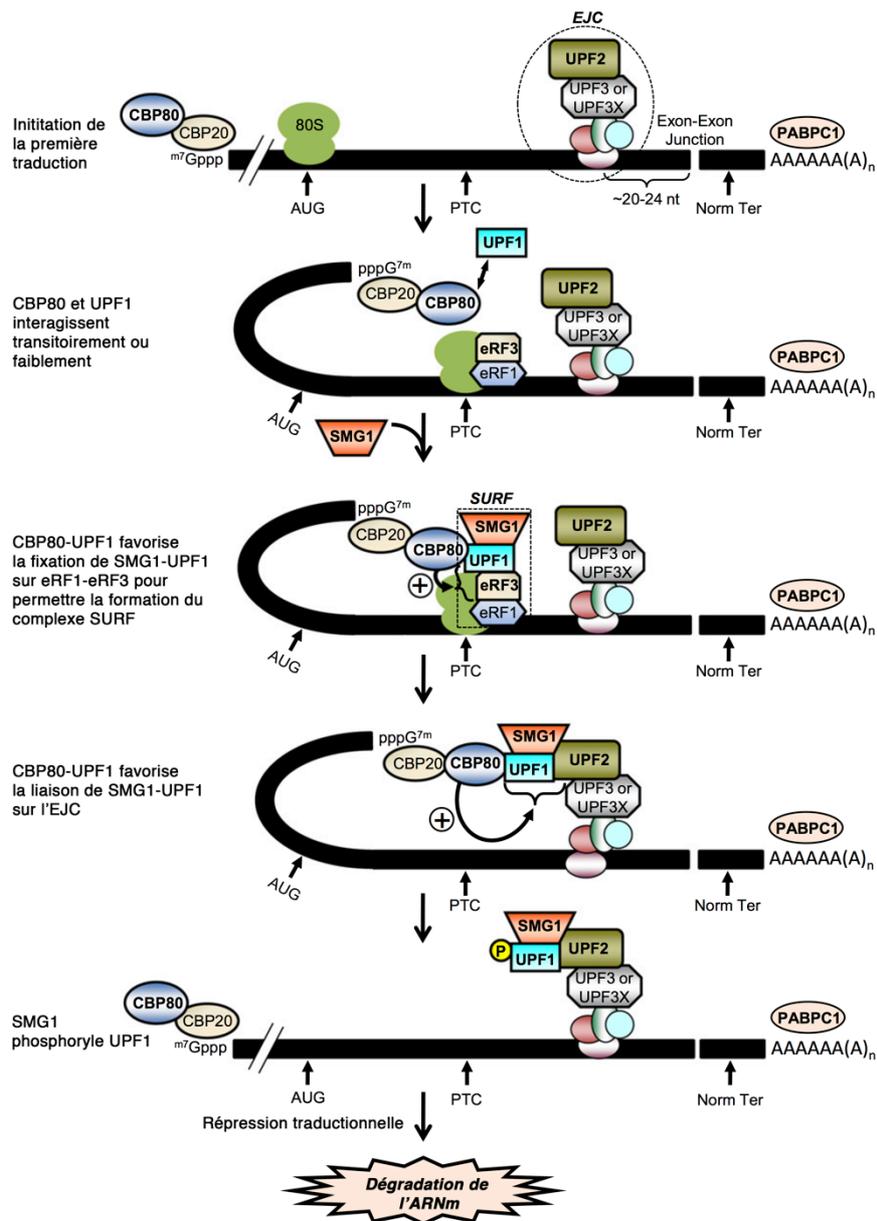
La numérotation des acides aminés se réfère aux protéines humaines. CH : domaine riche en cystéine-histidine; SQ : domaine riche en sérine-glutamine; MIF4G : milieu des domaines 4G; UBD : domaine de fixation à UPF1; MRR : motif de reconnaissance de l'ARN; EBM : motif de liaison des EJC; HEAT : huntingtine, facteur d'élongation 3 (EF3), protéine phosphatase 2A (PP2A), domaine kinase TOR1; FAT : domaine d'adhésion de kinase; PIKK : phosphatidylinositol 3-kinase liée au domaine de la protéine kinase; FATC : domaine C-terminale de FAT; PIN : domaine N-terminale de PERI; PC : région C-terminale riche en proline. *Figure issue de Karousis et al., 2016* <sup>139</sup>

### c. La voie NMD

Il y a plus de 30 ans, il a été découvert que les mutations qui interrompaient prématurément la phase ouverte de lecture réduisaient la demi-vie des ARNm chez la levure et les mammifères <sup>140,141</sup>. En 1993, le terme NMD (non-sense mediated mRNA decay) a été introduit et définissait un processus de dégradation spécifique des ARNm portant un codon stop prématuré (PTC – à ne pas confondre ici avec Peptidyl Transferase Center) <sup>142</sup>. La voie NMD était alors vue comme un mécanisme de contrôle qualité qui prévenait la synthèse de protéines tronquées potentiellement délétères. Ces dernières années, la définition de la voie NMD a beaucoup évolué depuis que les études globales du génome ont révélé que la stabilité de nombreux ARNm ne contenant pas de PTC était également contrôlée par la voie NMD. La NMD modulerait directement ou indirectement la stabilité d'environ 10% des ARNm des cellules de mammifères, participant ainsi de façon significative à la régulation de l'expression des gènes. Toutefois, les voies et fonctions biologiques régulées par la NMD commencent seulement à être identifiées <sup>139</sup>. Il a notamment été montré que les facteurs intervenant dans la voie NMD seraient nécessaires au développement embryonnaire chez les mammifères <sup>143</sup>, ils contribueraient également au développement du cerveau <sup>144</sup> et permettraient de moduler la réponse au stress <sup>145</sup>. Enfin la voie NMD régulerait l'expression de ses propres facteurs <sup>146,147</sup>.

#### ★ Les facteurs de la voie NMD

Malgré un nombre de travaux très importants consacrés à la voie NMD, les mécanismes moléculaires impliqués dans cette voie restent mal compris. Les facteurs centraux de la voie NMD sont au nombre de neuf (Figure 32). Les trois premiers ont été identifiés initialement chez *S. cerevisiae* et sont appelés UPF1, 2 et 3 (Up-Frameshift) <sup>148</sup>. Plus tard, les facteurs SMG1, 5, 6, 7, 8 et 9 ont été découverts (suppressors with morphogenetic effects on genitalia) <sup>149</sup>. UPF1 est une hélicase qui possède également une activité ATPase essentielle chez les eucaryotes ; UPF2 permet de lier les facteurs UPF1 et UPF3 et d'activer SMG1; UPF3 est en réalité constitué de 2 protéines homologues chez l'humain : UPF3A et UPF3B. UPF3A interviendrait seulement lorsque UPF3B ne serait pas présent en quantité suffisante. Cette protéine s'associe aux EJC (Exon-Junction Complex) un complexe protéique déposé 20 à 24 nucléotides en amont des jonctions exon-exon sur l'ARNm. UPF3 permettrait également de stimuler l'activité ATPase et hélicase d'UPF1. SMG1 est une protéine kinase responsable de la phosphorylation d'UPF1. Alors que SMG5 et 7 sont impliquées dans sa déphosphorylation via le recrutement de PP2A, une protéine phosphatase.



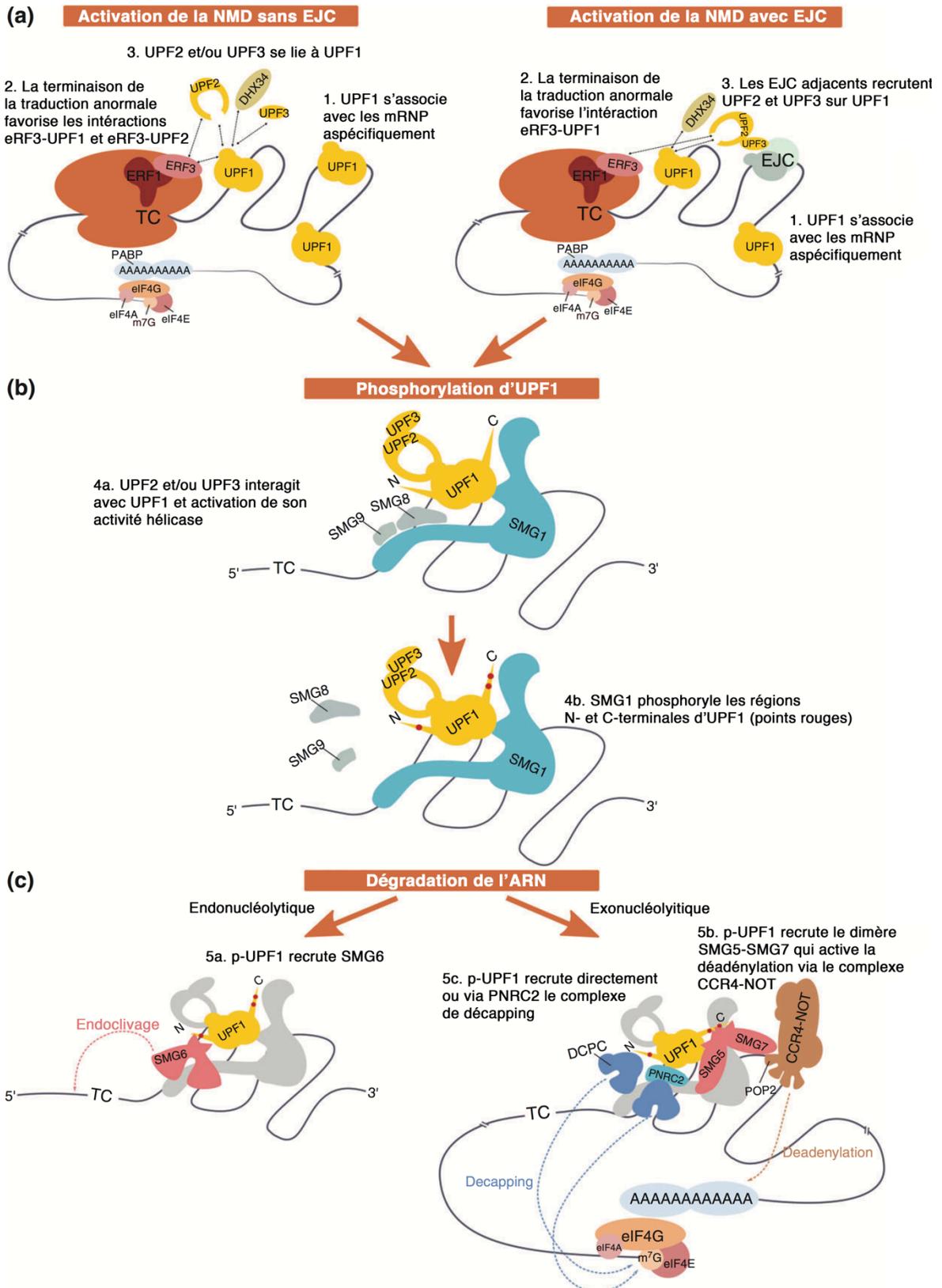
**Figure 33 : Modèle d'une voie NMD exclusive de la première traduction**

Lors de la toute première traduction, les ribosomes s'engagent sur un ARNm portant l'hétérodimère CBP80-CBP20 (CBC) sur sa coiffe et au moins un complexe exon-exon (EJC) si l'ARNm est issu de l'étape d'épissage. Lorsque les ribosomes atteignent le codon stop prématuré<sup>22</sup>, la protéine CBP80 interagit directement mais transitoirement avec le facteur UPF1. L'interaction CBP80-UPF1 favorise la liaison de UPF1-SMG1 sur eRF1-eRF3 pour former le complexe SURF. Cette liaison serait en compétition avec PABPC1 associée à la queue poly(A). CBP80 favorise ensuite la liaison de SMG1-UPF1 avec UPF2-UPF3 associé à l'EJC. UPF1 est alors phosphorylé par SMG1 et active les voies de dégradation de l'ARNm. *Figure adaptée de Hwang et al., 2010*<sup>150</sup>

SMG7 permet aussi de recruter le complexe de déadénylation CCR4-NOT qui interagit avec POP2 et qui intervient dans la dégradation de l'ARN. SMG6 est une endonucléase capable d'initier le clivage de l'ARNm à proximité du codon stop chez l'humain. Elle interagirait également avec UPF1 et les complexe EJC. SMG8 et 9 sont deux facteurs régulateurs de SMG1 et inactivent son activité hélicase.

### ★ **Modèle d'une voie NMD exclusive du premier cycle de la traduction**

La voie NMD est activée lorsqu'un ribosome rencontre un codon stop prématuré. Ce codon stop est efficacement reconnu comme tel si un complexe EJC est localisé plus de 50 nucléotides en aval du codon stop. Cette terminaison conduit à l'activation de la voie NMD et à la dégradation de l'ARNm. Dans ce modèle (Figure 33), la présence d'EJC représentait un signal essentiel pour identifier la terminaison prématurée et pour l'activation de la voie NMD <sup>151</sup>. Plusieurs travaux ont montré que les ARNm aberrants étaient sujets à la NMD uniquement lors du premier passage du ribosome <sup>152</sup>. En effet, ce dernier en traduisant éliminerait les complexes EJC et donc le signal permettant de reconnaître les transcrits comme aberrants <sup>153</sup>. De plus, ces ARNm nouvellement synthétisés et maturés venant d'être exportés dans le cytoplasme portent un complexe CBC (cap-binding complex) composé des protéines CBP80 et CBP20. Ce complexe est modifié indépendamment de la traduction et remplacé par eIF4F lors de son entrée dans le cytoplasme <sup>154</sup>. Or il a été observé que CBC interagit avec UPF1 via CPB80 et que cette interaction n'a pas lieu avec eIF4E <sup>150,155,156</sup>. Ces résultats iraient dans le sens que seule la première traduction est sujette à la NMD, car une fois CBC remplacée, ces transcrits seraient immunisés contre cette voie de dégradation. Toutefois, ces résultats ont été contredits par d'autres travaux dans lesquels il a été observé une interaction entre eIF4E-UPF1 chez les mammifères <sup>154,157</sup>. De plus, de nombreux ARNm sont régulés par la NMD, même après le premier cycle de traduction et donc en absence d'EJC <sup>158</sup>. Ces complexes EJC ne seraient donc pas indispensables à l'activation de la voie NMD <sup>159</sup>. Ces derniers résultats seraient cohérents avec un modèle unique de la voie NMD pour tous les organismes sur lequel je reviendrai dans la partie suivante <sup>160</sup>.



**Figure 34 : Modèle universel de la voie NMD**

(a) La terminaison anormale de la traduction se produit lorsque le codon de terminaison (TC) est éloigné de la queue poly(A). L'activation de la NMD nécessite l'interaction d'UPF1 et non PABP avec les facteurs de terminaison. (suite de la légende page suivante)

L'activation d'UPF1 repose sur le recrutement d'UPF2 et/ou UPF3. Dans ce modèle, il est proposé qu'UPF1 se lie de manière non spécifique à l'ARNm avant la traduction et interagit avec les facteurs de terminaison sur le ribosome positionné au niveau du TC. En absence d'EJC, UPF1 est activé par UPF2 et/ou UPF3. La présence d'EJC facilite le recrutement d'UPF2-UPF3 et va donc faciliter l'activation de la voie NMD. **(b)** La kinase SMG1 est activée lorsque ses homologues SMG8 et 9 sont dissociés, ce qui est facilité par UPF2 et DHX34. SMG1 phosphoryle UPF1 sur ses résidus Ser et Thr en N- et C-terminale. L'activité hélicase d'UPF1 est activé et stimule les voies de dégradations de l'ARN. **(c)** La phosphorylation d'UPF1 (p-UPF1) favorise la dégradation de l'ARNm par au moins trois mécanismes distincts : le clivage endonucléolytique par SMG6 (a), le recrutement du complexe CCR4-NOT médiée par SMG7 (b) et le recrutement du complexe de déadénylation (DCPC) directement par UPF1 ou par l'intermédiaire de PNRC2 (c). Les ARNm qui sont la cible de la voie NMD sont ensuite dégradés par les activités exonucléotiques cellulaires. *Figure issue de Karousis et al., 2016* <sup>139</sup>

### ★ Un modèle universel

Outre le fait que le modèle unifié soit commun à tous les organismes, ce modèle repose sur l'hypothèse que la voie NMD serait la conséquence d'une terminaison anormale de la traduction (Figure 34). Les ribosomes resteraient davantage sur les codons stop sujets à la NMD, ce qui impliquerait des différences mécanistiques entre le processus de terminaison classique et celui permettant de déclencher la voie NMD <sup>161</sup>.

Lors du processus canonique de la terminaison, PABPC1 chez les mammifères interagirait avec eRF3 via ses deux motifs PAM2 et cette interaction empêcherait l'activation de la voie NMD <sup>162,163</sup>. Ce résultat suggère que PABPC1 favoriserait la terminaison de la traduction classique par un mécanisme qui doit encore être identifié. Parallèlement, il a été rapporté qu'UPF1 interagirait avec eRF3 et que cette interaction se ferait en compétition avec PABPC1 <sup>96,164</sup>. Deux modèles différents alors quant à savoir quand et comment UPF1 interagirait avec eRF3 <sup>139</sup>. Soit, UPF1 serait directement recruté sur eRF3, soit UPF1 serait présent sur l'ARNm et interagirait avec eRF3 lorsque le ribosome rencontre un codon stop prématuré. Le plus probable serait qu'UPF1 se fixe directement à l'ARNm indépendamment de la traduction et de façon non spécifique. Plusieurs travaux ont montré qu'UPF1 est capable de se lier dans le 3'UTR d'ARNm, sur des ARN non codants et dans la phase codante des ARNm régulés ou non par la voie NMD <sup>165,166</sup>. UPF1 se lierait donc aux ARNm avant la traduction puis serait détaché de l'ARNm par les ribosomes. Ainsi, lors d'une terminaison anormale de la traduction sur un codon stop, le complexe eRF1-eRF3 serait recruté et il y aurait compétition entre UPF1 et PABPC1 pour interagir avec eRF3.

En absence d'EJC, le recrutement d'UPF2 serait beaucoup plus lent et PABPC1 pourrait encore interagir avec eRF3 en dissociant UPF1, ce qui aurait pour conséquence d'inactiver la NMD et de favoriser la terminaison classique de la traduction <sup>160</sup>. En présence des EJC, le complexe UPF2-UPF3 s'associe aux EJC, ce qui favorise grandement l'interaction UPF1-UPF2 et donc l'activation de la voie NMD <sup>159</sup>. Dans certaines circonstances, la voie NMD peut être activée en absence d'UPF2 et d'UPF3 qui serait alors remplacée par DHX34 <sup>167</sup>.

Le changement de conformation provoqué par la fixation d'UPF2 à UPF1 permet également de recruter SMG1 lié à SMG8 et 9. Avec l'aide d'UPF2, SMG8 et SMG9 se dissocient de SMG1, dont l'activité kinase est alors activée. SMG1 va phosphoryler les résidus sérine et thréonine en N- et C-terminal d'UPF1 <sup>168</sup> ce qui déclenche la dissociation des sous-unités ribosomique, la libération de SMG1, l'activation de l'activité hélicase d'UPF1 et les mécanismes de dégradation de l'ARNm <sup>169,170</sup>.

### ★ Les mécanismes de dégradation de l'ARNm de la voie NMD

Indépendamment des mécanismes exacts de l'identification des cibles de la NMD, UPF1 hyperphosphorylée servirait de plateforme de liaison pour les effecteurs SMG5, SMG6 et SMG7. Il est prouvé que, chez les mammifères, la dégradation rapide des cibles de la NMD peut être réalisée par au moins deux voies différentes (Figure 34), mais les mécanismes d'activation de l'une plutôt que l'autre ne sont pas clairs <sup>171</sup>.

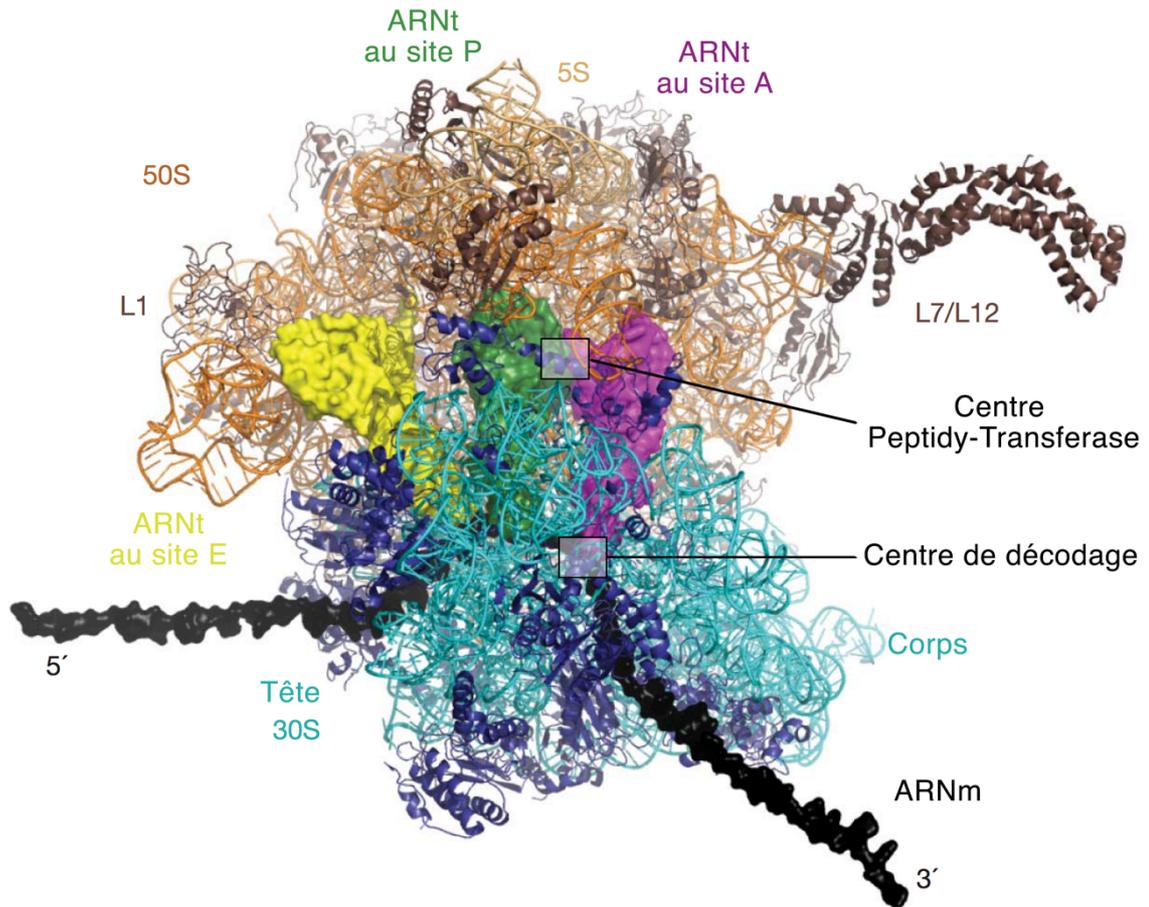
La voie majoritaire est déclenchée par l'interaction de SMG6 avec UPF1 <sup>172</sup>. SMG6 clive l'ARNm au niveau du codon stop prématuré dans les cellules humaines. Les fragments 3' d'ARN sont alors rapidement dégradés par Xrn1 et les fragments 5' sont digérés par l'exosome <sup>173,174</sup>.

La voie alternative de dégradation des ARN serait une voie de secours et serait déclenchée par le recrutement du complexe SMG5-SMG7 qui interagit avec les sérines phosphorylées d'UPF1 <sup>175</sup>. Ce complexe déphosphorylerait SMG1 puis s'associerait au complexe CCR4-NOT pour déadényler la queue poly(A) et permettre à XRN1 de dégrader l'ARNm.

UPF1 pourrait aussi s'associer au complexe de décoiffage DCP1A, DCP2 et PNRC2 sans être phosphorylé. Cette voie de déadénylation serait indépendante de SMG5, 6 et 7, et constituerait une troisième voie potentielle de dégradation des ARNm aberrants <sup>176,177</sup>.

### III. La fidélité de la traduction

Pour assurer une transmission et une expression exacte de l'information génétique, chaque étape du processus doit être fidèle. Chez les eucaryotes, le taux d'erreur de l'ADN polymérase lors de la réplication est de l'ordre de  $10^{-10}$ <sup>178</sup>. Lors de la transcription, la fréquence d'incorporation erronée (misincorporation) est d'environ  $10^{-5}$ <sup>179</sup>. La traduction est l'étape de l'expression génique qui est la moins fidèle avec un taux d'erreur qui varie entre  $10^{-3}$  et  $10^{-4}$ <sup>180</sup>. Cela représente pour une protéine moyenne de 400 codons, au moins un acide-aminé incorporé à tort (misincorporé)<sup>181</sup>. Le plus souvent, cette mauvaise incorporation n'est pas délétère, mais 10 à 50% des substitutions intégrées aléatoirement dans plusieurs gènes affectent la fonction de la protéine<sup>182</sup>. Au cours de la traduction, un acide aminé erroné peut être incorporé suite à une ou des erreurs d'aminocyclation des ARNt. Toutefois, grâce aux différents processus d'« editing » et de « proofreading » des aminoacyl ARNt synthétases, le taux d'erreur du chargement de l'acide aminé sur son ARNt est très faible, inférieur à  $10^{-6}$ . L'aminocyclation constitue une étape cruciale de la traduction, car une fois que le chargement de l'acide aminé a eu lieu, il est irréversible. De plus, le ribosome ne contrôle pas l'exactitude de l'acide aminé incorporé. Il peut aussi se produire des erreurs au cours de l'élongation ou de la terminaison. En effet, le ribosome sélectionne l'ARNt dont l'appariement codon-anticodon est correct, mais cette étape de discrimination peut parfois valider un ARNt dont l'appariement n'est que partiel. Le ribosome peut également être amené à changer de phase en cours de traduction (frameshift) ou à sauter une région de l'ARNm (bypassing). Lors de l'étape de terminaison, le codon stop peut être non reconnu permettant à un ARNt d'être incorporé à la place des facteurs de terminaison. Les erreurs consécutives à un frameshift ou au bypassing sont les erreurs les plus délétères pour la protéine. Elles sont donc les erreurs les plus contre sélectionnées. Ainsi la grande majorité des erreurs correspondent à l'incorporation d'un mauvais ARNt (misincorporation)<sup>183</sup>. Le rôle du ribosome étant d'assurer un compromis entre vitesse et fidélité de la traduction, celle-ci n'est donc pas maximale, mais optimale.



**Figure 35 : Centre de décodage et centre peptidyl-transférase du ribosome 70S**

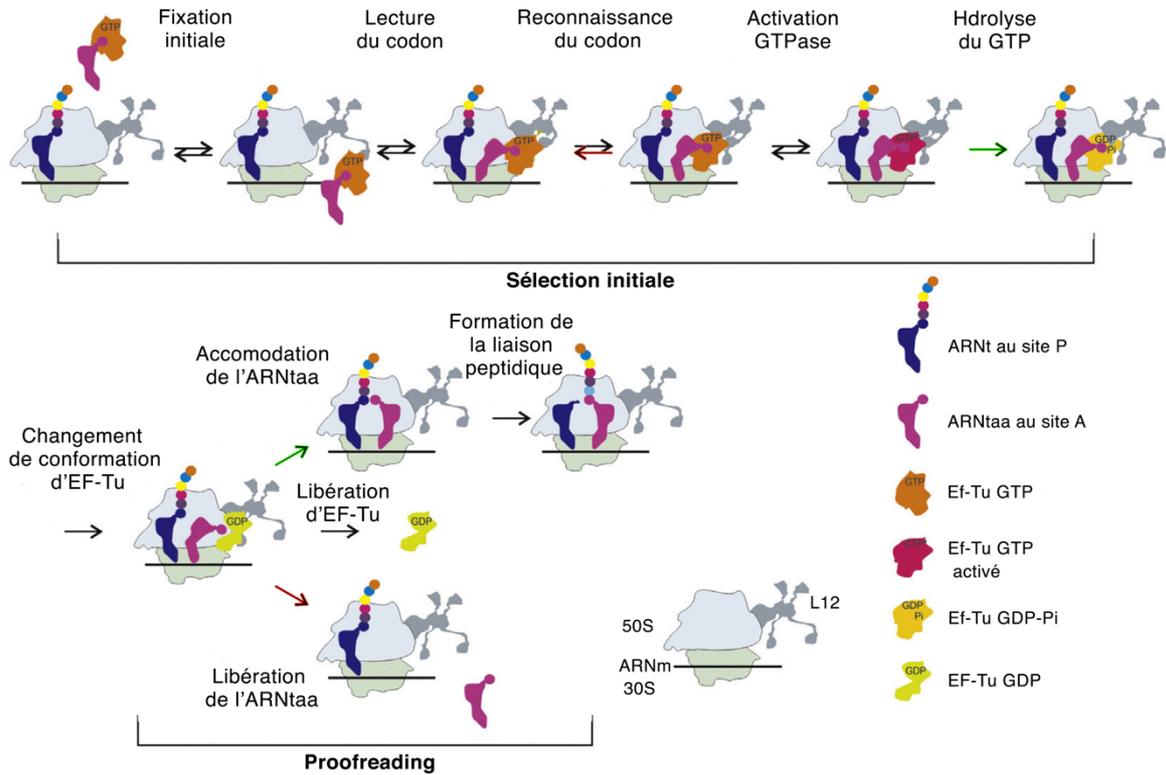
La grande sous-unité 50S est représentée en orange, la petite sous-unité 30S en turquoise et L'ARNr 5S en orange clair. Les ARNt au site A, P et E sont représentés respectivement en magenta, vert et jaune. Le codon de l'ARNm au site A est coloré en bleu. L'ARNm est représenté en noir. Les protéines L1 et L7/L12 sont de couleur brune. La localisation du centre de décodage et du centre peptidyl-transférase est indiquée. *Figure adaptée de Schmeing and Ramakrishnan, 2004*<sup>184</sup>

## 1. Le rôle du ribosome dans le maintien de la fidélité de la traduction

La fidélité de la traduction est basée sur l'activité des ARN ribosomiques qui forment le cœur du ribosome, en particulier le centre de décodage et le centre peptidyl-transférase, et également de certaines protéines ribosomiques (Figure 35). Le centre de décodage est formé de l'ARN 18S (16S chez les procaryotes) dans la petite sous-unité du ribosome et est impliqué dans la reconnaissance et la validation de l'appariement codon-anticodon entre l'ARNm et l'ARNt. Le centre peptidyl-transférase est formé de l'ARN 23S (23S chez les procaryotes) dans la grande sous-unité du ribosome et catalyse la liaison ester entre les acides aminés lors de l'élongation, ou hydrolyse la liaison entre le dernier ARNt et l'acide aminé lors de la terminaison. De plus, le décodage au niveau du ribosome doit se faire rapidement pour maintenir la vitesse élevée de synthèse des protéines, le ribosome incorporant environ 10 acides aminés par seconde. De nombreuses études de mutagenèse des gènes codant les protéines ou les ARN du ribosome <sup>185</sup>, ainsi que des analyses structurales, ont été menées afin de mieux comprendre les mécanismes moléculaires mis en jeu dans la fidélité de la traduction.

## 2. Les mécanismes de décodage

La compréhension des mécanismes impliqués dans le décodage a particulièrement progressé depuis l'obtention de plusieurs structures bactériennes haute résolution, d'abord des sous-unités ribosomiques, puis de l'ensemble du ribosome <sup>184</sup>. L'analyse de ces structures a été grandement améliorée par des études biochimiques et génétiques et par des structures de haute résolution du ribosome déterminées en cryo-EM. Plus récemment, les structures des sous-unités 40S et 60S eucaryotes <sup>186,187</sup> puis du ribosome 80S <sup>188</sup> ont mis en évidence des différences structurales entre les ribosomes bactériens et eucaryotes. Toutefois, le cœur du ribosome est fortement conservé et les étapes du cycle d'élongation sont pratiquement inchangées dans le domaine du vivant <sup>32</sup>. Étant donné que la majorité des données obtenues sur les mécanismes mis en jeu lors du décodage sont issues d'études structurales faites sur le ribosome procaryote, tous les éléments structuraux cités dans ce paragraphe feront donc référence à la nomenclature procaryote.



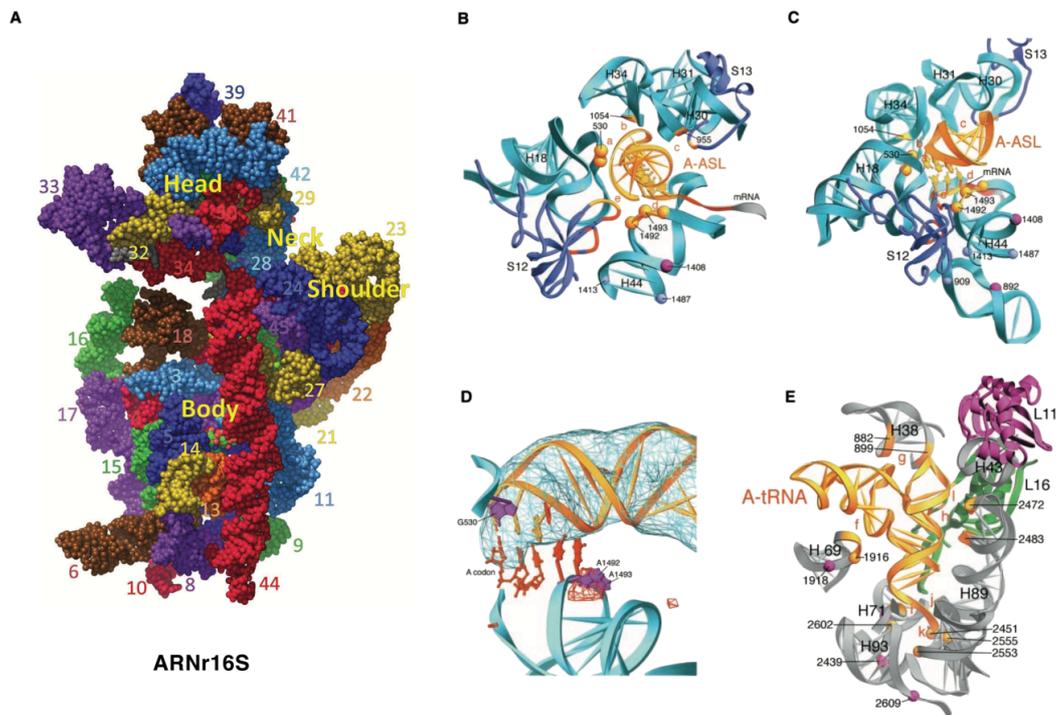
**Figure 36 : Processus de sélection des ARNt**

La sélection des ARNt lors de l'élongation se produit en deux étapes qui sont la sélection initiale et l'étape de relecture ou « proofreading », séparées par la réaction d'hydrolyse du GTP. Les ARNt proche-cognats peuvent être sélectionnés lors de la reconnaissance initiale mais sont rejetés lors l'étape de proofreading. Les flèches vertes indiquent les réactions qui sont accélérées pour un ARNt cognat, les flèches rouges les réactions qui sont plus rapides pour les ARNt proche-cognats. *Figure adaptée de Rodnina and Wintermeyer, 2016* <sup>189</sup>

## A. La sélection de l'ARNt

A la fin de l'initiation, le ribosome est positionné sur l'ARNm avec l'ARNt initiateur au site P et un site A libre qui contient le codon à décoder. Le ribosome sélectionne un ARNt dont l'appariement codon-anticodon est correct sur les trois bases. Toutefois, le pool d'ARNt se répartit en 3 classes : les ARNt cognat qui s'apparient complètement au codon, les ARNt proche-cognat qui s'apparient au codon avec un mésappariement d'une base et les ARNt non-cognat qui ne s'apparient pas au codon. Le ribosome doit pouvoir discriminer le bon ARNt dont l'anticodon est complémentaire du codon au site A pour maintenir un niveau de fidélité optimale compatible avec le fonctionnement de la cellule.

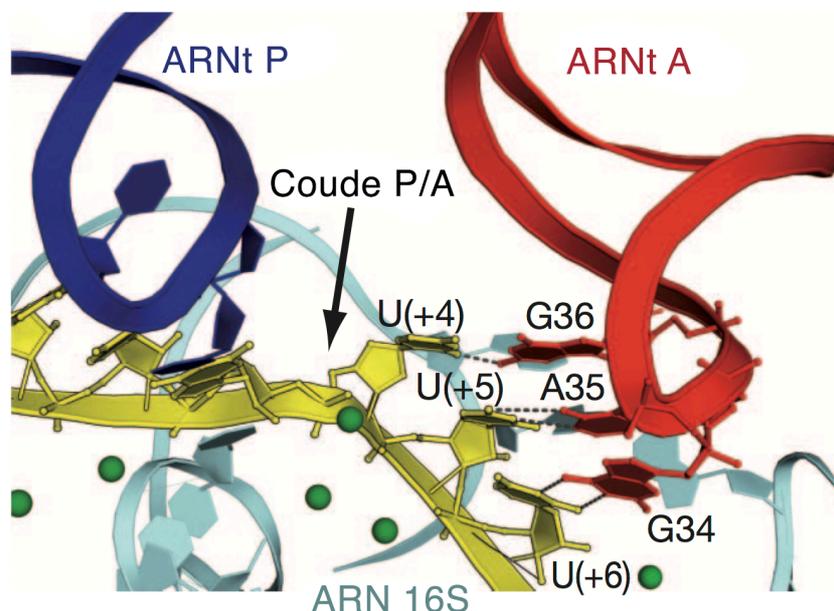
Cette sélection se fait en 2 étapes, la sélection initiale et la relecture (« proofreading »), séparées par une réaction d'hydrolyse du GTP<sup>189</sup> (Figure 36). Les ARNt non-cognat sont discriminés dès la première étape, en revanche les ARNt proche-cognat peuvent échapper à cette première sélection et sont alors rejetés lors de la deuxième phase. Pendant les deux étapes de la sélection, les ARNt sont discriminés sur la base de leur stabilité différente avec le ribosome et de la vitesse d'hydrolyse du GTP et d'accommodation de l'ARNt au site A. Lors de la première étape, le complexe ternaire, formé de l'ARNt aminoacylé et du facteur d'élongation associé au GTP (EF-Tu), est recruté via EF-Tu par le complexe protéique L7/L12 situé à la surface de la sous-unité 50S, près du site A. Une fois le complexe ternaire fixé au ribosome, l'ARNt est dans une configuration A/T, c'est-à-dire que seule la boucle anticodon de l'ARNt est située dans le site A. Il y a ensuite reconnaissance du codon, ce qui induit un changement de conformation du facteur d'élongation dont le domaine GTPase est alors activé. Le GTP est hydrolysé, ce qui libère un phosphate inorganique et provoque un réarrangement conformationnel d'EF-Tu. Débute alors la deuxième phase de « proofreading ». Le facteur d'élongation a perdu son affinité pour l'ARNt aminoacylé et se dissocie du ribosome. L'ARNt est alors libre de bouger dans le site A et son extrémité 3' est accommodée dans le centre peptidyl-transférase. L'ARNt aminoacylé cognat s'accommode au site A à la fois très rapidement et de manière efficace. En revanche, la plupart des ARNt proche-cognat sont rejetés à cause de leur faible stabilité de fixation et de leur accommodation plus lente. La liaison peptidique entre le nouvel acide aminé et la chaîne peptidique est catalysée dans le centre peptidyl-transférase situé entre les sites A et P.



**Figure 37 : Régions impliqués dans le décodage**

**A** - Représentation 3D de la structure de l'ARNr 16S d'*E. Coli* et de ses hélices (PDB : DAW7). *Figure adaptée de Sweeney et al., 2015* <sup>190</sup>

**B-D** - Structures cristallographiques de *Thermus thermophilus*. **B,C** - Deux vues de l'anticodon d'un ARNt-Ala lié à son codon au site A de la sous unité 30S du ribosome. **D** - Représentation de Fourier de l'ARNt-A superposée à la structure de l'ARNt-A et du codon au site A. **E** - Interaction de la tige-boucle D de l'ARNt-A avec la sous-unité 50S. *Figure adaptée de Yusupov et al., 2001* <sup>191</sup>



**Figure 38 : Interactions codon-anticodon dans le centre de décodage du ribosome 70S**

Le coude formé par l'ARNm entre le codon au site A et le codon au site P, stabilisé par l'ARNt au site P, les ions magnésium  $Mg^{2+}$  et l'ARNr 16S, contraignent le positionnement de la première base du codon. La boucle anticodon de l'ARNt au site A est représentée en rouge, celle de l'ARNt au site P en bleu, l'ARNm est coloré en jaune, et les ions magnésium correspondent aux billes vertes. *Figure adaptée de Demeshkina et al., 2012* <sup>46</sup>

## B. Les régions impliquées dans le décodage

Le centre de décodage est situé dans la petite sous-unité du ribosome, au niveau de l'ARNr 16S. En interagissant avec l'ARNm et l'ARNt, la petite sous-unité du ribosome joue un rôle prépondérant dans la reconnaissance codon/anticodon au site A du ribosome. Plusieurs régions de cet ARN ont été identifiées comme affectant la fidélité de la traduction (Figure 37) :

- L'hélice 44 porte les résidus A1492 A1493 et l'hélice 18, le résidu G530. Ces trois résidus forment le site de décodage et sont universellement conservés et essentiels <sup>191</sup>.

- Les hélices 8 et 14 pour lesquelles la présence de mutations conduit à une perte de fidélité <sup>192</sup>.

- L'hélice 27 sur laquelle trois mutations améliorent la fidélité de la traduction chez la levure <sup>193</sup>.

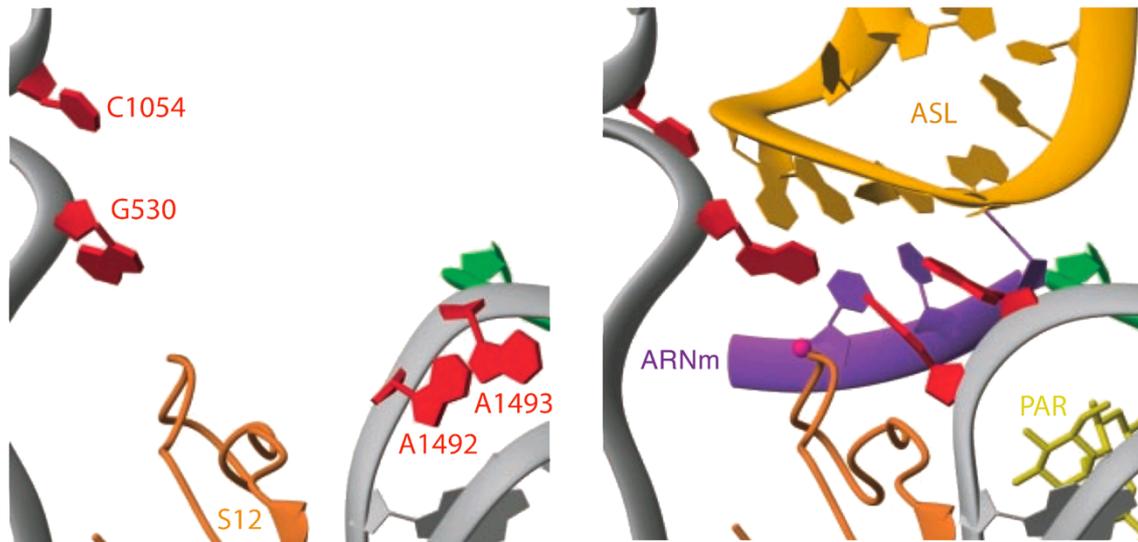
- L'hélice 34 est également impliquée dans la fidélité de la traduction <sup>194</sup>, en particulier le résidu conservé C1054 qui se situe à proximité du centre de décodage. Des mutations sur ce résidu perturbent la terminaison de la traduction chez *Escherichia coli*, mais également chez la levure *S. cerevisiae* <sup>195</sup>.

Plusieurs travaux ont également montré l'importance de résidus modifiés par pseudouridylation de l'hélice 69 de l'ARNr 23S dans la fidélité de la traduction <sup>196</sup>.

Deux protéines ribosomiques ont aussi été mises en évidence pour leurs rôles dans la fidélité de la traduction chez les procaryotes :

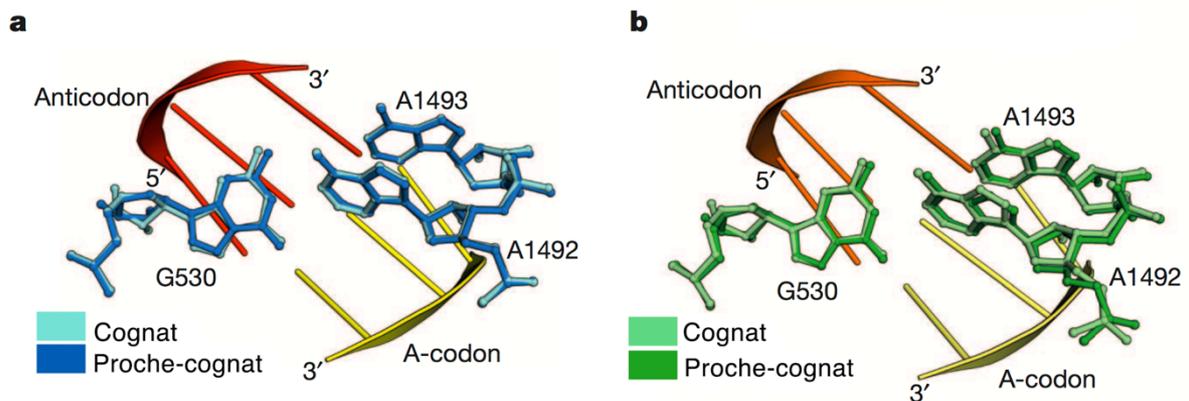
- La protéine eS12 (sous-unité 30S) se situe à proximité du centre de décodage et interagit avec l'ARNr 16S. Une approche par mutagenèse dirigée réalisée chez *E. coli* a montré que deux hélices très conservées de la protéine ont un rôle dans la sélection de l'ARNt au site A. Les mutants obtenus présentent une augmentation de la fidélité de la traduction <sup>197</sup>.

- Des mutants dans les protéines uS5 et uS4, appelés RAM (Ribosome Ambiguity Mutants), sont responsables d'une diminution de la fidélité incluant le passage de codon stop, l'incorporation de mauvais acides aminés et des décalages de phase de lecture <sup>198</sup>.



**Figure 39 : Conformation du site A dans la petite sous-unité 30S**

A gauche, état natif de la sous-unité 30S. Les bases A1492 et A1493 sont empilées à l'intérieur de l'hélice 44, la base G530 est en conformation *syn*. A droite, reconnaissance du codon par l'ARNt. Lorsque le codon et la boucle anticodon se fixent dans le site A du ribosome, A1492 et A1493 basculent à l'extérieur de l'hélice 44 pour vérifier l'interaction codon-anticodon et la base G530 passe en conformation *trans* pour interagir avec A1492, la 2ème position de l'anticodon et la 3ème base du codon. Les bases conservées de l'hélice 44 sont colorés en rose, l'ARNm en violet, la boucle anticodon en jaune, la protéine S12 en marron et les ions Mg<sup>2+</sup> sont représentés par des billes magenta. *Figure adaptée de Ogle et al., 2001*<sup>199</sup>



**Figure 40 : Conformation des bases conservées A1492, A1493 et G530 dans le site de décodage**

**a** - Appariements d'un ARNt cognat (bleu clair) versus un proche cognat portant un mismatch en première position (bleu foncé). **b** - Appariements d'un ARNt cognat (vert clair) versus un proche cognat portant un mismatch en seconde position. La conformation globale des bases universellement conservées G530, A1492 et A1493 de l'ARNr 16S lors de l'appariement d'un ARNt cognat ou proche-cognat (avec un mésappariement en 1<sup>ère</sup> ou 2<sup>ème</sup> position). Les anticodons sont représentés en rouge et orange, les codons au site A sont colorés en jaune et jaune clair. *Figure adaptée de Demeshkina et al., 2012*<sup>46</sup>

### C. Les mécanismes du décodage

L'obtention de structures cristallographiques de la petite sous-unité 30S du ribosome procaryote en complexe avec des ARNt cognat et proche-cognat a permis d'avoir un premier aperçu du fonctionnement du site de décodage <sup>199,200</sup>.

#### *a. Contrôle de l'appariement codon-anticodon par les bases A1492-A1493*

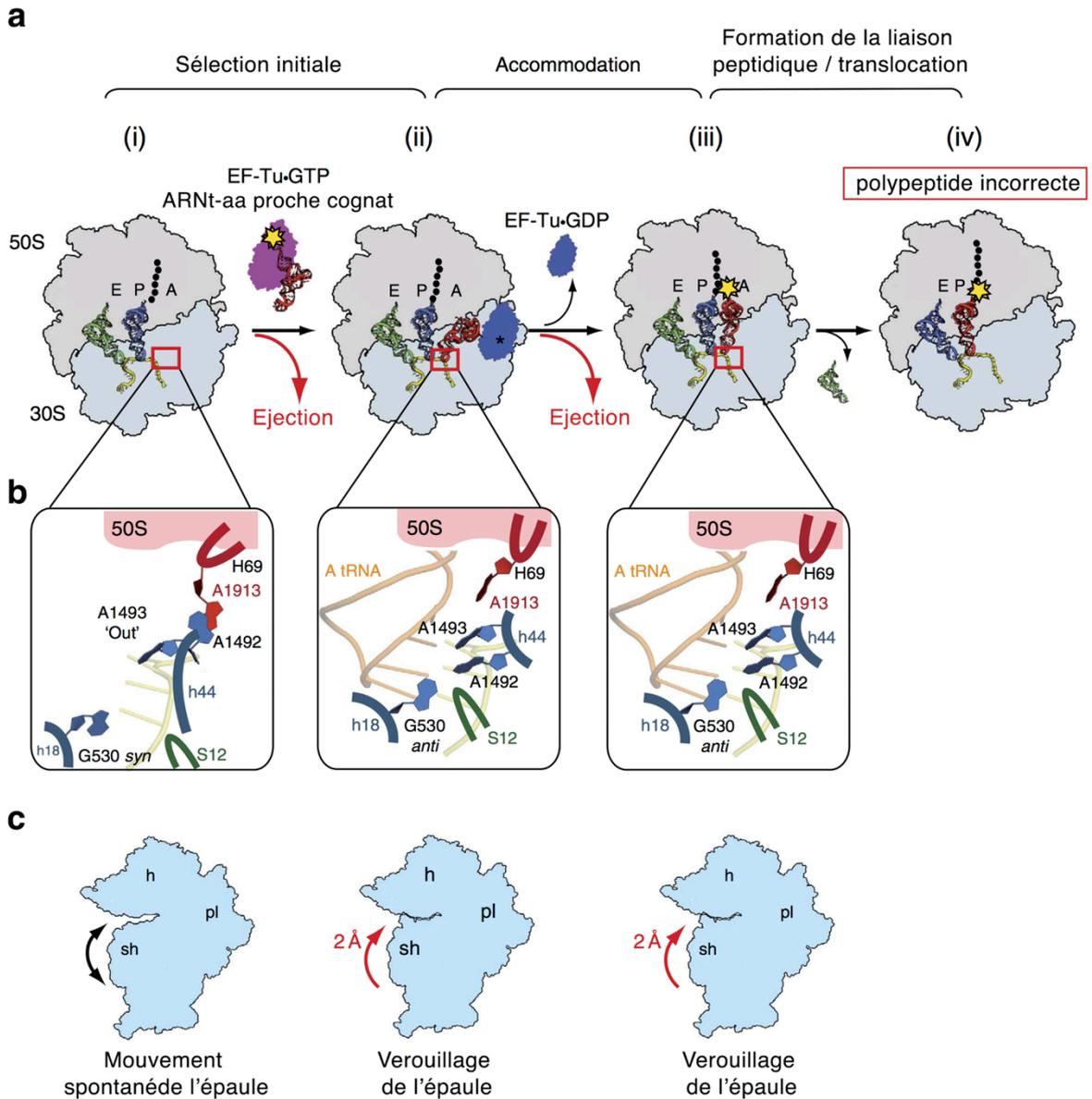
Lors de l'entrée d'un ARNt au site A, la reconnaissance du duplex codon-anticodon correct induit des changements globaux de configuration. La sous-unité 30S passe d'une configuration « ouverte » à une configuration « fermée ». En présence d'un ARNt cognat ou proche-cognat au site A, la tête et l'épaule de la sous-unité 30S se referment autour du centre de décodage et rétrécissent le tunnel d'entrée de l'ARNm <sup>200</sup>. La fermeture de la sous-unité 30S semble être un prérequis pour former le centre de décodage et y induire des changements de configuration locaux <sup>46</sup>.

Au niveau du centre de décodage, la boucle anticodon de l'ARNt (cognat ou proche-cognat) s'apparie au codon du site A, ce qui entraîne un changement de configuration locale des bases A1492, A1493 (hélice 44) et G530 (hélice 18) de l'ARNr 16S (Figure 38). Les bases A1492 et A1493 basculent à l'extérieur de la boucle de l'hélice 44 et interagissent respectivement avec la première et la deuxième paire de bases de l'appariement codon-anticodon. La base G530 passe d'une conformation *syn* à une conformation *trans* et interagit avec la deuxième position de l'anticodon et la troisième base du codon (Figure 39). La stabilité de ce complexe ARNm-ARNt est favorisée par des interactions avec la position C518 de l'hélice 18 et le résidu Ser50 de la protéine ribosomique S12 <sup>199</sup>. Dans ce modèle, les deux premiers appariements de l'hélice codon-anticodon sont donc étroitement vérifiés par le ribosome pour discriminer entre des appariements de type Watson-Crick et des mésappariements, alors que l'environnement de la troisième position (position « wobble ») semble adapté pour accepter d'autres géométries d'appariements.

#### *b. Contrôle de l'appariement codon-anticodon via les contraintes exercées par le centre de décodage*

#### ★ Nouvelles données structurales

Récemment, l'obtention de structures cristallographiques du centre de décodage du ribosome procaryote 70S avec un ARNt cognat ou proche-cognat au site A, en présence d'un ARNt au site P ont remis en cause le modèle précédent <sup>46</sup>.



**Figure 41 : Mécanisme de l'infidélité traductionnelle**

**a** - Mauvaise incorporation d'un acide aminé par le ribosome. Les principales étapes du processus de sélection des ARNt sont présentés, y compris l'hydrolyse du GTP (en astérisque noir) par le facteur d'élongation EF-Tu et l'interaction codon-anticodon dans le centre de décodage (i-iv) **b** - Conformation des principaux nucléotides du centre de décodage sans ARNt (à gauche), appariés à un ARNt cognat ou proche-cognat au cours de l'étape de reconnaissance (milieu) et à l'étape finale de l'accommodation (à droite). Les nucléotides essentiels de l'ARNr 16S et 23S sont indiqués respectivement en cyan et en rouge. La protéine ribosomique S12 qui restreint l'appariement à la seconde position du codon, est représentée en vert. Les trois nucléotides du codon d'ARNm dans le site A sont numérotés selon le système standard. **c** - Vue globale de la 30S lors de la sélection de l'ARNt. Le panneau de gauche montre le mouvement spontané du domaine de l'épaule (flèches noires) lorsque le centre de décodage est inoccupé; les panneaux du milieu et de droite montrent que l'épaule est déplacée et stabilisée lors de la liaison d'un ARNt proche-cognat lors des étapes de sélection initiale et d'accommodation. sh, h et pl désignent respectivement les domaines de l'épaule, la tête et la plate-forme de la petite sous-unité ribosomique. *Figure adaptée de Rozov et al., 2015* <sup>47</sup>

Tout d'abord, l'ARNm formerait un coude entre les codons au site A et au site P et serait stabilisé par l'ARNt au site P, l'ARNr 16S et un ion  $Mg^{2+}$ , qui contraindrait le positionnement de la première base du codon (Figure 38) <sup>201</sup>. De plus, contrairement à ce qui était pensé jusqu'ici, les appariements U-G et G-U en première ou deuxième position du codon seraient possibles et n'adopteraient pas une conformation de type wobble, mais formeraient un duplex de type Watson-Crick. D'autre part, les bases A1492, A1493 et G530 de l'ARNr 16S interagissent de la même manière avec ces paires U-G et G-U qu'avec des appariements canoniques Watson-Crick. Le mouvement des bases A1492 et A1493 ne permettrait pas de vérifier la spécificité de l'interaction codon-anticodon, mais formerait en réalité une partie statique du centre de décodage, définissant ses propriétés spatiales et chimiques (Figure 40). Les contraintes imposées par le centre de décodage au niveau des bases 1 et 2 du codon sur les appariements des ARNt proche-cognat proviendraient de différences de géométrie par rapport aux appariements des ARNt cognat. Les distorsions induites par les mésappariements se propageraient dans toute la structure de l'ARNt, ce qui le déstabiliserait et faciliterait son rejet hors du ribosome.

Les auteurs expliquent que les nouvelles structures 70S ayant permis l'élaboration de ce modèle diffèrent des précédentes qui étaient basées uniquement sur la petite sous-unité 30S <sup>46</sup>.

#### ★ Précisions sur le mécanisme de contrôle de l'appariement codon/anticodon

Dans un troisième modèle <sup>202</sup>, la grande sous-unité ribosomique et en particulier l'hélice 69 joue un rôle crucial. Cette hélice qui forme un pont entre les deux sous-unités agit comme un régulateur du résidu A1492 de l'ARNr 16S dans le centre de décodage. La sélection de l'ARNt débute par la liaison de l'ARNt au site A libre (Figure 41a,i) où le résidu A1493 est en position pour interagir avec le duplex codon/anticodon. Le résidu A1492 est positionné sur le résidu A1913 de l'hélice 69 de l'ARNr 23S (Fig. 41b, à gauche). Le coude de l'ARNm entre les sites A et P contraint le positionnement de la première base du codon. L'entrée d'un ARNt cognat ou non cognat provoque alors deux réarrangements majeurs : Tout d'abord, lors de la fixation de l'ARNt, la boucle anticodon entre en contact avec l'hélice 69 et casse l'interaction A1492-A1913. Ceci provoquerait entre autres le repositionnement du résidu A1492 qui interagirait avec le codon en deuxième position (Figure 41b, milieu). Le second réarrangement correspond au léger mouvement du domaine de l'épaule de la petite sous-unité vers la boucle anticodon (Figure 41c), ce qui rapproche le résidu G530 au niveau de la seconde position du codon et finalise la formation du centre de décodage

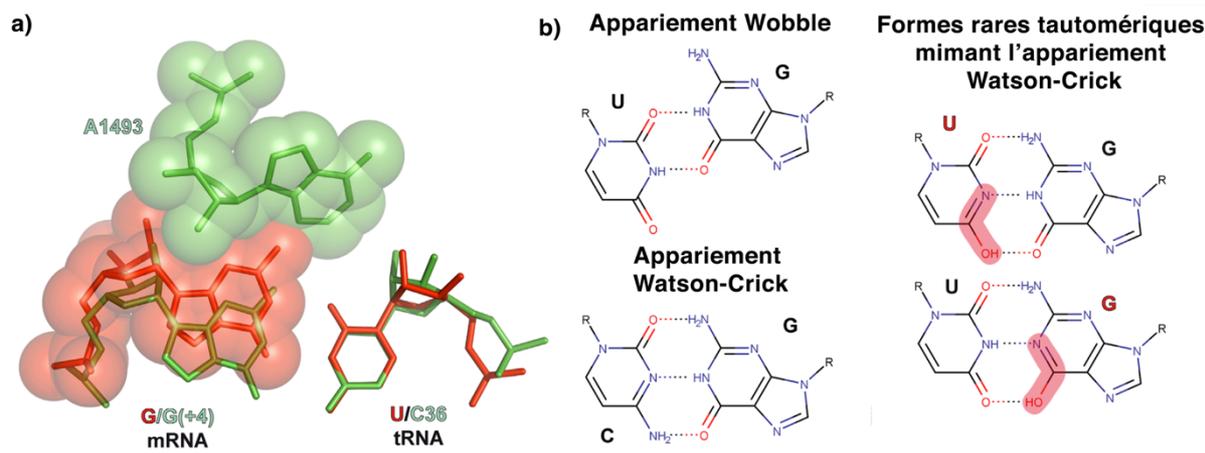


Figure 42 : Comparaison des appariements Watson-Crick, wobble et tautomérique rare

a) Hypothétique superposition d'un appariement wobble (G-U) en rouge sur un appariement canonique en première position (G-C) en vert, ce qui démontre un clash possible avec le résidu A1493 de l'ARN 16S. b) Géométrie d'un appariement non canonique wobble<sup>203</sup>, canonique Watson-Crick (C-G) et d'un appariement de type Watson-Crick formé par des états tautomères rares de l'uracile et de la guanine. Cette forme rare permet un appariement stable en première et deuxième position de l'appariement G-U. Figure adaptée de Rozov et al., 2016<sup>204</sup>

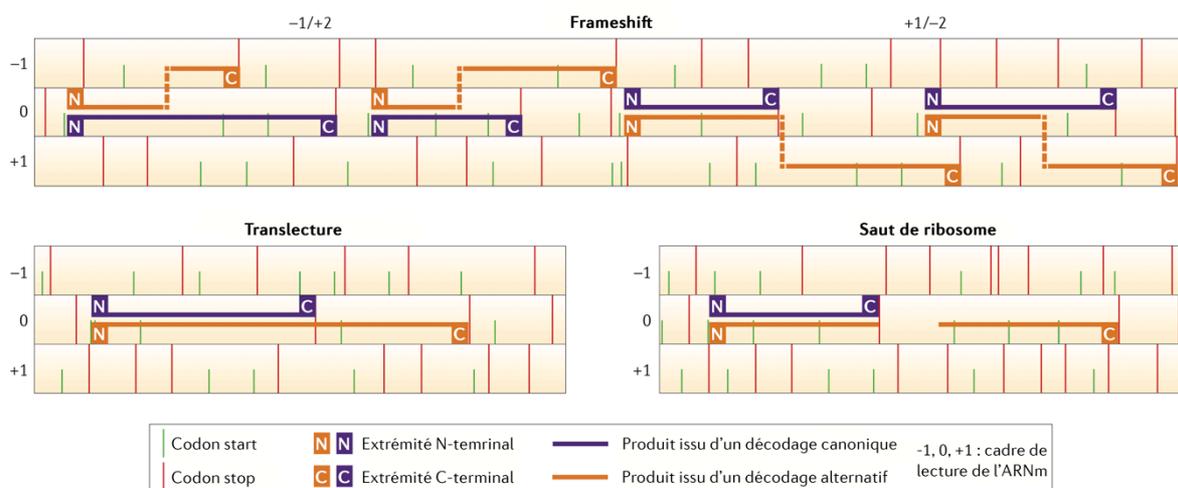


Figure 43 : Relation entre séquences nucléotidiques et protéines issues d'un décodage alternatif

Les séquences nucléotidiques des trois principaux événements de recodage sont représentées avec leurs trois cadres de lecture différents. Le codon start (vert) et stop (rouge) sont représentées par des lignes verticales. Les séquences qui sont traduites en protéines sont présentées sous forme de barres horizontales dont les extrémités amino- (N-ter) et carboxy-terminales (C-ter) sont indiquées. Les barres violettes correspondent au décodage standard et les barres oranges à des événements alternatifs. Les transitions entre les cadres de lecture sont indiquées par des pointillés. Figure adaptée de Baranov et al., 2015<sup>205</sup>

(Figure 41b, milieu). Ce mouvement permettrait également de bloquer l'ARNt, de telle sorte que la complémentarité du duplex codon-anticodon soit testée. Un ARNt serait alors efficacement incorporé en raison de sa capacité à former un appariement Watson Crick stable sur les deux premières positions du codon. A contrario, la plupart des ARNt proche-cognat seront éjectés en raison de l'encombrement stérique au niveau de l'appariement codon-anticodon ou du coût énergétique nécessaire pour que l'anticodon de l'ARNt soit correctement positionné dans le centre de décodage (Figure 41a, i-ii). Toutefois, certains ARNt proche-cognat échapperaient à cette discrimination du fait de leur capacité à former une interaction Watson-Crick via une forme tautomère rare de leurs bases (Figure 42). Ces ARNt seraient alors accommodés au site A puis transloqués au site P, ce qui conduirait à l'incorporation d'un acide aminé incorrecte dans la chaîne polypeptidique (Figure 41a, iii-iv).

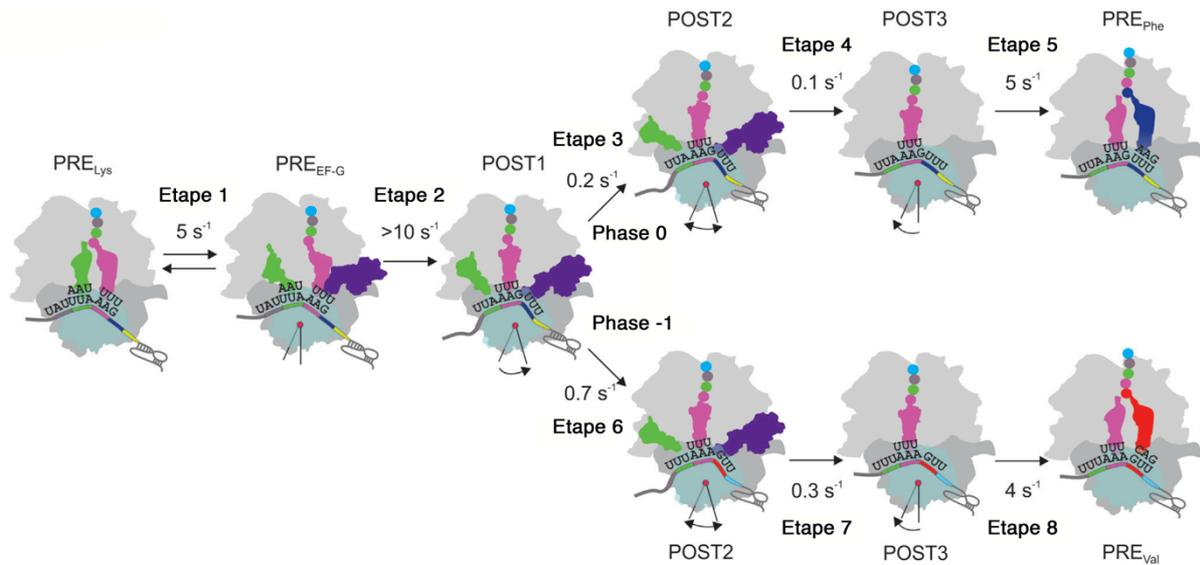
### 3. Le recodage

Alors que la fidélité de la traduction est généralement élevée, certains ARNm contiennent des signaux qui altèrent la lecture du code génétique en modifiant les règles de décodage et qui ont pour conséquence la synthèse d'une protéine différente. Ces erreurs programmées de la traduction sont regroupées sous le terme « recodage » et sont en compétition avec les règles classiques de la traduction. Par conséquent, seule une fraction des ribosomes effectuera du recodage. Ce processus a été découvert chez les virus qui l'utilisent fréquemment pour réguler la synthèse de leurs protéines. Quatre types d'événements de recodage peuvent se produire au cours de la traduction : le décalage du cadre de lecture en +1 ou en -1 et le saut de ribosome qui ont lieu durant l'élongation et le passage du codon stop qui a lieu lors de la terminaison <sup>206</sup> (Figure 43).

#### A. Le décalage du cadre de lecture (Frameshift)

Le décalage du cadre de lecture, plus couramment appelé « Frameshift », correspond au décalage du ribosome d'un nucléotide en 5' (Frameshift -1) ou en 3' (Frameshift +1) et à la poursuite de l'élongation dans une autre phase de lecture que celle de départ.

La plupart des cas frameshift nécessitent à la fois une séquence glissante qui favorise le glissement des ARNt et un élément stimulateur qui augmente l'efficacité du glissement en induisant une pause du ribosome. En effet, en absence d'éléments stimulateurs, la fréquence de cet événement est relativement faible, environ  $3.10^{-5}$



**Figure 44 : Modèle cinétique du frameshift -1**

Partitionnement cinétique du frameshift -1. Le glissement se produit au cours de la translocation des deux ARNt liés à la séquence glissante (tRNA<sup>Leu</sup> et du peptidyl ARNt<sup>Lys</sup>). Tout d'abord, il y a recrutement de EF-G (étape 1) sur le complexe de pré-translocation (PRE), ce qui facilite le mouvement rapide de l'ARNt (étape 2) dans un état chimérique de post-translocation (POST1). Toutefois, les étapes suivantes sont inhibées par la présence du pseudo-noeud. Puis, l'ARNt<sup>Leu</sup> poursuit son mouvement découpées ici en deux étapes. Tout d'abord, l'ARNt<sup>Leu</sup> se déplace sur la sous-unité 50S dans un état POST2 sans que sa distance à la sous-unité 30S ne change (étapes 3 et 6 respectivement en phase 0 et -1). Deuxièmement, l'ARNt<sup>Leu</sup> et la sous-unité 30S s'écartent (étapes 4 et 7) dans un état de POST3. Les étapes 3 et 4 sont particulièrement lentes pour l'ARNt qui reste dans en phase 0, ce qui limite le taux de fixation de l'ARNt<sup>Phe</sup> (étape 5). En revanche, le mouvement de l'ARNt<sup>Leu</sup> sur les ribosomes qui se sont décalés en phase -1, est plus rapide (étape 6). Ce mouvement est suivi de la dissociation de l'ARNt<sup>Leu</sup> de la sous-unité 30S, puis de la rotation de la du domaine tête de la sous-unité 30S, de la dissociation de EF-G (étape 7) et enfin de la liaison de l'ARNt<sup>Val</sup> (étape 8). *Figure adaptée de Caliskan et al., 2014*<sup>207</sup>

par codon. Il existe plusieurs types d'éléments stimulateurs dont les plus connus sont les structures secondaires simples de types tige-boucle, ou les pseudonœuds<sup>208</sup>. Les peptides naissants sont également connus pour faciliter le frameshift<sup>209</sup>. La nécessité de la présence de ces éléments stimulateurs pour activer le frameshift fournit un mode de régulation de la traduction très important. Grâce à la présence d'éléments régulateurs positifs et négatifs, il est possible de moduler très finement l'expression de certains gènes.

#### a. Le décalage du cadre de lecture en -1

Le frameshift -1, dont le mécanisme est le mieux compris, est un événement durant lequel les ARNt se décalent d'une base en 5' de l'ARNm. Le ribosome a besoin d'au minimum deux signaux pour avoir une efficacité de décalage significative : une séquence glissante et une structure secondaire en 3'. La séquence glissante correspond à la région sur laquelle le ribosome va effectuer une pause et se décaler d'un nucléotide en 5'. Cette séquence est généralement heptamérique et de consensus X XXY YYZ, où X=N, Y=A ou U et Z≠G. L'élément stimulateur peut être une structure secondaire de type pseudo-nœud ou tige-boucle située en 3' de la séquence de glissement. La présence de cet élément permet d'augmenter de 10 à 30 fois l'efficacité de décalage de la séquence glissante<sup>210</sup>. Chez les eucaryotes, la distance optimale entre ces 2 signaux varie entre 6 et 8 nucléotides et permettrait au ribosome d'exercer son activité hélicase au niveau de la structure secondaire<sup>211</sup>.

Les données issues de la structure cryo-EM d'un ribosome de mammifère au cours d'un événement de frameshift -1 ont montré que le pseudo-nœud bloquait le ribosome à une étape intermédiaire de la translocation, provoquant un repliement de l'ARNt au site A et contraignant le ribosome à se décaler d'un nucléotide en 5'<sup>212</sup>. De plus, ces résultats ont mis en évidence l'interaction directe du facteur d'élongation eEF2 avec l'ARNt au site P. Récemment, les résultats de cinétiques et de smFRET ont permis d'affiner ce modèle du mécanisme de frameshift -1 chez les procaryotes (Figure 44)<sup>213,214</sup>. Le facteur d'élongation procéderait à de multiples tentatives de liaison et resterait lié au ribosome jusqu'à ce que la translocation soit complète. Le ribosome serait bloqué dans un état pivoté et le mouvement vers l'arrière de la sous-unité 30S serait entravé. Ainsi, lors du frameshift -1, les ribosomes sont piégés dans un état métastable dans lequel les mouvements de la 30S sont entravés très probablement par le pseudo-nœud/tige-boucle bloquant l'entrée du tunnel de l'ARNm. Il y a alors deux voies possibles pour libérer le ribosome bloqué dans cet état métastable : soit le ribosome attend que l'ARNm se déroule, soit, l'appariement codon-anticodon se rompt et l'ARNm glisse en 5' d'un nucléotide. Le second scénario serait plus avantageux,

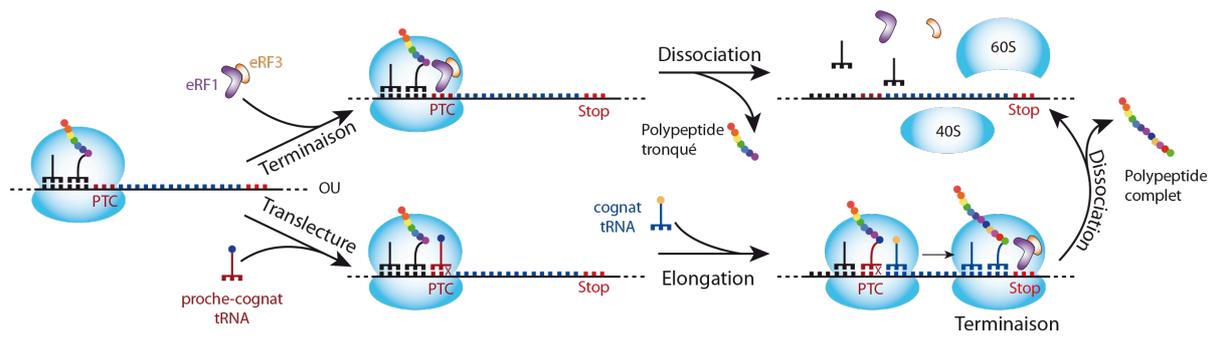
car il permet de déplacer la base de la structure secondaire à l'entrée du tunnel de l'ARNm de telle sorte que le centre hélicase du ribosome déroule activement la structure secondaire.

Le mécanisme de frameshift -1 a été retrouvé essentiellement chez les virus et les éléments transposables. L'exemple le plus connu est le VIH-1 (Virus de l'Immunodéficience Humaine de type 1), qui utilise ce mécanisme pour réguler le ratio entre les protéines Gag et Gag-Pol. Un seul exemple de frameshift -1 ayant un rôle biologique a été identifié chez les procaryotes : le gène *dnaX* qui code les sous-unités  $\gamma$  et  $\tau$  de l'ADN polymérase III. La synthèse de la protéine  $\gamma$  résulte du frameshift -1 qui conduit les ribosomes à s'arrêter à un codon stop présent uniquement en phase -1. La forme longue  $\tau$  est traduite par décodage classique. Chez les eucaryotes, il existe deux gènes cellulaires eucaryotes connus pour être régulés par frameshift -1. Il s'agit du gène *Edr* de souris (Embryonal carcinoma Differentiation Regulated), indispensable au développement prénatal<sup>215</sup>, et du gène humain *MA3*<sup>216</sup>.

#### b. Le décalage du cadre de lecture en +1

Le frameshift +1 est un événement durant lequel les ARNt se décalent d'une base en 3' sur l'ARNm. Le décalage de cadre de lecture se produit sur une séquence minimale tétramérique XXX X. Dans la plupart des cas, le codon adjacent stimule le décalage et la séquence forme alors un heptamère XXX XXX X. Contrairement au frameshift -1, il n'y a pas de séquence consensus<sup>217</sup>, mais les codons favorisant le glissement ont été identifiés<sup>218</sup>. Deux mécanismes peuvent expliquer un événement de décalage du cadre de lecture en +1 : soit l'ARNt au site P se décale d'un nucléotide en 3', soit l'ARNt rentrant au site A s'apparie directement dans la phase +1<sup>219</sup>. La différence réside dans l'affinité de l'ARNt au site P pour son codon et dans sa capacité à glisser. Quelle que soit le mécanisme impliqué, le frameshift +1 nécessite une reconnaissance lente du codon présent au site A ainsi qu'un décodage efficace du codon situé en phase +1. Celle-ci peut être déclenchée par la présence d'un codon rare, par un codon stop peu efficace ou par une structure secondaire de type tige-boucle ou pseudonoeud.

Chez les procaryotes, le seul cas connu est celui du gène *prfB* codant le facteur de terminaison RF2 (Release Factor 2). Le frameshift +1 permet à RF2 d'autoréguler son abondance<sup>220</sup>. Plusieurs gènes cellulaires sont régulés par frameshift +1 chez les eucaryotes : c'est le cas de l'antizyme OAZ qui régule la concentration en polyamines, des molécules indispensables au bon fonctionnement de nombreux processus cellulaires et à la croissance<sup>221</sup>. Chez la levure, les gènes *EST3* et *ABP140* sont régulés



**Figure 45 : La translecture**

Représentation schématique de la terminaison de la traduction et du processus de translecture. Lorsqu'un codon stop entre au site A du ribosome, qu'il soit prématuré<sup>22</sup> ou non, il y a une compétition entre le facteur de terminaison eRF1 et un ARNt proche-cognat. S'il y a terminaison, les facteurs de terminaison eRF1 et eRF3 entrent au site A du ribosome et induisent la terminaison de la traduction ainsi que la dissociation des deux sous-unités ribosomique. Toutefois, un ARNt proche-cognat peut également entrer au site A du ribosome et décoder le codon stop, ce qui provoque de la translecture. Le ribosome continue alors sa traduction dans la même phase de lecture jusqu'au prochain codon stop. *Figure adaptée de Bidou et al., 2012*<sup>222</sup>

par ce mécanisme. Pour ces deux gènes, le frameshift +1 est nécessaire à l'expression du gène et à la synthèse de la protéine active<sup>223,224</sup>.

## B. Le passage du codon stop

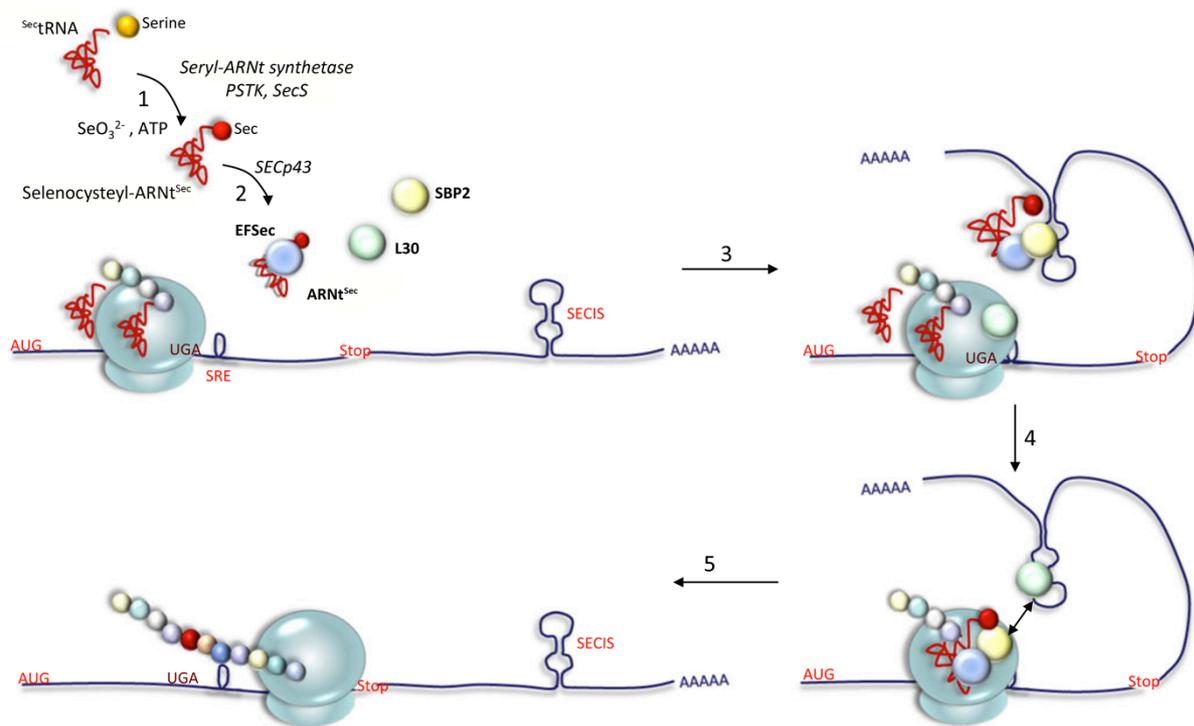
Les facteurs de terminaison sont capables de reconnaître les codons stop très efficacement, ce qui permet la synthèse de protéines complètes et fonctionnelles. Toutefois, plusieurs événements modifient le décodage du codon stop en permettant au ribosome de lire ces codons et d'y incorporer un acide aminé. Le ribosome continue alors sa traduction en phase jusqu'au prochain codon stop.

### a. La translecture

Lorsqu'un codon stop entre au site A du ribosome, il y a une compétition entre le facteur de terminaison eRF1 et un ARNt proche-cognat. Le processus d'incorporation d'un ARNt proche-cognat est appelé translecture et permet de poursuivre la synthèse protéique, en phase, jusqu'au prochain codon de terminaison (Figure 45). De manière générale, la compétition est largement favorable à la terminaison qui est plus lente et plus précise que l'élongation. En effet, les facteurs de terminaison sont capables de différencier les codons UGG des codons UGA alors que seul le résidu en position wobble est différent, ce qu'un ARNt ne serait pas capable de réaliser. Au niveau d'un codon stop, le taux de translecture de base se situe généralement entre 0,01% et 0,1% chez les mammifères et 0,3% chez la levure<sup>225</sup>. Néanmoins cette fréquence peut être influencée par des éléments en *cis* et *trans* qui seront détaillés dans la section suivante :

#### 4. Les facteurs connus pour modifier la fidélité de la terminaison de la traduction.

Le mécanisme de translecture est utilisé de manière contrôlée par beaucoup de virus de plante à ARN, c'est le cas du virus de la mosaïque du tabac (TMV) pour lequel la translecture est indispensable à la synthèse de la réplicase<sup>226</sup>. La séquence identifiée dans ce virus permet une efficacité de translecture de 2 % dans les cellules de souris<sup>227</sup> 4 à 5 % chez les plantes<sup>226</sup> et de 15 à 25% chez *S. cerevisiae*<sup>228</sup>. Chez les eucaryotes, il existe quelques gènes cellulaires régulés par translecture, ce qui pourrait suggérer que ce phénomène est utilisé comme moyen de régulation de l'expression génique. En particulier, deux gènes possédant un motif compatible avec un événement de translecture ont été identifiés chez *S. cerevisiae*. Il s'agit des gènes PDE2 (PhosphoDiEsterase 2) et IMP3 (Interacting with Mpp10p). La translecture du codon stop de PDE2 ajoute une courte séquence de 20 acides aminés qui déstabilise la protéine et induit sa dégradation<sup>229</sup>.



**Figure 46 : Synthèse et incorporation de la sélénocystéine chez les eucaryotes**

La synthèse de la sélénocystéine a lieu sur son ARNt<sup>Sec</sup> spécifique <sup>22</sup>. L'interaction de l'ARNt<sup>Sec</sup> avec les facteurs de traduction Sec-spécifiques EFSec et SBP2, en présence de SECp43 (2, 3), permet son recrutement au site A du ribosome <sup>22</sup>. Lorsque le résidu sélénocystéine est incorporé dans la séquence du peptide en réponse au codon reprogrammé UGA, la traduction reprend jusqu'au prochain codon stop (5). *Figure adaptée de Castets et al., 2012* <sup>230</sup>

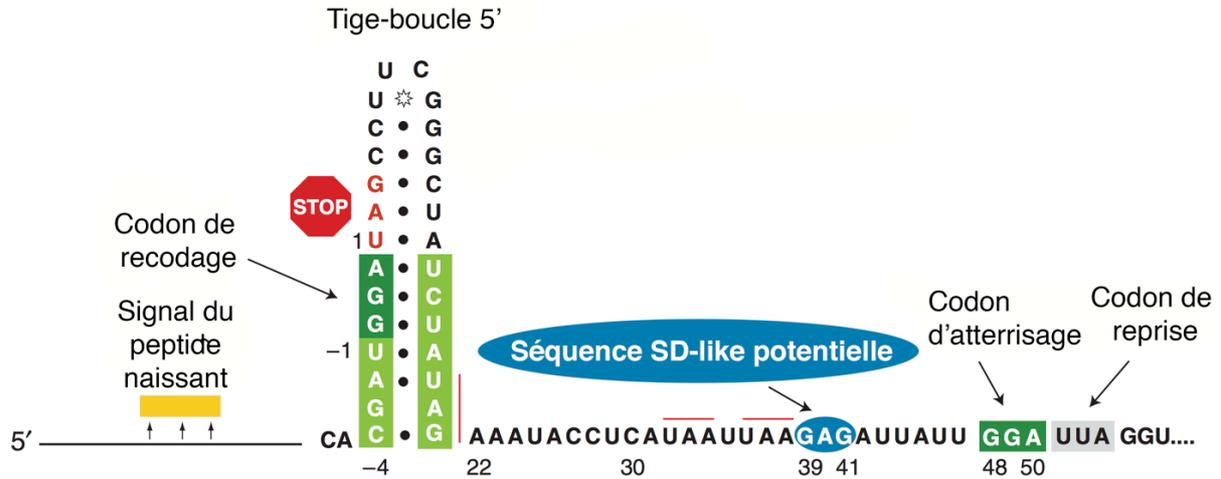
Récemment, il a été proposé que eIF3, le facteur d'initiation de traduction eucaryote, pourrait être impliqué dans la translecture programmée des 3 codons stop dans un contexte défavorable. Dans ce mécanisme, qui semble être évolutivement conservé, eIF3 interagit avec le complexe de prétermination, où il empêche eRF1 de reconnaître la troisième position wobble du codon stop. En conséquence, un ARNt proche-cognat ayant un mésappariement à cette position peut décoder le codon stop, ce qui permet de poursuivre la synthèse de la protéine <sup>231</sup>.

#### *b. Incorporation de la sélénocystéine*

La sélénocystéine (Sec ou U) est le 21<sup>ème</sup> acide aminé naturel du code génétique, et dérive de la sérine. Elle est incorporée dans les sélénoprotéines qui protègent les cellules contre les stress oxydatifs. Le sélénium est un micronutriment essentiel et un élément trace pour la plupart des organismes, incluant les humains qui possèdent pourtant 25 gènes codant les sélénoprotéines. Ils sont souvent évoqués dans un contexte médical, car une déficience en sélénocystéine est responsable de plusieurs maladies graves tels que la maladie de Keshan et Kashin-Beck. Des études épidémiologiques ont rapporté que le sélénium protège contre le cancer et d'autres maladies neurodégénératives grâce à ses propriétés antioxydantes <sup>232</sup>.

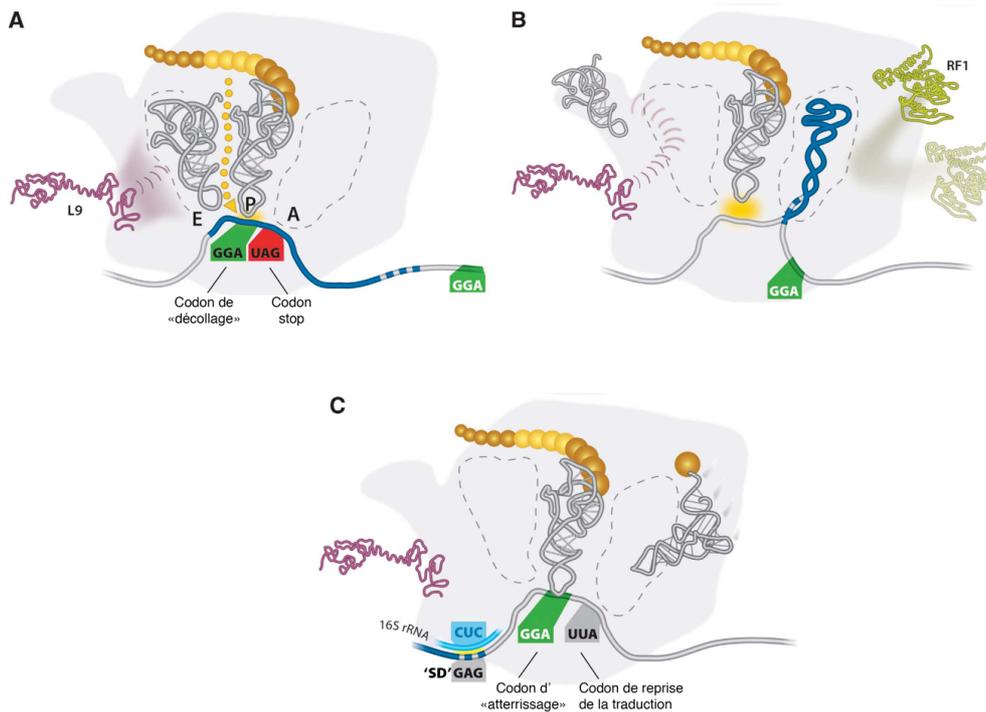
L'incorporation de la sélénocystéine se fait par translecture du codon stop UGA par un ARNt<sup>[Ser]Sec</sup> spécifique. C'est le seul ARNt connu qui gouverne sa propre voie de biosynthèse. Chez les eucaryotes, l'expression de l'ARNt<sup>[Ser]Sec</sup> nécessite quatre enzymes contre une seule pour les autres ARNt. SerRs permet de charger une Sérine sur l'ARNt, puis PSTK phosphoryle la sérine qui est ensuite convertie en sélénocystéine par SepSecS. Enfin, l'activation du sélénium est réalisée par SPS1/SPS2 <sup>233</sup>. En revanche, la synthèse de l'ARNt<sup>[Ser]Sec</sup> est réalisée en une seule étape chez les bactéries.

Chez les bactéries, l'incorporation de l'ARNt<sup>[Ser]Sec</sup> dépend de la présence du codon UGA, d'une structure tige-boucle appelée SECIS (Selenocysteine Insertion Sequence) située immédiatement en 3' du codon UGA, et du facteur SelB <sup>206</sup>. Chez les eucaryotes, le facteur d'élongation EFSec se fixe spécifiquement sur l'ARNt<sup>[Ser]Sec</sup> avec l'aide du facteur SBP2 (SECIS Binding Protein 2) <sup>234</sup> et interagit avec une des deux structures SECIS connues situées dans le 3'UTR <sup>233</sup> (Figure 46).



**Figure 47 : Site du recodage du gène 60 chez le Bactériophage T4**

Le « saut » du ribosome au codon GGA (en vert) est déclenché par la présence du codon stop UAG (en rouge) au site A du ribosome. Il est également stimulé par des signaux contenus dans le peptide naissant (en jaune) et la structure secondaire en tige-boucle. Le ribosome « atterrit » au codon GGA suivant, 50 nucléotides plus loin, en aval d'une potentielle séquence Shine-Dalgarno-like (en bleu), pour reprendre la traduction au codon UUA (en gris). *Figure adaptée de Wills et al., 2008* <sup>235</sup>



**Figure 48 : Mécanisme du saut de ribosome**

**A** - La présence du codon stop UAG (en rouge) au site A va induire une pause permettant à l'ARNm de former une structure secondaire représentée dans la partie B de la figure. **B** - Les appariements intra-ARNm vont induire la formation d'un tige-boucle (en bleu) qui va empêcher d'RF1 au site A et induire la sortie de l'ARNt du site E. Le mouvement vers l'avant de l'ARN va permettre de résoudre la structure. **C** - Le retour à un ARNm linéaire et l'appariement du codon GAG au 3' de l'ARNr 16S (bleu clair) favorisent le recrutement de l'ARNt peptidyl au codon GGA. Le décodage standard reprend au codon UUA adjacent en 3'. *Figure adaptée de Wills et al., 2008*

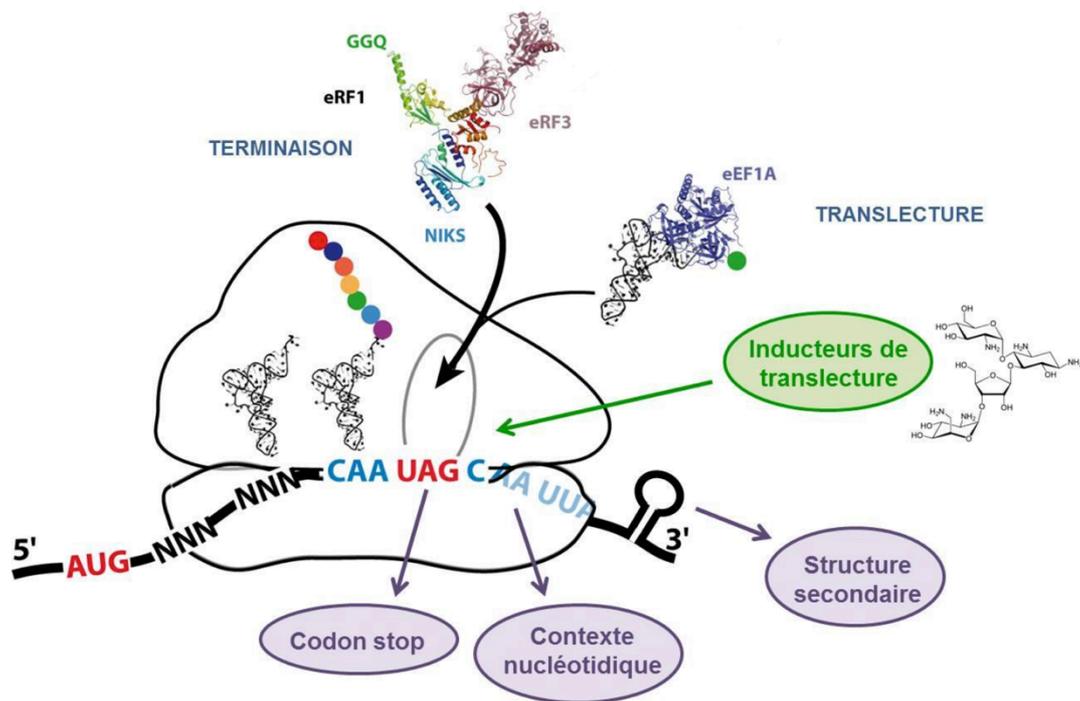
### c. Incorporation de la pyrrolysine

La pyrrolysine (Pyl ou O) est le 22<sup>ème</sup> acide aminé naturel et dérive de l'acide aminé lysine. Elle est retrouvée seulement chez les espèces d'archées méthanogènes *Methanosarcinna* et chez les espèces bactériennes *Desulfitobacterium*. Les archées utilisent la pyrrolysine pour produire du méthane. Ce processus nécessite l'expression des plusieurs gènes dont trois (MttBn MtbB et MtmB) contiennent un codon stop UAG qui doit être traduit<sup>236</sup>. Le codon stop UAG est reconnu par un ARNt<sup>Pyl</sup> spécifique à la pyrrolysine qui est incorporé lors de la traduction sans signaux spécifiques. Cette incorporation est particulièrement élevée et pourrait être expliquée par la présence d'une tige-boucle PYLIS (pyrolysine insertion elements) en aval du codon stop<sup>206</sup>. L'ARNt<sup>Pyl</sup> est plus court que les autres ARNt naturels et possède donc une structure secondaire légèrement différente. La tige de l'anticodon est formée de 6 paires de bases au lieu de 5 et certaines modifications conservées universellement sont absentes. Toutefois, l'ARNt<sup>Pyl</sup> est reconnu par le facteur d'élongation classique<sup>236</sup>.

### C. Le saut de ribosome

Le saut de ribosome, également appelé « Bypassing » est un événement très rare qui a été décrit initialement dans le gène 60 du bactériophage T4. Ce gène code une topoisomérase essentielle pour ce phage infectant la bactérie *E. coli*. Au cours de l'élongation, des éléments en *cis* et *trans* vont permettre au ribosome de suspendre l'élongation sur un codon de « décollage » GGA et de la reprendre 50 nucléotides en aval sur un codon d' « atterrissage » GGA. L'efficacité de cet événement est de 50 %<sup>237</sup> (Figure 47).

Dans le modèle proposé par Wills et ses collaborateurs, le codon stop induirait une pause du ribosome lors de son entrée au site A. Le peptide naissant déstabiliserait suffisamment le ribosome et l'interaction codon-anticodon au site P pour permettre à la structure secondaire de se reformer dans le site A. Cela implique que les extrémités 5' et 3' de l'ARNm soient tirées à l'intérieur du ribosome. Une fois la structure formée, le ribosome serait capable de détecter que son site A est occupé et engagerait un cycle de translocation permettant une reprise de la traduction au codon d'atterrissage. Cette reprise serait facilitée par la présence d'une région de type Shine-Dalgarno<sup>235</sup> (Figure 48).



**Figure 49 : Les facteurs modulant la fidélité de la terminaison de la traduction**

La compétition entre terminaison et translecture peut être modulée par des facteurs agissant en *cis* ou en *trans*. Les facteurs *cis* présents sur l'ARNm sont la nature du codon stop, son contexte nucléotidique ou encore la présence d'une structure secondaire de type tige-boucle. Parmi les facteurs *trans*, les molécules inductrices de translecture telles que les aminoglycosides peuvent impacter l'efficacité de la terminaison. Les facteurs *cis* sont indiqués en violet, les facteurs *trans* en vert. Sur l'ARNm, le codon stop est en rouge et le contexte nucléotidique en bleu. *Figure issue du mémoire de thèse de Sandra Blanchet, 2014*

## 4. Les facteurs connus pour modifier la fidélité de la terminaison de la traduction

Dans la cellule, la terminaison de la traduction est toujours en compétition avec la translecture qui correspond à l'incorporation d'un ARNt proche-cognat au niveau du codon stop et à la poursuite de la traduction dans la même phase. Le processus de décodage étant très fidèle, cette compétition est généralement favorable à la terminaison. Toutefois, certains facteurs peuvent moduler son efficacité soit en favorisant l'incorporation d'un ARNt proche-cognat soit en défavorisant la reconnaissance des codons stop. Deux types de facteurs ont été mis en évidence : les facteurs *cis* tels que la nature du codon stop, le contexte nucléotidique en 5' et en 3' du codon stop et la présence de structures secondaires. Des facteurs agissant en *trans* peuvent également affecter la fidélité de la traduction, c'est notamment le cas de certaines molécules telles que les antibiotiques de la famille des aminoglycosides<sup>238</sup> (Figure 49).

### A. Les facteurs agissant en *cis*

#### a. La nature du codon stop

Les trois codons stop ne sont pas reconnus par les facteurs de terminaison avec la même efficacité<sup>239</sup>. Plusieurs travaux ont montré que le codon UAA est celui qui induit la meilleure efficacité de terminaison, suivi des codons UAG et UGA<sup>72,240</sup>. Ces résultats s'expliqueraient à la fois par le nombre d'ARNt naturels suppresseurs proche-cognat présents dans la cellule et par leur capacité à reconnaître les codons stop.

Il y a plusieurs ARNt proche-cognat capable de reconnaître chacun des codons stop, mais leur utilisation dans le processus de translecture ne semble pas être aléatoire. Plusieurs études<sup>241-244</sup>, principalement sur des ARNm viraux et de levure, ont montré que le codon UGA est le plus fréquemment décodé par l'ARNt<sup>Trp</sup> (W), mais il peut également être lu par un ARNt<sup>Cys</sup> (C) ou ARNt<sup>Arg</sup> (R). Le codon stop UAG est décodé principalement par les ARNt<sup>Gln</sup> (Q) et ARNt<sup>Trp</sup> (W), mais les ARNt<sup>Tyr</sup> (Y) et ARNt<sup>Lys</sup> (K) ont aussi été retrouvés. Enfin, le codon UAA serait décodé par l'ARNt<sup>Gln</sup> (Q). Une étude plus récente réalisée dans l'équipe chez la levure, a utilisée une approche basée sur la spectrométrie de masse pour identifier les acides aminés incorporés au codon stop dans un système rapporteur *in vivo*<sup>77,245</sup>.

CODON STOP						
UGA	>	UAG	>	UAA	SYSTEME UTILISE	REFERENCE
C>U~A>G		G>C>U=A		C>U~A>G	Yeast	Bonetti et al.
C~U>G≥A		C≥U>>G≥A		C≥U>>G>A	Mammalian	McCaughan et al.
C>>A>U~G		U>C>G~A		C>G>U~A	Mammalian	Manuvakhova et al.
C>A~G~U		C>A~G~U		C>A~G~U	Mammalian	Floquet et al.
C>A>G>U					Mammalian	Beznoskova et al.

**Tableau 1 : Influence de la nature du nucléotide +4 sur la translecture basale**

Le codon stop correspond aux bases +1, +2 et +3. Les résultats des travaux obtenus sur l'influence de la nature du nucléotide +4 sur la translecture sont présentés dans ce tableau. Les nucléotides sont classés par ordre d'influence sur la translecture pour chacun des trois codons stop. Figure adaptée de *Dabrowski et al., 2015*<sup>246</sup>

Nous avons confirmé que le codon UGA est décodé par le tryptophane, la cystéine ou l'arginine, mais nous avons aussi démontré que les codons UAA et UAG peuvent être décodés par la glutamine, la tyrosine ou la lysine. Les mêmes ARNt proche-cognat ont été identifiés chez la levure par la suite <sup>247</sup>, ce qui confirme nos résultats.

L'anticodon de l'ensemble de ces ARNt s'apparie systématiquement sur au moins deux des trois nucléotides du codon stop, ce qui suppose que le ribosome ne tolérerait pas deux appariements non conventionnels. De plus, l'incorporation d'arginine sur UGA, de glutamine sur UAG et UAA confirme la possibilité d'un appariement U-G en première position. La cystéine et le tryptophane étant incorporés sur le codon stop UGA, cela implique que des appariements G-A, G-G et C-A sont acceptés par le ribosome lorsqu'ils sont en troisième position sur la base wobble. Bien qu'il soit encore actuellement impossible de prédire théoriquement l'acide aminé qui sera incorporé dans le peptide synthétisé lors de la translecture, ces travaux ont montré que l'identité du codon stop affecte non seulement la nature du résidu incorporé par translecture, mais aussi sa fréquence d'incorporation. A l'inverse, il a été suggéré que le contexte du codon stop n'a aucune incidence sur l'identité ou la proportion des acides aminés incorporés au cours du processus de translecture <sup>245</sup>.

#### *b. Le contexte nucléotidique*

Le contexte nucléotidique entourant le codon stop constitue un des éléments ayant le plus impact sur l'efficacité de translecture, mais aussi celui qui est le moins bien compris.

#### **✱ Le contexte en 3'**

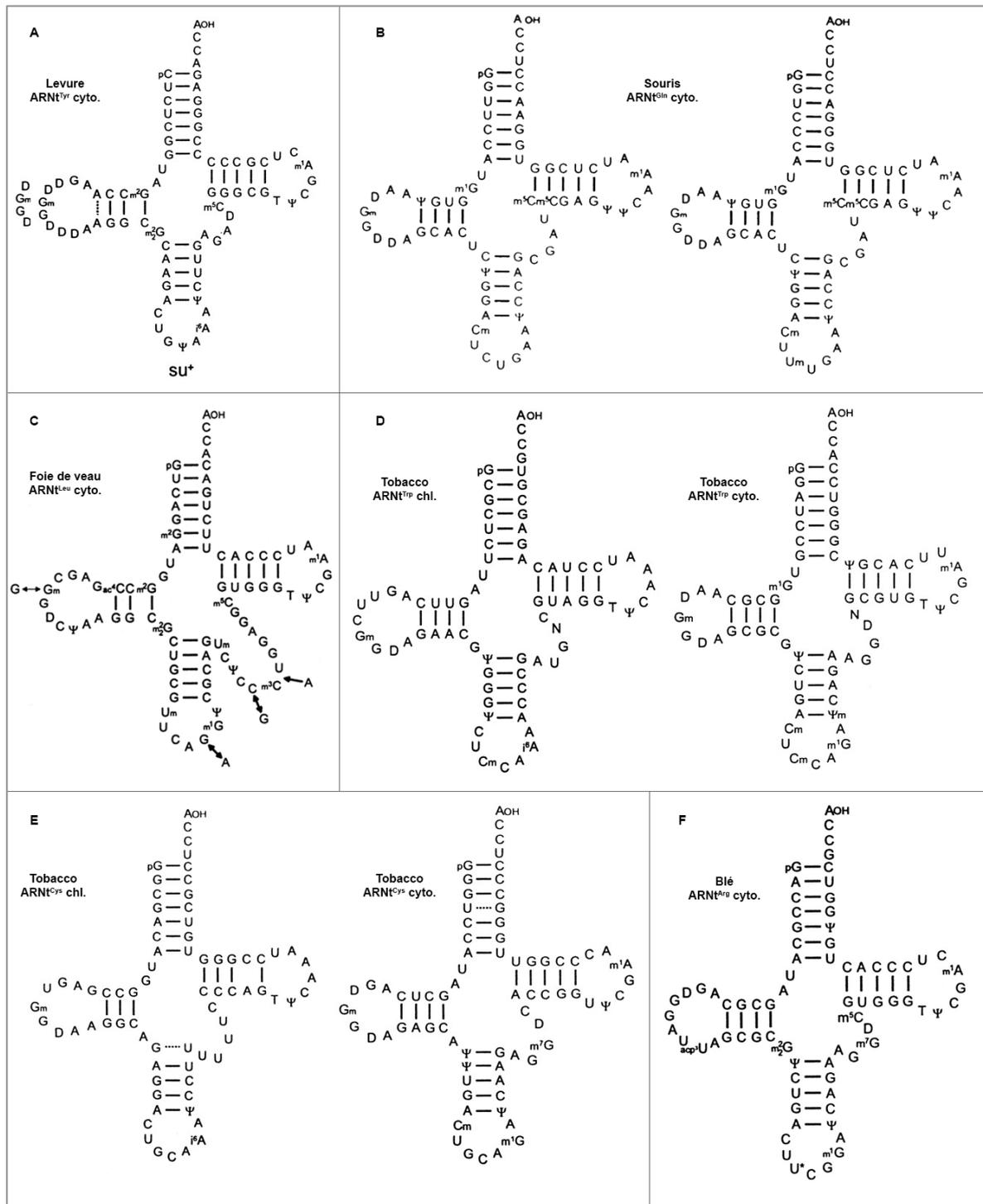
Chez les bactéries et les eucaryotes, la base qui suit immédiatement le codon stop (position +4) exerce la plus forte influence sur l'efficacité de translecture. Ces résultats ont conduit à l'hypothèse que le signal de terminaison de la traduction serait constitué d'une séquence de quatre nucléotides et non uniquement par le codon stop lui-même <sup>248</sup>. Cette hypothèse a été très récemment vérifiée chez les eucaryotes avec l'observation en cryo-EM du codon stop et de la position +4 dans le site A du ribosome <sup>76</sup>. Les auteurs de ces travaux ont également montré que la présence d'une purine en +4 stabiliserait le complexe de terminaison. Toutefois, l'influence de la nature du nucléotide en position +4 sur la translecture est encore controversée <sup>246</sup> (Tableau 1). S'il semble admis qu'une cytosine en position +4 après les codons stop UGA et UAA induise le plus fort taux de translecture, aucun consensus n'a été trouvé pour UAG.

Néanmoins, ces résultats ne permettent pas de prédire l'efficacité de translecture d'une mutation non-sens. Deux études récentes de l'influence des contextes sur l'efficacité de terminaison chez la levure et chez les mammifères ont révélé que l'influence du contexte ne se limitait pas au résidu +4, mais s'étendait jusqu'à la position +9<sup>225,240</sup>. Chez la levure, la séquence 3' la plus favorable à la translecture serait CARNBA (où R=A/C, N=A/T/G/C, B=C/G/T). Une autre étude, réalisée sur des cellules de mammifères, a également déterminé que la présence d'un uracile en position +5 était corrélée à un taux plus élevé de translecture<sup>249</sup>. Cela suggère que, mis à part la présence d'un C en position +4 qui semble être un prérequis universel pour une translecture efficace, l'influence du contexte nucléotidique peut être variable selon les organismes.

### ★ Le contexte en 5'

La conservation au cours de l'évolution de la séquence en 5' du codon stop chez *E. coli* et chez l'humain a conduit au postulat que cette séquence en amont jouerait également un rôle dans l'efficacité de la terminaison de la traduction<sup>250</sup>. Il a en effet été montré que les positions -1 et -2 peuvent moduler le niveau de translecture chez les bactéries et la levure. La présence d'une adénine sur ces deux positions stimulerait la translecture du codon UAG et probablement des autres stop<sup>228,251</sup>. Chez les mammifères, le contexte en 5' est également important. Il a également été montré qu'une adénine en position -1 induisait une forte translecture, alors qu'à l'inverse, la présence d'un U était associée au plus faible taux de translecture<sup>227</sup>. Dans une seconde étude chez les mammifères, une purine en position -1 favoriserait la translecture, mais aucun consensus n'a été trouvé en position -2<sup>249</sup>. L'influence du contexte en 5' semble donc être beaucoup plus subtile que le contexte en 3'<sup>246</sup>.

La façon dont le contexte nucléotidique influence la terminaison de la traduction est encore mal comprise. Plusieurs hypothèses ont été envisagées afin d'expliquer l'influence des nucléotides en 5' du codon stop. Chez la levure, le contexte 5' interagirait directement avec l'ARNt au site P du ribosome. Les acides-aminés de la chaîne naissante pourraient également interagir avec les facteurs de terminaison<sup>252</sup>. Cette hypothèse a toutefois été contredite par Tork et ses collaborateurs, qui ont suggéré que l'adénine en position -1 ou -2 pourrait induire de la translecture en modifiant la structure de l'ARNm au niveau du site P, ce qui influencerait l'incorporation d'un ARNt au site A<sup>251</sup>.



**Figure 50 : Structure secondaire des ARNt supresseurs**

Séquence nucléotidique des ARNt supresseurs tyrosine (A), glutamine (B), leucine (C), tryptophane (D), cystéine (E) et arginine (F). **A** - L'ARNt<sup>Tyr</sup> de *S. cerevisiae* porte l'anticodon est GΨA et décode le codon UAG. **B** - Les deux isoaccepteurs ARNt<sup>Gln</sup> proviennent du foie de souris. La forme ayant la modification UmUG au niveau de l'anticodon est un fort supresseur des codons UAG et UAA. **C** - Les deux isoaccepteurs ARNt<sup>Leu</sup> sont issus de foie de veau et portent l'anticodon CAG ou CAA et décodent le codon UAG. **D** - Les deux isoaccepteurs ARNt<sup>Trp</sup> issus de *N. rustica*, portent l'anticodon CmCA et induisent la translecture du codon UGA. L'ARNt chloroplastique (chl.) serait plus efficace que la forme cytosolique (cyto.), ce qui s'expliquerai par la présence de la modification i<sup>6</sup>A portée uniquement par l'ARNt chloroplastique. **E** - Les ARNt<sup>Cys</sup> cytoplasmique et chloroplastique de *N. rustica* portent l'anticodon GCA et peuvent décoder le codon UGA. L'ARNt chloroplastique serait également plus efficace que celle cytoplasmique. **F** - L'ARNt<sup>Arg</sup> provenant du blé porte l'anticodon U\*CG et plus rarement ICG et décode le codon UGA. *Figure adaptée de Beier and Grimm, 2001*<sup>253</sup>

c. *Les structures secondaires*

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

B. **Les facteurs agissant en trans**

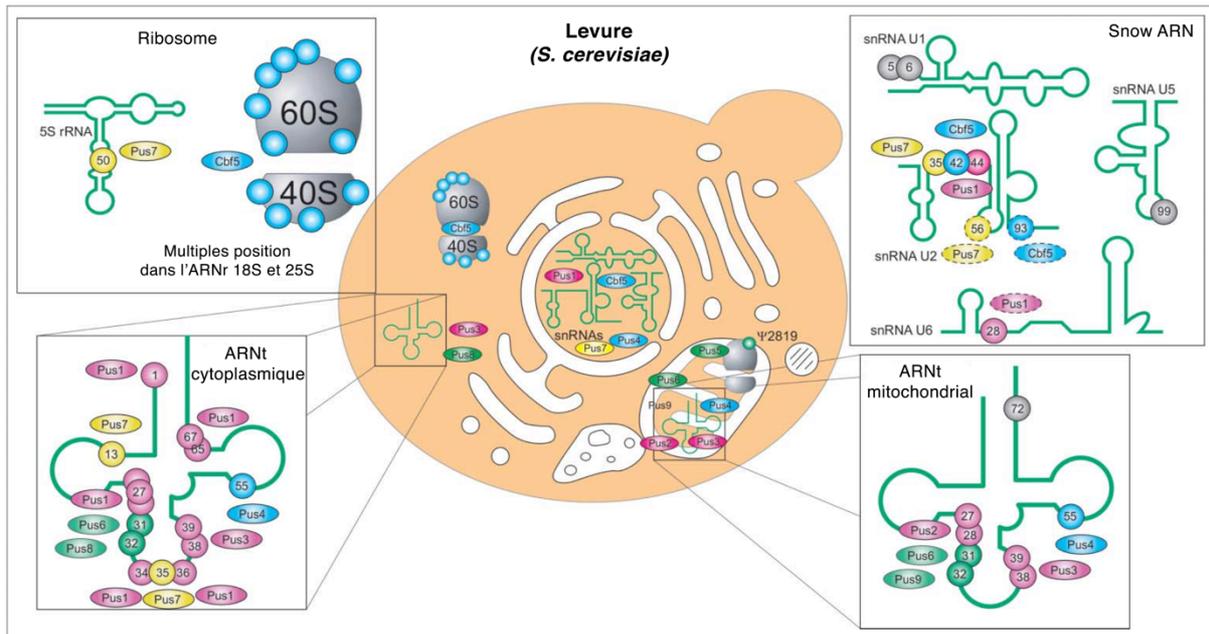
a. *Les ARNt supprimeurs*

En 1975, il a été rapporté que certains ARNm eucaryotes possédaient un codon stop traduit et codaient une protéine ayant une fonction cellulaire <sup>255</sup>. Les ARNt supprimeurs sont définis comme des ARNt capables de décoder des codons stop au cours de la traduction. Ces ARNt supprimeurs naturels sont des ARNt proche-cognat qui ne sont pas mutés. A contrario, les ARNt supprimeurs issus de gènes d'ARNt normaux qui subissent une modification dans leur anticodon, ce qui leur permet de lire un codon stop <sup>256</sup>.

★ **Supprimeurs des codons UAG/UAA**

Le premier exemple de suppression naturelle de codon stop chez les eucaryotes a été isolé chez la Drosophile où le codon UAG est décodé par l'ARNt<sup>Tyr</sup> portant l'anticodon GΨA (Figure 50A). La pseudouridylation permet de stabiliser l'interaction codon-anticodon et de décoder les codons stop UAG et UGA. L'ARNt<sup>Tyr</sup> modifié a aussi été retrouvé chez les plantes où le codon UAG est efficacement décodé (10-30%) <sup>257</sup>.

La seconde classe d'ARNt supprimeurs des codons UAG/UAA est celle des ARNt<sup>Gln</sup> (Figure 50B). Deux isoaccepteurs avec un anticodon CUG ou U\*UG existeraient chez tous les procaryotes et eucaryotes <sup>258</sup>. Chez la souris, la forme minoritaire de l'ARNt<sup>Gln</sup> portant l'anticodon UmUG serait capable de stimuler la translecture sur le codon UAG <sup>259</sup> alors que chez la levure, les deux formes de l'ARNt<sup>Gln</sup> ont cette capacité <sup>245</sup>. L'interaction de ces deux ARNt isoaccepteurs avec les codons stop UAG et/ou UAA nécessitent un appariement non conventionnel G-U sur la première position, dont on sait désormais que ce dernier est possible grâce à une forme tautomère rare des bases <sup>46,204</sup>.



**Figure 51 : Répartition des pseudouridines et pseudouridines synthases chez la levure**

La localisation cellulaire des enzymes et substrats ainsi que la position du substrat sont donnés pour les membres de la famille TruA : Pus1p, Pus2p et Pus3p (rose), les membres de la famille RluA : Pus5p, Pus6p, Pus8p96 et Pus9p96 (vert), l'homologue de TruD : Pus7p (jaune) et l'homologue de TruB : Pus4p (bleu), ainsi que les homologues des ARN-guides TruB : Cbf5 (bleu) pour le snARN (small nuclear RNA) U2 et snARN U5. Les sites de modification sans activité enzymatique connus sont indiqués en gris. La grande sous-unité mitochondriale ne contient qu'une seule modification à la position 2819 réalisée par Pus5. Notez que pour plus de clarté, seules quelques modifications du ribosome sont représentées alors qu'au moins 44 réalisées par Cbf5 sont connues à ce jour. Les snARN U2 en position 56 et 93 et snARN U6 en position 28 ont un contour en pointillés en raison de leur inductibilité. Pus7p est représentée dans le cytoplasme avec un contour en pointillé car sa localisation varie entre le noyau et le cytoplasme suite à un choc thermique. *Figure adaptée de Spenkuch et al., 2015* <sup>260</sup>

L'ARNt<sup>Leu</sup> a aussi deux formes isoacceptrices permettant de décoder le codon stop UAG<sup>261</sup>. Ces deux formes portent l'anticodon CAA et CAG (Figure 50C). La reconnaissance du codon stop nécessite un appariement non conventionnel sur la première (G-U) et seconde position<sup>262</sup>.

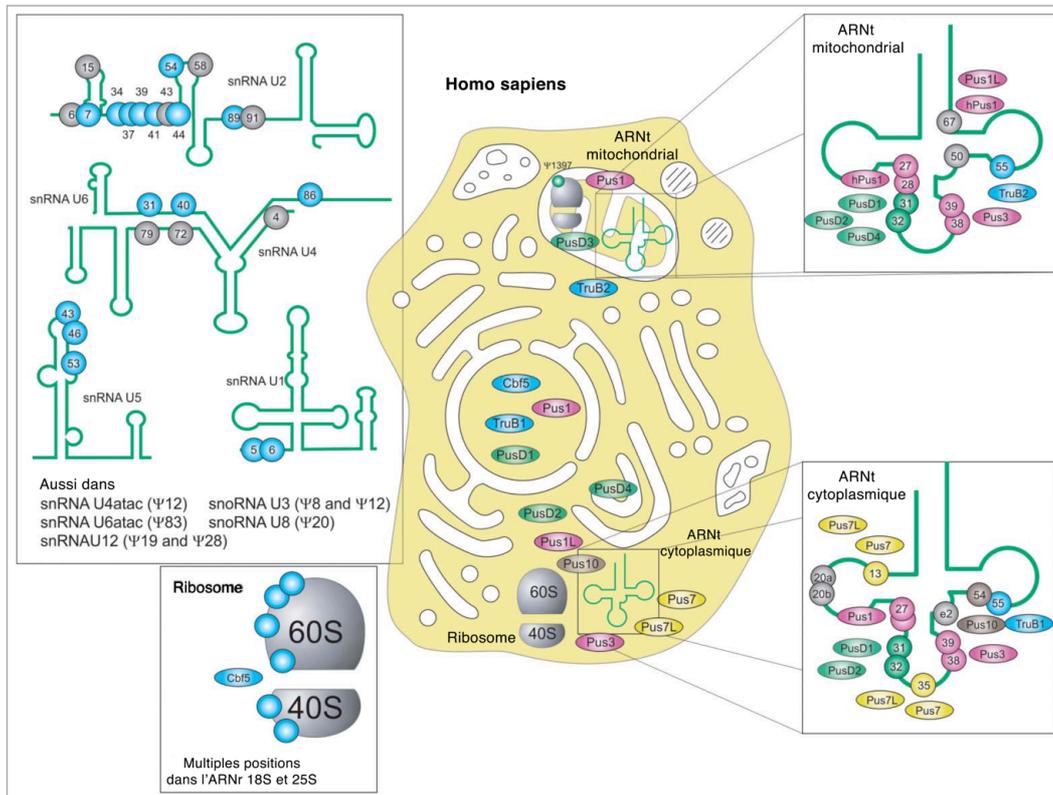
### ★ Suppresseurs du codon UGA

Deux ARNt supprimeurs naturels ont été identifiés chez les plantes, permettant de lire le codon UGA<sup>263</sup>. Ces deux formes correspondent aux ARNt<sup>Trp</sup> chloroplastes et cytosoliques dont l'anticodon est CmCA (Figure 50D). Étrangement, l'ARNt<sup>Trp</sup> chloroplaste induit davantage de translecture. Il a été proposé que les modifications i<sup>6</sup>A et ms<sup>2</sup>i<sup>6</sup>A portées exclusivement par l'ARNt<sup>Trp</sup> chloroplaste stabiliseraient les interactions codon-anticodon en augmentant l'effet d'empilement de l'anticodon sur les nucléotides voisin, ce qui favoriserait les interactions non canoniques à la troisième position (Wobble)<sup>264,265</sup>. Ces deux formes d'ARNt<sup>Trp</sup> ont été retrouvées chez le lapin (lysats de réticulocytes de lapin - RRL), mais la séquence de l'ARNt<sup>Trp</sup> n'a pas été déterminée et il n'est pas exclu que l'ARNt supprimeur soit en réalité l'ARNt<sup>Trp</sup> mitochondrial avec un anticodon U\*CA. Ce dernier étant le seul ARNt<sup>Trp</sup> encodé par le génome mitochondrial du lapin<sup>242,258</sup>.

De nombreuses études ont rapporté que les deux isoaccepteurs ARNt<sup>Cys</sup> portant l'anticodon GCA chez les plantes peuvent décoder le codon UGA (Figure 50E). Comme les ARNt<sup>Trp</sup>, la forme chloroplaste serait plus efficace que celle cytosolique<sup>243</sup>. Dans les cellules animales, il n'y a aucune donnée disponible mettant en évidence une possible activité de translecture de l'ARNt<sup>Cys</sup>, toutefois, la forte homologie de séquences des gènes codant l'ARNt<sup>Cys</sup> avec les plantes suggère que les ARNt<sup>Cys</sup> des animaux sont également des ARNt supprimeurs<sup>253</sup>.

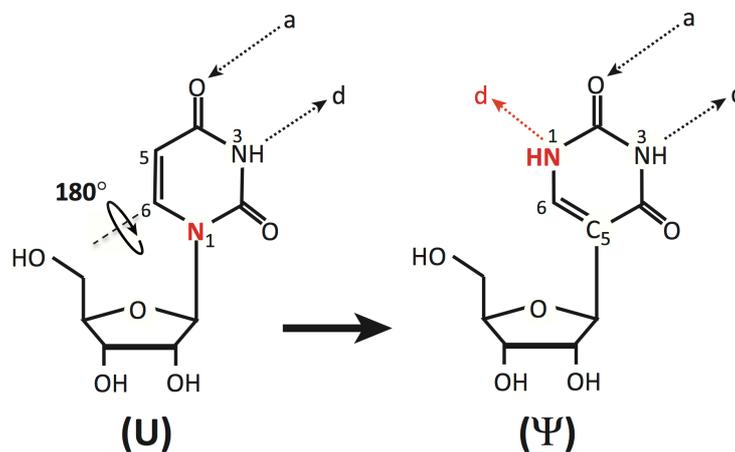
La troisième classe d'ARNt supprimeurs du codon UGA a été exclusivement caractérisée chez les plantes. Les six codons arginine CGN et AGR sont lus par quatre isoaccepteurs chez *E. Coli* et *S. Cerevisiae*, et cinq dans le cytosol des plantes (Figure 50F). Cet ARNt<sup>Arg</sup>, dont l'anticodon est U\*CG, est absent de la levure et d'*E. Coli* et décode particulièrement efficacement le codon UGA. La translecture du codon stop UGA a aussi été détectée dans les cellules de poulet, humaines et d'insecte<sup>266</sup>.

Ces données contredisent le dogme de non-ambiguïté du code génétique et montrent qu'un codon stop peut avoir plusieurs significations. Le codon UGA peut médier la fin de la synthèse d'un polypeptide ou permettre l'incorporation du tryptophane, de l'arginine, de la cystéine ou d'une sélénocystéine.



**Figure 52 : Répartition des pseudouridines et pseudouridines synthases chez l'*Homo sapiens***

L'emplacement cellulaire des substrats et des position des substrat sont donnés pour les membres de la famille TruA Pus1, Pus1L, Pus3 (rose), les membres de la famille TruB : TruB1, TruB2 et CBF5 (bleu), les membres de la famille TruD : Pus7 et Pus7L, membres de la famille RluA : PusD1, PusD3 et PusD4 (vert) et Pus10 (marron). En plus de l'ARNt et de l'ARNr, les snARN et snoARN sont modifiés. D'après les connaissances actuelles, les position des pseudouridylations des snARN U2, U4 et U6 sont exclusivement formés par H/ACA Box RNPs. Les positions représentés en bleu ont des ARN guides connus ou putatifs, tandis que les positions en gris n'en ont pas. Les ARN télomérases ne sont pas représentés ici. *Figure adaptée de Spenkuch et al., 2015*<sup>260</sup>



**Figure 53 : Schéma représentant l'isomérisation de l'uridine (U) en pseudouridine (Ψ)**

Représentation schématique des structures de l'uridine et de la pseudouridine. Ψ dérive de l'uridine par une réaction d'isomérisation dans lequel la base est tournée de 180° le long de l'axe N3-C6 et forme une liaison en C5-C10. L'atome d'azote en première position de l'uridine et le donneur de liaison hydrogène supplémentaire dans Ψ sont indiqués (en rouge). Les accepteurs de liaison hydrogène (a) et les donneurs de liaison hydrogène (d) sont également indiqués. *Figure issue de Ge and Yu, 2013*<sup>267</sup>

De même, le codon de terminaison UAG peut provoquer l'incorporation de la tyrosine, de la glutamine ou de la leucine et de pyrrolysine. La lecture erronée des codons de terminaison est réalisée par une variété d'ARNt supprimeurs naturels. Les ARNt supprimeurs de codon stop (à l'exception de l'ARNt sélénocystéine) étant des ARNt proche-cognat, leur reconnaissance des codons stop nécessite des appariements de bases non canoniques <sup>253</sup>.

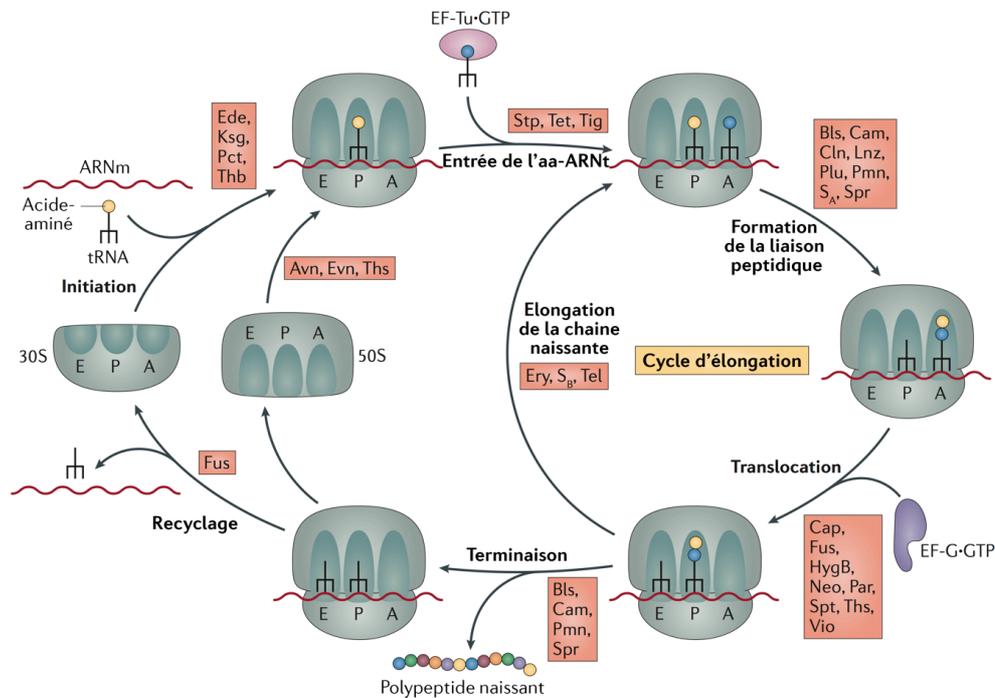
#### *b. Les pseudouridines*

Découvert en 1951, il a été identifié à ce jour plus de 150 modifications nucléotidiques dont le rôle est d'affiner la conformation, la structure et la fonction de l'ARN (Figures 51 et 52). Malgré les nombreux travaux publiés ces 15 dernières années, le rôle de la modification de nombreux nucléotides et le mécanisme impliqué dans la formation de la pseudouridylation sont toujours très largement incompris <sup>260</sup>.

La pseudouridine est un isomère de l'uridine (U) (Figure 53). Initialement, la liaison N1-C1' liant la base uracile au sucre est cassée. L'uracile libéré est tourné de 180° sur l'axe N3-C6 et établit une nouvelle liaison avec le sucre (C5-C1') <sup>267</sup>. Cette modification crée un donneur de liaison hydrogène supplémentaire dotant  $\Psi$  de propriétés chimiques différentes de l'uridine et des autres nucléotides connus. En effet, si l'uridine et  $\Psi$  peuvent tous deux former un appariement de type Watson-Crick avec l'adénosine, l'appariement de  $\Psi$  avec les 4 bases est bien plus stable <sup>117,268</sup>. Cette modification est réalisée par six familles de pseudouridine synthases : TruA, TruB, TruD, RsuA, RluA et Pus10p, dont le repliement et le mécanisme seraient très proches <sup>269,270</sup>. Lors de la dégradation des ARN, les pseudouridines peuvent être dégradés chez les eubactéries <sup>271</sup> ou inactivés par déphosphorylation du carbone 5' du ribose chez les mammifères <sup>272</sup>.

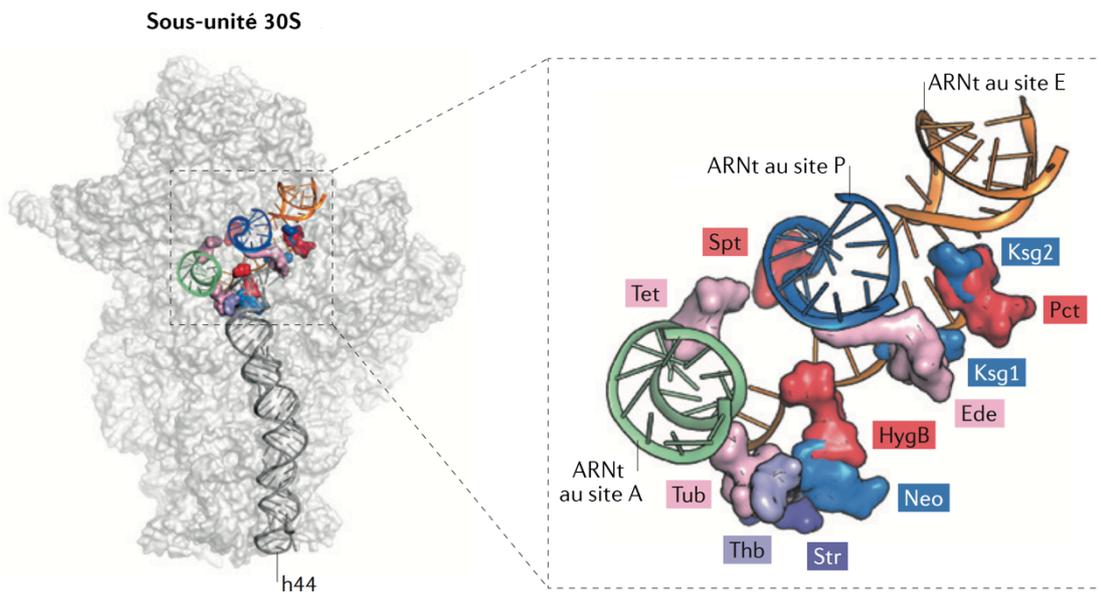
Les modifications de  $\Psi$  les plus conservées stabilisent la structure tertiaire de l'ARNt, qui peut être soit en position 32, 39 ou 55. Les changements de conformation provoqués par la présence de  $\Psi$  en position 38 et 39 modifient la reconnaissance de l'anticodon, diminuant la fidélité de la traduction et permettant au ribosome de changer de phase de lecture chez la levure <sup>273-275</sup>. Plusieurs ARNt sont également modifiés au niveau des positions 34, 35 et 36. La pseudouridylation de la position 35 conférerait une stabilisation plus importante de l'anticodon en renforçant la liaison U33- $\Psi$ 35. Quant aux modifications 34 et 36, leur rôle est encore inconnu.

Les modifications des bases de l'ARNr sont très conservées et sont associées à des domaines fonctionnels importants du ribosome. Une inactivation globale de la pseudouridylation de l'ARNr affecte les interactions ribosome-ligands ce qui a pour conséquence d'augmenter la fidélité du ribosome chez les levures et les mammifères.



**Figure 54 : Cible des antibiotiques au cours de la traduction bactérienne**

Schéma représentant les étapes de la traduction ciblées par les différents antibiotiques bactériens : édéine (Ede), kasugamycine (Ksg), pactamycine (Pct), thermorubine (Thb), avilamycine (Avn), évernimicine (Evn), thiostreptone (Ths), streptomycine (Stp), tétracyclines (Tet), tigécycline (Tig), blasticidine S (Bls), chloramphénicol (Cam), clindamycine (Cln), linézolide (Lnz), pleuromutilines (Plu), puromycine (Pmn), streptogramine A (SA), sparsomycine (Spr), capréomycine<sup>255</sup>, viomycine (Vio), hygromycine B (hygB), néomycine<sup>276</sup>, paromomycine (Par), l'acide fusidique (Fus), spectinomycine (Spt), érythromycine (Ery) streptogramine B (SB) et téli-thromycine (Tel). *Figure adaptée de Wilson et al., 2014<sup>277</sup>*



**Figure 55 : Site de fixation des antibiotiques sur la 30S**

Vue d'ensemble et élargie des sites de liaison des antibiotiques le long du canal de fixation de l'ARNm de la sous-unité 30S. Les sites de fixation des ARNt A, P et E sont respectivement représentés en vert, bleu et orange. Antibiotiques présents : tétracycline (Tet; PDB : 4G5K-N), spectinomycine (Spt; PDB : 2QOU-X), kasugamycine (Ksg1 and Ksg2; PDB : 2HHH), pactamycine (Pct; PDB : 1HNX), édéine (Ede; PDB : 1I95)14, hygromycine B (HygB; PDB : 3DF1-4), néomycine (Neo; PDB : 4GAQ/R/S/U), streptomycine (Str; PDB : 1FJG), thermorubine (Thb; PDB : 3UXQ-T) et tuberactinomycine (Tub; PDB : 3KNH-K). *Figure adaptée de Wilson et al., 2014<sup>277</sup>*

En absence de ces modifications, les évènements de recodages tels que le frameshift et la translecture sont fortement diminués. L'initiation sur les IRES et l'incorporation des ARNt seraient également affectées négativement en absence de la pseudouridylation des bases de l'ARNr<sup>278</sup>.

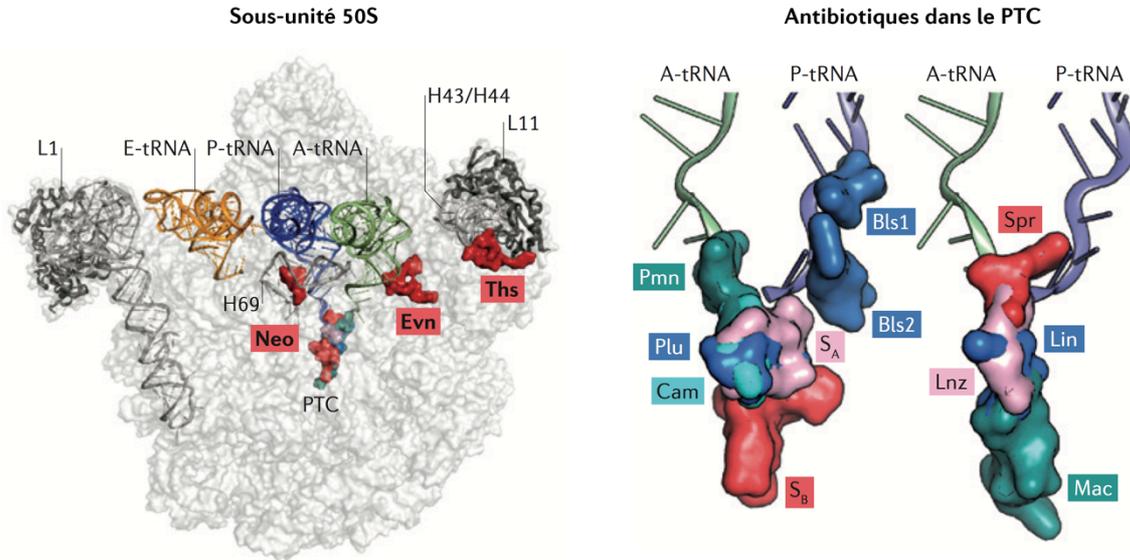
Les bases de l'ARNm sont aussi modifiées chez les eucaryotes et les archées<sup>267</sup>. Il a été montré que la modification de la première base des trois codons stop favorisait l'incorporation d'un ARNt proche-cognat ou non-cognat en stabilisant l'interaction codon-anticodon. Chez la levure, ΨAG serait décodé majoritairement par une sérine puis une thréonine, ΨUA par les ARNt sérine et thréonine de manière équivalente, et ΨGA par la tyrosine puis la phénylalanine<sup>279</sup>. Ce décodage sélectif des codons stop pseudouridynilés a été rationalisé récemment par le groupe de Pan<sup>280</sup>. Étant donné le grand nombre de codons sens contenant une uridine (34 des 61 codons sens contiennent une ou plusieurs uridines), la pseudouridylation de l'ARNm ciblé laisse présager d'une expansion du code génétique. La présence de Ψ dans l'ARNm augmenterait l'efficacité de traduction, mais le mécanisme impliqué doit encore être élucidé<sup>281</sup>.

### c. Les antibiotiques

La traduction bactérienne est la cible de nombreux antibiotiques (Figure 54). La quasi-totalité d'entre eux se fixent sur le ribosome et plus particulièrement sur les ARNr qui jouent un rôle majeur dans le décodage. Ces molécules constituent des outils permettant d'étudier et de mieux comprendre les différentes étapes de la traduction. Malgré la taille importante du ribosome, peu de sites sont ciblés par l'arsenal actuel d'antibiotiques<sup>277</sup> :

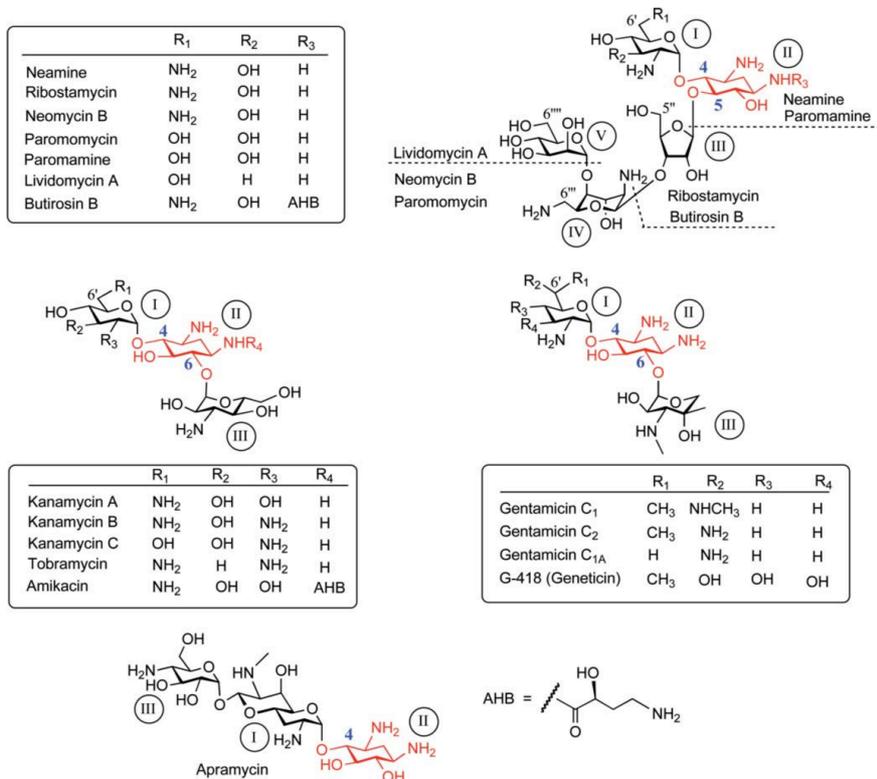
Sur la sous-unité 30S, les sites de liaison des antibiotiques sont regroupés le long du trajet de l'ARNm et de l'ARNt (Figure 55). Les antibiotiques comme l'édéine et la kasugamycine inhibent l'étape d'initiation de la traduction en empêchant une interaction stable entre l'ARNt initiateur et le codon de départ. La majorité des autres antibiotiques qui cible la 30S inhibent l'élongation en interférant soit avec l'entrée d'un ARNt au site A (tétracyclines et streptomycines), soit au cours de la translocation du ribosome (aminoglycosides, hygromycine B, néomycine, pactamycine et la spectinomycine).

Sur la sous-unité 50S, la plupart des sites de fixation des antibiotiques sont regroupés au niveau ou à proximité du centre peptidyl-transférase<sup>22</sup> (Figure 56). Parmi les exceptions, il y a les ortosomycines qui interagissent avec les hélices H89 et H91 de l'ARNr 23s et empêchent la liaison du facteur IF2 et les thiopeptides qui interagissent avec l'hélice 43 et 44 de l'ARNr 23S et bloquent la fixation d'EF-G, EF-Tu et IF2.



**Figure 56 : Site de fixation des antibiotiques sur la 50S**

Vue d'ensemble et rapprochée des sites de liaison à la sous-unité 50S et sur le centre peptidyl transférase<sup>22</sup>. Les sites de fixation des ARNt A, P et E sont respectivement représentés en vert, bleu et orange. Antibiotiques représentés : néomycine (Neo; PDB : 4GAQ/R/S/U), evermicine (Evn) et thioestreptone (Ths; PDB : 3CF5), blasticidine S (Bls1 and Bls2; PDB : 1K88), sparsomycine (Spr; PDB : 1M90), lincomycine (Lin; PDB : 3OFX/Y/Z/O), linezolid (Lnz; PDB : 3DLL), macrolides (Mac; PDB : 1K9M), puromycine (Pmn; PDB : 1M90), pleuromutilins (Plu; PDB:1XBP), chloramphenicol (Cam; PDB:3OFA-D), streptogramines A et B (SA et SB; PDB : 1SM1). *Figure adaptée de Wilson et al., 2014*<sup>277</sup>



**Figure 57 : Structure des aminoglycosides**

Structures des dérivés 2-désoxystreptamine (2-DOS) 4-, 4,5- et 4,6-disubstitués qui ciblent le site A procaryote. La structure 2-DOS est indiquée en rouge et la numérotation de carbones en bleu. *Figure issue de Hainrichson et al., 2007*<sup>282</sup>

Pour être utilisables chez l'homme, les antibiotiques doivent cibler le ribosome bactérien sans trop perturber la traduction eucaryote. Ce sont ces quelques différences au niveau des ARNr et parfois des protéines ribosomiques qui confèrent au ribosome eucaryote cette résistance aux antibiotiques. Certaines de ces molécules ont tout de même la capacité de se fixer sur le ribosome eucaryote, mais avec une affinité moindre. C'est notamment le cas des antibiotiques de la famille des aminoglycosides.

#### d. Les aminoglycosides

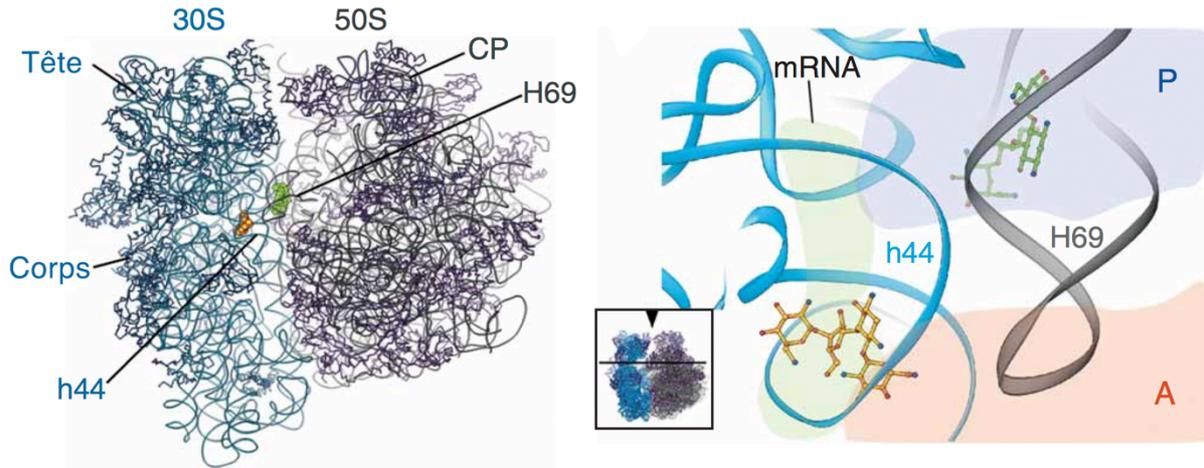
##### ✧ Structure et diversité

En 1944, les antibiotiques de la famille des aminoglycosides ont été isolés pour la première fois à partir de bactéries du sol <sup>283</sup>. Suite à la découverte de la streptomycine, d'autres molécules produites naturellement par les *Streptomyces* (néomycine, kanamycine) et par des actinobactéries de type *Micromonospora* (gentamicine) ont été isolées. Depuis, plusieurs dérivés semi-synthétiques ont été synthétisés pour contrer les résistances bactériennes (amikacine, arbekacine).

D'un point de vue de leur structure, les aminoglycosides sont des molécules polaires et polycationiques. La majorité des aminoglycosides naturels se composent d'un noyau commun, appelé 2-déoxystreptamine (2-DOS) auquel sont attachés des sucres (groupements glucidiques) sur les positions 4, 5 et 6 (Figure 57). Selon la nomenclature actuelle, le cycle 2-DOS est numéroté cycle II et représente dans la plupart des structures le cycle central. Les groupements glucidiques sont liés à la position 4 du 2-DOS (cycle I) et à la position 5 ou 6 (cycle III). Deux autres cycles glucidiques peuvent s'ajouter et sont numérotés IV et V. Les dérivés 2-DOS 4,5- et 4,6-disubstitués sont les deux classes d'antibiotiques aminoglycosides les plus utilisés cliniquement <sup>282</sup>.

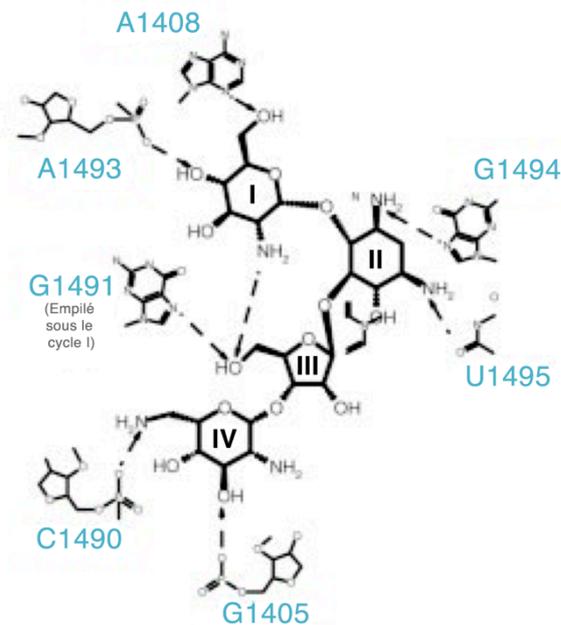
##### ✧ Mode d'action et effets

Les aminoglycosides présentent une activité contre de nombreuses bactéries Gram-négatives, certaines bactéries Gram-positives et mycobactéries non tuberculeuses. Cette activité est due à leur liaison au ribosome <sup>284</sup>. Grâce à cette interaction, les aminoglycosides réduisent la fidélité de la traduction entraînant une augmentation de l'incorporation d'aminoacyl-ARNt proche-cognat au site A du ribosome, ce qui conduit à une accumulation de protéines non fonctionnelles ou tronquées aboutissant à la mort cellulaire bactérienne.



**Figure 58 : Fixation des aminoglycosides dans le ribosome 70S**

A gauche, vue globale du ribosome et des deux sites de fixation des aminoglycosides en vert au niveau de la sous-unité 50S et en doré au niveau de la sous-unité 30S. L'ARNr 16S est en bleu clair, l'ARNr 23S en gris, l'ARNr 5S en violet, les protéines en bleu foncé (30S) et magenta (50S), CP : protubérance centrale. A droite, vue rapprochée des deux sites de fixation sur l'hélice 69 et l'hélice 44, avec la néomycine en exemple. La molécule se fixant sur h44 est en doré, celle se fixant sur H69 est en vert. L'ombre verte correspond à l'ARNm, la bleue au site P et l'orange au site A. *Figure adaptée de Borovinskaya et al., 2007*<sup>285</sup>



**Figure 59 : Sites d'interactions de la paromomycine dans le centre de décodage**

Digramme chimique de la paromomycine, antibiotique de la famille des aminoglycosides, comprenant le cycle 2-DOS (II) et 3 groupements glucidiques (I, III, IV). Ses interactions avec les résidus spécifiques de la sous-unité 30S sont indiquées. *Figure adaptée de Carter et al., 2000*<sup>286</sup>

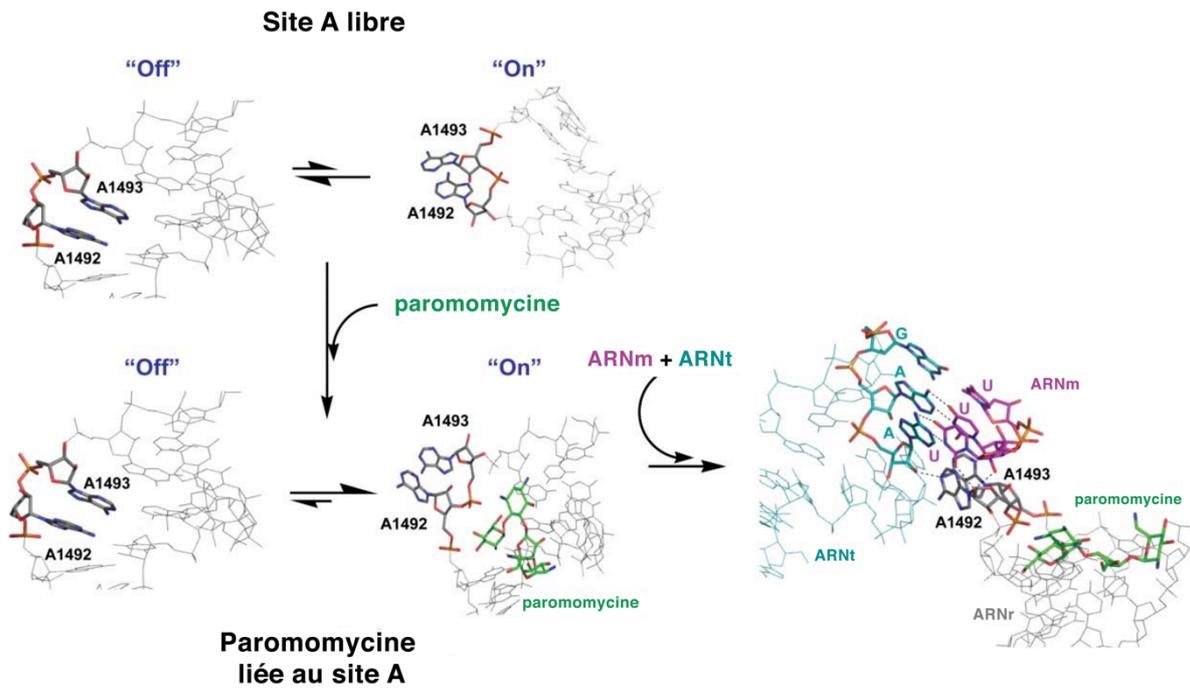
Toutefois, l'apramycine est le seul aminoglycoside dont le mécanisme d'action ne consiste pas à diminuer la fidélité de la traduction, mais à bloquer la translocation du ribosome, via des interactions avec la protéine ribosomale S12<sup>287</sup>.

Les aminoglycosides sont des molécules chargées positivement qui ont une forte affinité pour l'ARNr chargé négativement. Ils peuvent se fixer sur la grande sous-unité où ils impactent la translocation et le recyclage du ribosome et sur la petite sous-unité où ils influencent la fidélité du décodage<sup>285</sup> (Figure 58). Les aminoglycosides se fixent de manière spécifique au cœur du centre de décodage, en interagissant avec les résidus de l'hélice 44 sur l'ARNr 16S. Un des modèles les plus étudiés est celui de la fixation de la paromomycine sur le ribosome bactérien qui a permis d'obtenir des informations sur le processus de décodage<sup>286</sup>.

Le cycle I de la paromomycine se positionne dans l'hélice 44 par empilement avec la base G1491 et mime un nucléotide en créant une liaison hydrogène avec A1408. Une liaison hydrogène relie également le cycle I au résidu A1493, ce qui participerait au maintien du basculement des deux adénines (A1492 et 1493). Le cycle II (2-DOS) possède des liaisons étroites avec les bases et les phosphates des nucléotides. Le cycle III a de faibles contacts avec l'ARNr. Le cycle IV interagit avec les deux côtés de l'hélice 44<sup>288</sup> (Figure 59).

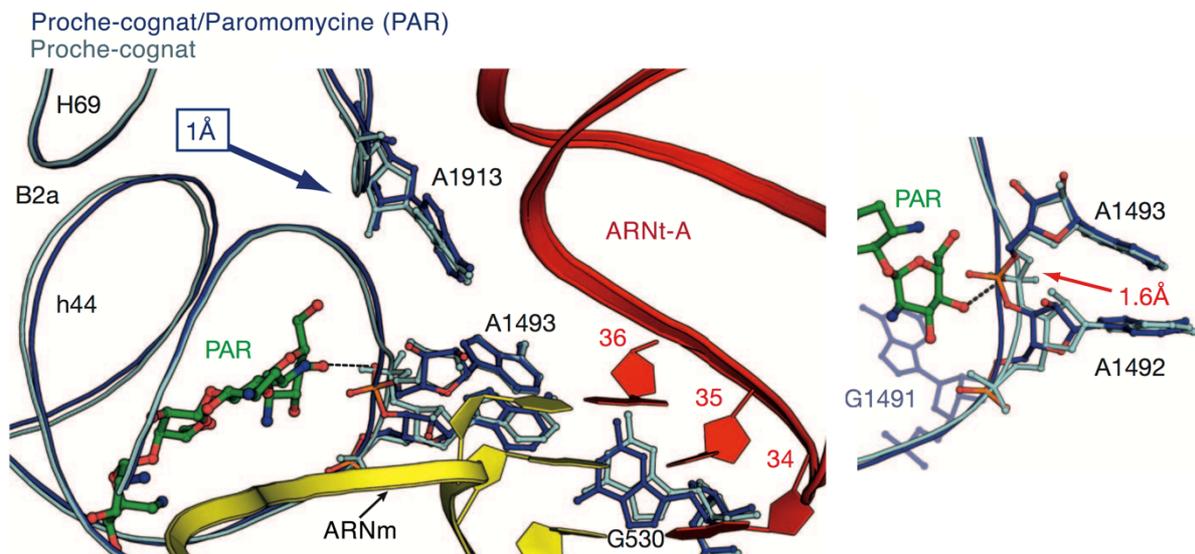
La fixation de la paromomycine dans le centre de décodage induit un réarrangement des adénines 1492 et 1493 qui basculent à l'extérieur de l'hélice. Dans cette orientation, qui est proche de celle observée lors de la fixation d'un ARNt proche-cognat, les deux adénines seraient en mesure d'interagir au niveau du petit sillon de l'hélice codon-anticodon<sup>286</sup> (Figure 60). Des analyses cinétiques ont aussi montré que la paromomycine réduisait la vitesse de dissociation des ARNt proche-cognat du site A, mais accélérât les deux réactions qui ont lieu lors de la sélection des ARNt, le déclenchement de l'activité GTPase et l'accommodation au site A<sup>289</sup>.

Récemment le rôle des bases A1492 et A1493 dans le décodage a été remis en question par Demeshkina et ses collaborateurs, l'impact de la paromomycine sur le site de décodage a également été réanalysé<sup>46</sup>. Leur structure du ribosome procaryote 70S avec un ARNt proche-cognat au site A en présence de paromomycine révèle un mouvement de l'hélice 69 de la grande sous-unité accompagné d'un réarrangement du pont B2a entre l'hélice 44 et 69 (Figure 61). Ces distorsions seraient probablement causés par la fixation de la paromomycine dont la liaison au ribosome provoque un fort décalage du groupement phosphate du résidu A1493. Ces changements locaux induisent le déplacement de l'hélice H69 via B2a vers l'ARNt qui améliore l'interaction de la surface de l'hélice H69 avec la tige-D de l'ARNt.



**Figure 60 : Fixation de la paromomycine de centre de décodage bactérien**

Au site de décodage, les bases A1492 et A1493 sont dans un équilibre conformationnel avec une prédominance de la conformation « off » dans laquelle elles sont positionnées vers l'intérieur de l'hélice. La fixation de la paromomycine fait basculer l'équilibre vers une conformation « on » dans laquelle les adénines sont orientées vers l'extérieur de l'hélice. Ce basculement a lieu même en absence d'ARNm et d'ARNt. Dans la conformation « on », les bases A1492 et A1493 sont capables d'effectuer des liaisons hydrogènes avec la mini-hélice codon-anticodon. La paromomycine est en vert, l'ARNm en magenta, l'ARNt en cyan, l'ARNr en gris. (PDB : 1IBL et 1J5E). *Figure adaptée de Hainrichson et al., 2008*<sup>282</sup>



**Figure 61 : Structure de la paromomycine dans le centre de décodage bactérien**

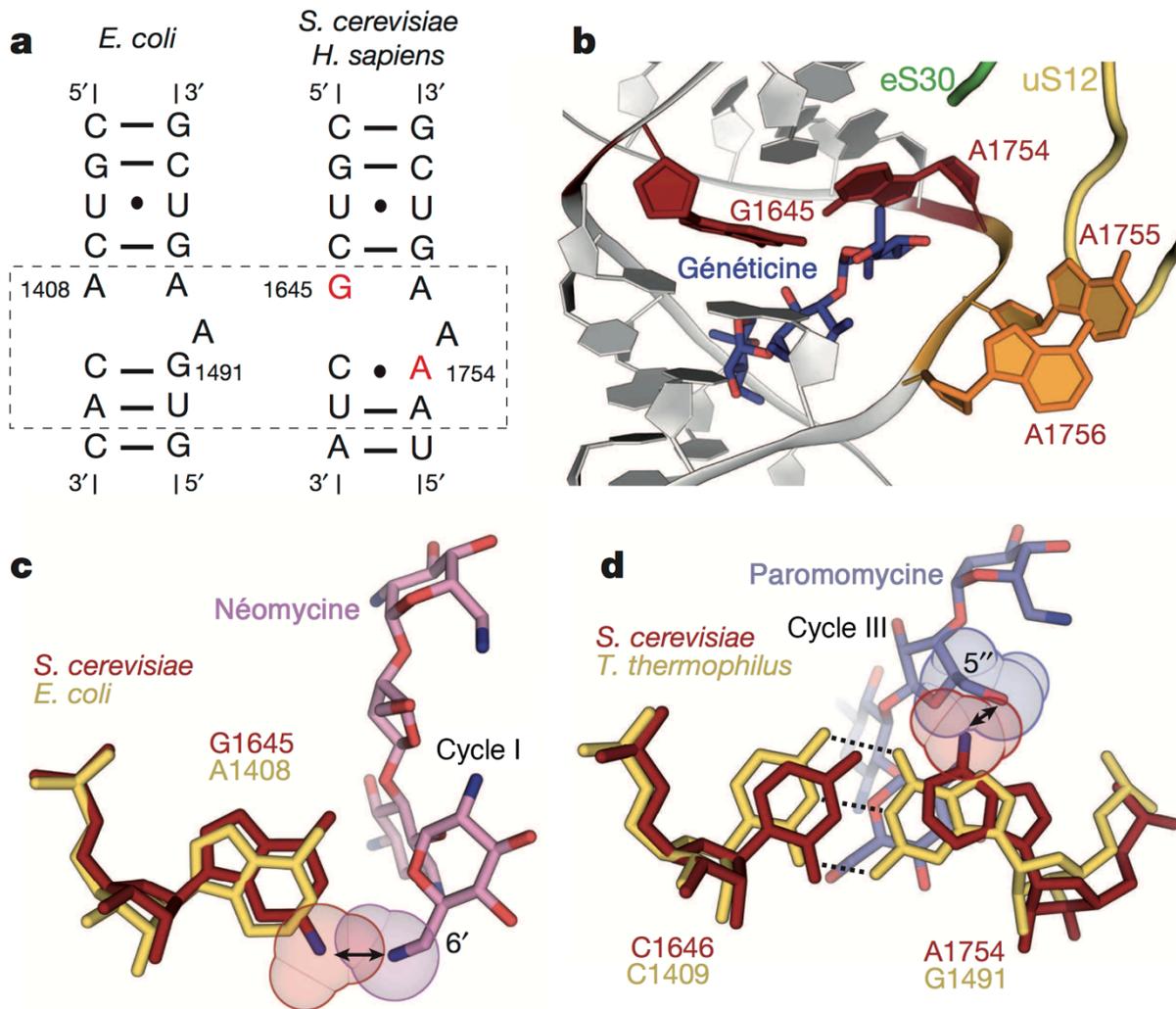
A gauche, réarrangements des hélices 44 et 69 de l'ARNr en présence d'un ARNt proche-cognat suite à la fixation de la paromomycine. L'ARNm est en jaune, l'ARNt au site A en rouge, la paromomycine est représentée en vert. A droite, changement de la position du phosphate de la base A1493 induit par la paromomycine. *Figure adaptée de Demeshkina et al., 2012*<sup>46</sup>

De plus, le déplacement du groupement phosphate de A1493 permet aussi de relâcher le site de décodage au niveau du codon au site A et modifie la déformation de l'hélice codon-anticodon formée avec l'ARNt proche-cognat. Le positionnement de la base A1493 limite également l'entrée des facteurs de terminaison.

### ★ Affinité pour le ribosome eucaryote

Bien que le centre de décodage du ribosome soit conservé entre les organismes procaryotes et eucaryotes, les nucléotides responsables de l'affinité et de la liaison des aminoglycosides à l'ARNr 16S procaryotes (A1408 et G1491) sont absents du centre de décodage du ribosome mammifère (18S; G1645 et A1754) (Figure 62a) <sup>290</sup>. De plus, la conformation du résidu G1645, qui est différente de celle de A1408 chez les bactéries, agit comme une barrière sur le ribosome eucaryote, en restreignant l'entrée des aminoglycosides ayant un substituant différent d'un 6'-hydroxyle sur le cycle I (Figure 62c) <sup>291</sup>. Pourtant, la paromomycine qui contient ce 6'-hydroxyle est faiblement active sur les ribosomes eucaryotes cytosoliques. Les auteurs ont également identifié une seconde barrière structurelle due au résidu A1754 qui est déplacé chez les eucaryotes en raison de son interaction non canonique avec C1646, ce qui restreint le logement d'un groupement 5''-hydroxyle du cycle III (Figure 62d).

La sélectivité des antibiotiques de la famille des aminoglycosides pour le ribosome procaryote est due à leur forte affinité pour le ribosome procaryote, ce qui n'est pas le cas pour le ribosome eucaryote. Toutefois, plusieurs molécules de cette famille sont capables de se lier suffisamment aux ribosomes eucaryotes pour perturber la fidélité de la traduction. Ainsi, la proportion d'ARNt proche-cognat incorporé est augmentée en présence de ces molécules <sup>292</sup>. Cette capacité a été retrouvée tout d'abord avec la streptomycine puis a été identifiée avec la gentamicine, l'amikacine, la paromomycine, la généticine (G418), la lividomycine et la tobramycine dans de nombreuses lignées cellulaires <sup>288</sup>. Récemment, la structure de plusieurs aminoglycosides dont la généticine a été obtenue dans le ribosome eucaryote <sup>291</sup> (Figure 62b). Comme attendu, la généticine se lie dans la poche des aminoglycosides (hélice 44 de l'ARNr18S) et induit le basculement des bases universellement conservées A1755 (A1492) et A1756 (A1493). L'utilisation de ces molécules dans le cadre des maladies génétiques provoquées par un codon stop prématuré est étudiée depuis plus d'une dizaine d'années et a fait l'objet de mon projet de thèse, elle sera développée dans le chapitre suivant : IV. Les maladies à codon stop.



**Figure 62 : Structure d'aminoglycosides dans le centre de décodage eucaryote**

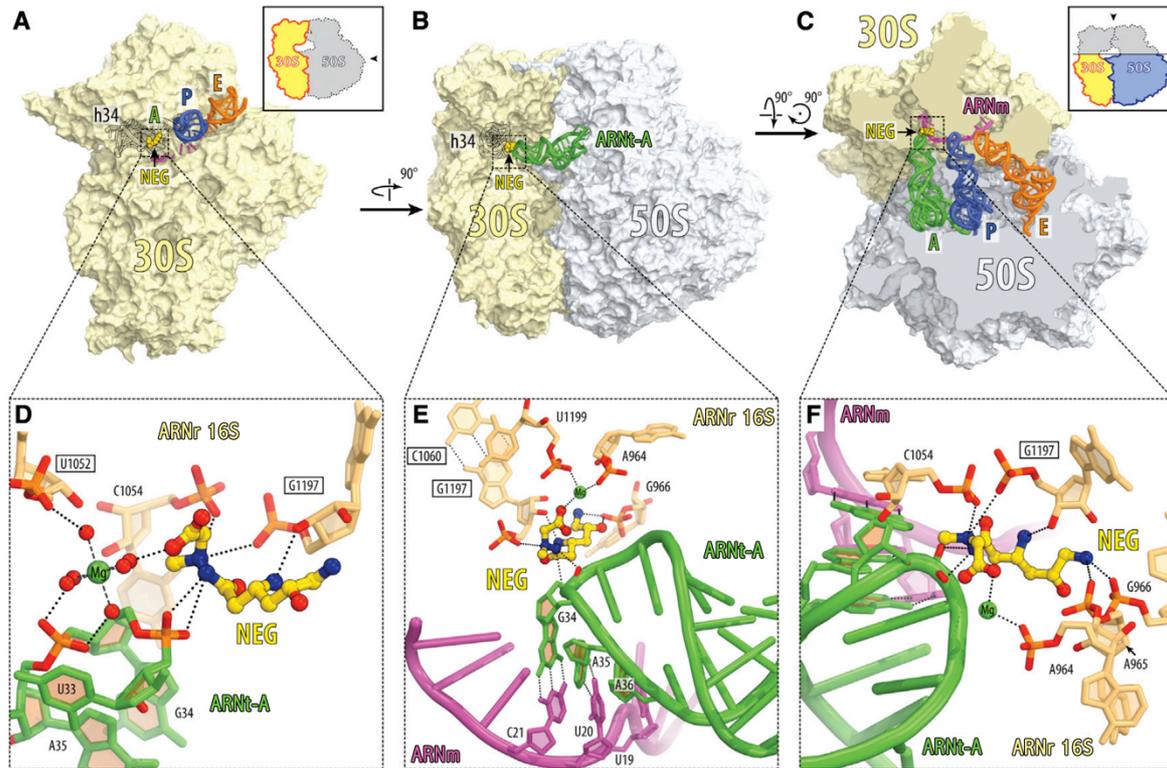
**a** - Le site de liaison des aminoglycosides est différent chez les eucaryotes. Structure secondaire de l'hélice 44 chez les bactéries (16S ARNr, à gauche) et chez la levure et l'humain (ARNr 18S, droite). **b** - La généticine (bleu) se lie à l'hélice 44 et induit le basculement des bases A1755 et A1756<sup>52</sup>. **c** - La conformation du résidu G1645 dans la sous-unité 18S de levure ARNr restreint la liaison des aminoglycosides portant un groupement amine en 6' du cycle I, comme illustré avec la néomycine (rose, PDB : 3QAN). **d** - La conformation du résidu A1754 dans la levure restreint la liaison des aminoglycosides portant un groupe 5'' hydroxyle sur le cycle III, comme le montre le schéma de la structure de la paromomycine (bleu, PDB 3UZ3). La structure de l'ARNm a été prise de PDB code adhésion 4KZZ. *Figure adaptée de Garreau de Loubresse et al., 2014*<sup>291</sup>

### e. La négamycine

La négamycine est un antibiotique dipeptide naturel non apparenté aux aminoglycosides, issu d'une souche de *Streptomyces* et isolée dans les années 1970. Cette molécule est efficace contre les bactéries Gram-négatives multirésistantes et contre certaines bactéries Gram-positives. La structure chimique de la négamycine, [2-[(3R, 5R)-3,6-diamino-5-hydroxyhexanol]-1-méthyl-hydrazino]acetic acid, et sa voie de synthèse ont été identifiées dans les années 1970 par plusieurs équipes <sup>293</sup>. De nombreuses études biochimiques ont suggéré que cet antibiotique inhibe la synthèse des protéines, mais son mode d'action est resté énigmatique, car diverses études *in vitro* ont suggéré qu'il pourrait inhiber l'initiation, le décodage et la terminaison de la traduction <sup>294</sup>. En 2003, l'équipe de Matsuda a montré que la négamycine avait également un effet inducteur de translecture sur le ribosome eucaryote <sup>295</sup>. Leurs données montrent également que son site de fixation se situait sur l'hélice 44 de la petite sous unité ribosomique. Cependant, les structures cristallographiques récentes de la négamycine dans le ribosome 70S de *Thermus thermophilus* et d'*E. Coli* contredisent ces données et révèlent qu'elle interagit avec l'hélice 34 de l'ARNr 16S <sup>296</sup> (Figure 63) :

Dans le ribosome, la négamycine est positionnée à proximité des résidus A1196 et C1054 de l'hélice 34. Le groupe carbonyle en C6 est directement lié à un ion magnésium, et le groupement amine en C4 se prolonge vers le squelette phosphate du résidu A1197. C2 pointe vers le squelette phosphate du résidu G966, et le groupe amino-terminal de la négamycine est positionné dans le solvant. Ce mode de liaison est supporté par les relations structure-activité des dérivés de la négamycine couplés à des essais de traduction *in vitro* <sup>294</sup>.

La négamycine forme des interactions avec l'anticodon de l'ARNt au site A. Le groupement amine (N4) de la molécule pourrait former une liaison hydrogène avec le nucléotide 34 de l'ARNt et son groupement carboxy-terminal interagit également avec l'ARNt via un ion magnésium, renforçant ainsi davantage son interaction avec l'ARNt. La capacité de la négamycine à induire de la translecture a longtemps été associée à un renforcement de la liaison ARNt au site A, augmentant ainsi les erreurs de décodage <sup>297,298</sup>. Toutefois, bien que ces nouvelles données cristallographiques soient cohérentes avec ce modèle, elles ne permettent pas de le confirmer <sup>294</sup>.



**Figure 63 : Fixation de la négamycine dans le ribosome 70S**

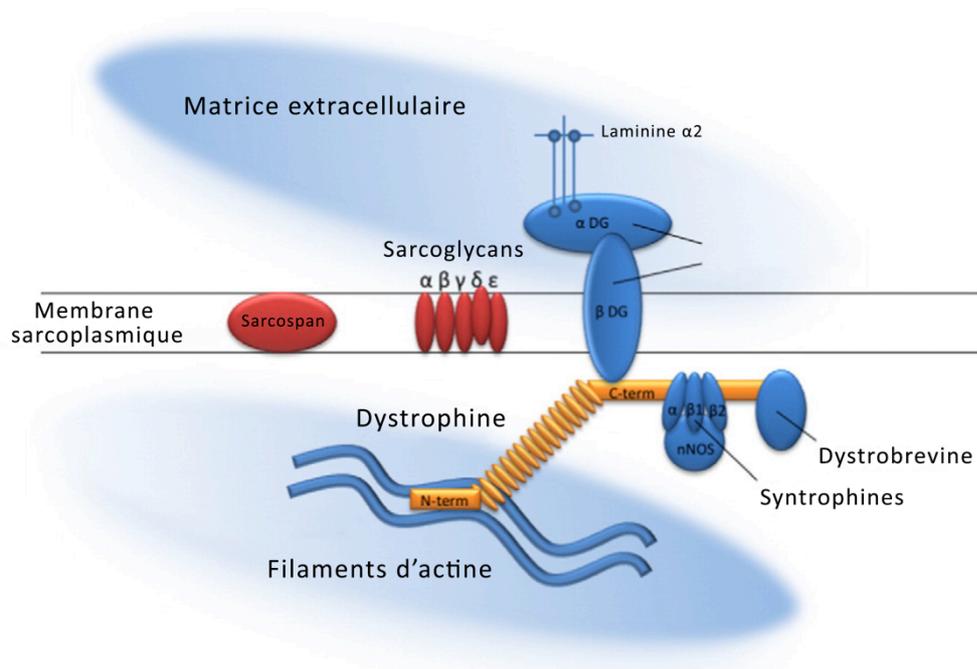
**A-C** - Vue d'ensemble du site de liaison de la négamycine (jaune) dans le ribosome 70S de *T. thermophilus* vu sous trois angles différents. La sous-unité 30S est représentée en jaune clair; la sous-unité 50S en gris clair; l'ARNm est représenté en magenta et les ARNt sont affichés en vert (site A), bleu foncé (site P) et orange (site E). **A** - La sous-unité 30S est vue de l'interface (30S-50S, la 50S et l'ARNt au site A ont été retirés pour plus de clarté). Dans **A** et **B**, l'emplacement de l'hélice 34 de l'ARNr 16S est mis en évidence par un contour noir. **D-F** - Vues rapprochées du site de liaison de la négamycine représentées sur les panneaux **A**, **B** et **C** respectivement. L'ion magnésium est représenté en vert. Les nucléotides dont les mutations chez *E. coli* confèrent une résistance à la négamycine sont encadrés. Figure adaptée de Polikanov et al., 2014 <sup>294</sup>

## IV. Les maladies à codon stop

La modification d'un seul nucléotide d'une séquence d'ADN peut conduire à l'apparition d'un codon stop prématuré <sup>22</sup>. Cette mutation, appelée mutation non-sens, entraîne la synthèse d'une protéine tronquée qui le plus souvent est non fonctionnelle et la dégradation de l'ARNm par la voie NMD. Une méta-analyse de plus de 7500 mutations non-sens dans 995 gènes responsables de maladies génétiques a montré que ces mutations non-sens représentent 11% des mutations impliquées dans les maladies génétiques humaines <sup>299</sup>. Parmi ces maladies génétiques, les plus étudiées sont la mucoviscidose <sup>300</sup> et la myopathie de Duchenne <sup>22</sup>. En raison de leur prévalence, d'importantes ressources ont été consacrées à la recherche de traitements pour ces maladies. Toutefois, selon le « National Institutes of Health Office of Rare Diseases Research » (NIH - <http://rarediseases.info.nih.gov>) et le « National Organization for Rare Disorders » (NORD - <http://www.rarediseases.org>), il existe plus de 7000 maladies génétiques présentes dans la population humaine. La grande majorité d'entre elles sont des maladies rares, car elles touchent entre 1 personne sur 1000 et 1 sur 200000 selon les pays, soit plus de 30 millions de personnes dans le monde. Sur l'ensemble des maladies génétiques identifiées, les mutations stop sont retrouvées pour 2,400 d'entre elles <sup>288</sup>. La nécessité de développer des stratégies thérapeutiques de suppression des PTC représente un enjeu important. Depuis 1996, une stratégie thérapeutique basée sur l'utilisation de molécules favorisant la translecture se développe afin de réexprimer une protéine entière et fonctionnelle à partir d'un gène interrompu par un codon stop prématuré <sup>301</sup>. D'autres stratégies comme les ARNt supprimeurs, l'inhibition de la voie de dégradation NMD ou le saut d'exon ont également vu le jour dans le même objectif.

### 1. Les pathologies humaines concernées

Entre 10% et 30% des maladies génétiques humaines sont liées à l'apparition d'une mutation non-sens dans la phase codante d'un gène <sup>302</sup>. Les premières maladies génétiques à avoir fait l'objet d'essais de suppression traductionnelle à visées thérapeutiques sont la mucoviscidose et la myopathie de Duchenne <sup>303</sup>. Depuis, plusieurs dizaines de pathologies humaines liées à des codons stop prématurés sont étudiées <sup>304</sup>. Toutefois, je m'attarderai ici principalement sur les deux maladies les plus étudiées, DMD et CF, et sur les cas de cancers pouvant être liés à un PTC.



**Figure 64 : Schéma du complexe glycoprotéiques associé à la dystrophine**

La dystrophine est une protéine essentielle pour la préservation de l'intégrité de la membrane cellulaire du muscle squelettique du cœur et des muscles lisses. La dystrophine est également exprimée dans certaines cellules non-musculaires, où sa fonction est moins bien définie. Dans les cellules musculaires, la dystrophine relie la face cytoplasmique de la membrane sarcoplasmique aux filaments d'actine à l'intérieur des fibres musculaires et agit comme un amortisseur qui stabilise la membrane. En l'absence de dystrophine, les contractions musculaires entraînent la détérioration voir la désintégration de la cellule. En revanche, si le domaine tige de la dystrophine (RD) est conservée, la dystrophine est partiellement fonctionnelle et fournit au moins une fonction de stabilisation. *Figure adaptée de Kole and Krieg, 2015* <sup>305</sup>

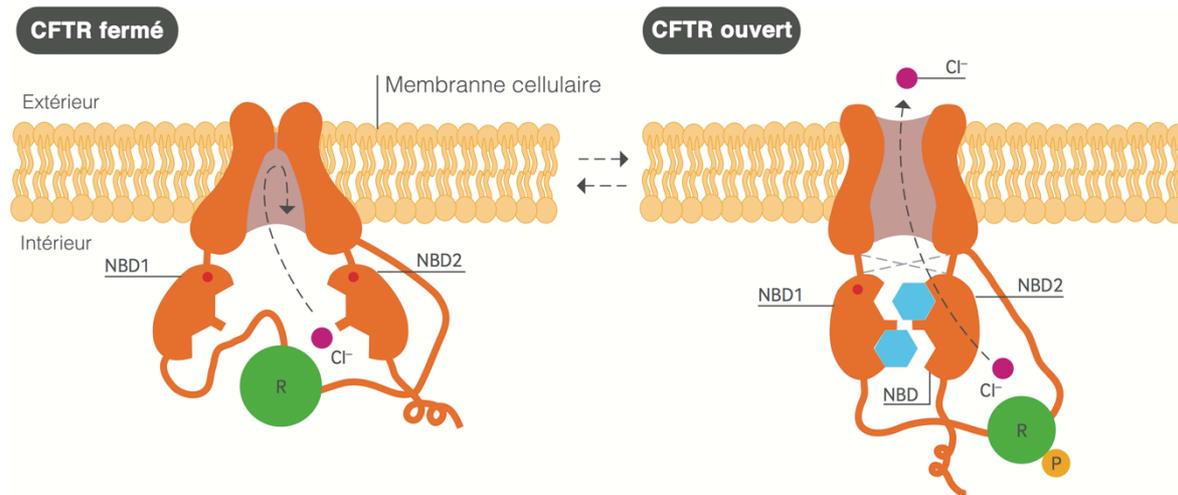
## A. La myopathie de Duchenne

La dystrophie musculaire de Duchenne est la maladie neuromusculaire la plus répandue, affectant jusqu'à 1 naissance masculine sur 3600 dans le monde. Elle est causée par la présence de mutations dans le gène de la dystrophine sur le chromosome X et aboutit à une perte de la fonction motrice (vers 10 ans) associée à une atteinte cardiaque et respiratoire <sup>22</sup>. La dystrophie musculaire de Becker (BMD) est une forme plus atténuée de dystrophinopathie, avec une incidence de 1/18000 naissances masculines. La BMD apparaît généralement plus tardivement que la DMD, entre 5 ans et 15 ans et la gravité varie selon les patients <sup>306</sup>.

Le gène de la dystrophine est le plus grand gène décrit chez l'homme (2,4 Mb) et son ARN messager (14 Kb) est principalement exprimé dans les muscles squelettiques et cardiaques, mais aussi en petite quantité dans le cerveau <sup>307</sup>. Dans le muscle sain, la dystrophine est localisée sur la surface intracellulaire du sarcolemme le long des fibres musculaires. Elle s'assemble avec le complexe glycoprotéique présent sur la membrane plasmique et sur les myofibrilles pour former le complexe de glycoprotéine associée à la dystrophine (DGC). La fonction essentielle de la dystrophine dans le muscle est de stabiliser les fibres musculaires pendant les contractions en se liant aux filaments d'actine via son domaine N-terminal et à la  $\beta$ -dystroglycane via son domaine C-terminal <sup>308</sup> (Figure 64).

La myopathie de Duchenne est le plus souvent associée à la présence de mutations telles que des délétions (65%), des duplications (6% - 10%), des mutations ponctuelles (environ 25%) comprenant les mutations faux-sens, les variations de site d'épissage et les mutations non-sens (13-15%), ou d'autres réarrangements qui interrompent le cadre de lecture <sup>309,310</sup>. La majorité de ces mutations conduisent à une perte de l'expression de la protéine dystrophine entraînant une atrophie musculaire sévère, une insuffisance respiratoire et cardiaque et la mort du patient avant l'âge de 30 ans. La perte de dystrophine perturbe le complexe DGC, ce qui provoque une instabilité de la membrane et la nécrose des fibres <sup>306</sup>. Toutefois, la dystrophine ayant une demi-vie *in vivo*, estimée à 26 semaines chez la souris *mdx*, cela suggère qu'une possible régression de la maladie pourrait avoir lieu en accumulant la dystrophine <sup>304</sup>.

Le modèle de classique pour étudier la DMD est la souris *mdx* (C57BL/10-Dmd<sup>mdx</sup>). Ces souris possèdent l'homologue murin du gène de la dystrophine portant la mutation non-sens (3185C>T) dans l'exon 23 et conduisant à l'apparition d'un codon stop UAA <sup>311</sup>. La dégénérescence musculaire est progressive :



**Figure 65 : Schéma de la structure du régulateur de la conductance transmembranaire de la fibrose kystique (CFTR)**

La structure de régulation est représentée en configuration fermée (à gauche) et ouverte (à droite). Les deux domaines transmembranaires forment le pore du canal. L'activation du canal est contrôlée par les deux domaines de liaison aux nucléotides (NBD1, NBD2) contenant chacun une séquence d'interaction avec l'ATP, en plus d'un domaine de régulation (R), qui contient de nombreux sites de phosphorylation (P). L'activation classique nécessite la phosphorylation du domaine R. Les NBD se lient à l'ATP et l'hydrolysent, ce qui provoque l'ouverture du canal via les domaines transmembranaires et les boucles extracellulaires, qui permettent également de stabiliser la protéine. *Figure adaptée de Quon et al., 2016*<sup>312</sup>

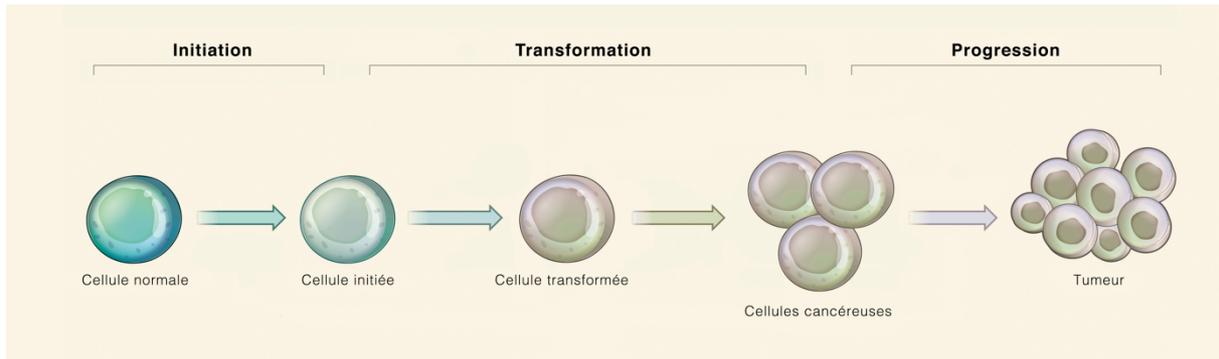
A 8 semaines, la force musculaire des membres antérieurs décroît, reflétant la nécrose du muscle tandis que la force des membres postérieurs est la même que celle d'une souris sauvage. L'utilisation de la souris *mdx* comme modèle préclinique a maintenant été standardisée<sup>313</sup>. Alors que chez l'homme, la DMD conduit à une perte précoce de la fonctionnalité du muscle, les souris *mdx* présentent des symptômes plus légers ainsi qu'une espérance de vie plus élevée, probablement en raison de la présence de l'utrophine, un homologue de la dystrophine<sup>314</sup>.

A ce jour, il n'existe pas de remède pour les patients atteints de DMD, et les interventions médicales actuelles sont davantage axées sur la prévention et la gestion des complications. Cependant, le développement de stratégies thérapeutiques pour réduire les effets pathologiques primaires et secondaires est largement étudié aujourd'hui. Ce point sera abordé dans la partie 2. *Stratégie thérapeutique*.

## B. La mucoviscidose

La mucoviscidose<sup>300</sup> est une maladie monogénique autosomale récessive rare qui touche plus de 70000 personnes dans le monde avec une incidence moyenne d'une naissance sur 3200. Elle est bien plus fréquente chez les personnes de race blanche d'origine européenne du Nord avec une prévalence de 1/3000 naissances, alors qu'en Afrique elle est de 1/12 000 et en Asie de 1/350 000.<sup>312</sup> Les personnes atteintes de mucoviscidose possèdent une mutation dans le gène CFTR codant un régulateur transmembranaire. Cette protéine est un canal ionique perméable au chlore, au thiocyanate et au glutathion et dont l'activation dépend de l'AMPc. Ce canal est impliqué dans la régulation du transport du chlore à travers les membranes cellulaires, ce qui permet de fluidifier le mucus. La protéine CFTR possède 2 domaines transmembranaires qui lui permettent de s'ancrer dans la membrane, 2 domaines cytoplasmiques de liaison aux nucléotides (NBD-1 et -2) et un domaine régulateur (R) (Figure 65). Elle est retrouvée dans de nombreux tissus, mais son expression dans les voies respiratoires, le petit et gros intestin, les canaux pancréatiques, biliaires, déférents, et les glandes sudoripares est directement liée à la pathogenèse de la maladie<sup>315</sup>.

Plus de 240 mutations sont connues pour déclencher la maladie. Malgré cette importante diversité allélique, 85 à 90% des personnes de race blanche sont porteurs d'au moins une copie de la mutation F508del (délétion d'une phénylalanine), qui est de loin la plus commune et qui conduit à la synthèse d'une protéine mal repliée et dégradée précocement dans le réticulum endoplasmique. Les 10 à 15% restants présentent des mutations ponctuelles dont 5 à 10% sont des mutations non-sens<sup>316</sup>.



**Figure 66 : La cancérogénèse**

Le développement d'un cancer résulte de la succession de trois étapes : l'initiation, la promotion et la progression tumorale. La première phase, l'initiation, correspond à l'accumulation d'anomalies génétiques. Au cours de la transformation, la cellule va acquérir des caractéristiques la rendant immortelle et lui permettant d'entrer dans une phase d'expansion clonale. Enfin, les cellules cancéreuses forment une tumeur et envahissent les autres organes et forment des métastases. *Figure adaptée de Tasselli et al., 2012*<sup>317</sup>

L'absence de CFTR ou la présence d'une forme partiellement fonctionnelle entraîne une déshydratation des sécrétions qui obstruent les structures glandulaires des voies respiratoires, biliaires et/ou pancréatiques, entraînant l'apparition de lésions tissulaires. L'atteinte principale se situe au niveau des poumons, car la viscosité des sécrétions favorise la colonisation par des bactéries et l'inflammation des tissus. Les patients atteints de mucoviscidose peuvent également présenter des défauts de l'immunité innée et adaptative <sup>318</sup>.

S'il n'existe aucun traitement curatif à ce jour, plusieurs molécules sont entrées en phase clinique pour soit potentialiser l'activité du canal CFTR, soit corriger son mauvais repliement ou encore permettre sa réexpression dans le cas de mutations non-sens <sup>312</sup>. Dans ce dernier cas, il a été montré que la réexpression de 5% de la protéine CFTR suffirait à atténuer la progression de la maladie <sup>319</sup>, tandis que 10 à 35% permettraient de réduire de façon significative la destruction des tissus pulmonaires <sup>320</sup>.

## C. Les cancers

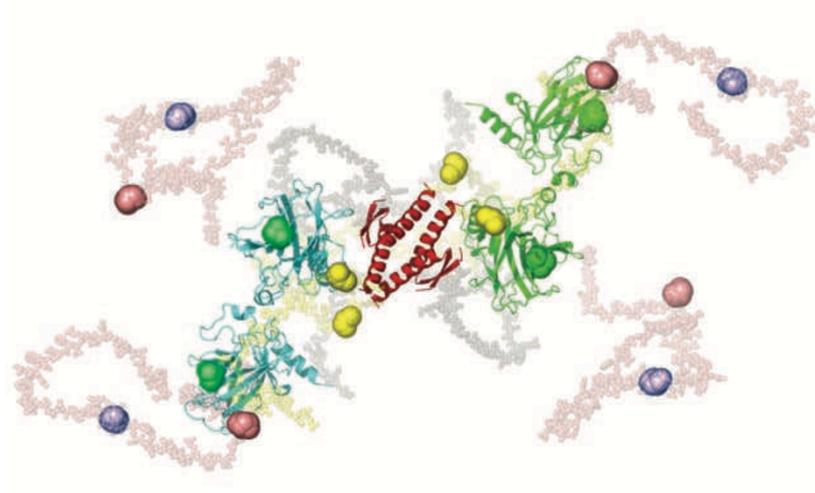
### a. La cancérogenèse

#### ★ La transformation cancéreuse

Le développement d'un cancer, appelé cancérogenèse, est un processus complexe au cours duquel une cellule normale subit une série progressive d'altérations conduisant à l'acquisition des capacités prolifératives et invasives ainsi qu'un important potentiel métastatique. Ces altérations sont classiquement définies en trois étapes <sup>321</sup> (Figure 66) :

- L'initiation : Lors de la première phase de la tumorigénèse, l'ADN de la cellule est muté par des agents cancérogènes chimiques ou physiques. Ces mutations induisent l'activation de gènes, appelés oncogènes, dont l'altération ou l'hyperexpression favorise la cancérogenèse, et/ou inactivent des gènes suppresseurs de tumeurs, également appelés antioncogènes. Le plus souvent, l'oncogène *KRAS* et l'antioncogène *P53* sont impliqués dans l'apparition des cancers. Ces cellules « initiées » ont le potentiel de former des tumeurs, mais d'autres mutations sont nécessaires afin qu'elles puissent se développer de façon autonome.

- La promotion : La seconde phase de la tumorigénèse se caractérise par l'accumulation d'un ensemble de modifications conférant à la cellule initiée un avantage sélectif (cellule transformée). Ces caractéristiques la rendent immortelle et lui permettent d'entrer dans une phase d'expansion clonale grâce à une augmentation de la prolifération cellulaire et/ou une diminution de la mort cellulaire).



**Figure 67 : Structure tridimensionnelle de p53 sous forme de tétramères**

Représentation de la protéine p53 sous forme de tétramère via son domaine de tétramérisation (rouge). Le domaine N-terminal est représenté en rose et le C-terminal en jaune. La protéine P53 est active lorsqu'elle sous forme de tétramère. *Figure adaptée de Whibley et al., 2009* <sup>322</sup>

- La progression tumorale : Au cours de cette dernière étape, les cellules continuent d'acquérir d'autres mutations qui leur confèrent un potentiel de malignité de plus en plus important (cellules malignes). Ces nouvelles propriétés leur permettent de former une tumeur et d'envahir d'autres organes (métastases).

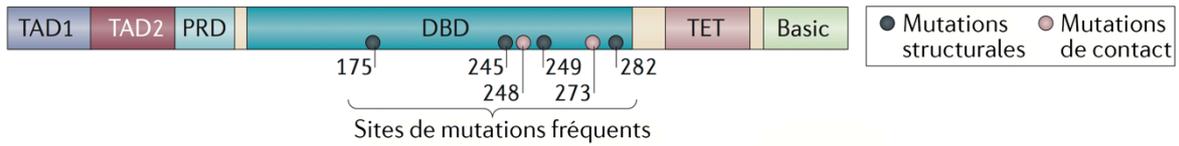
### ★ Les gènes associés au cancer

Contrairement à la mucoviscidose ou à la dystrophie musculaire, dans lesquels la mutation d'un seul gène provoque l'apparition de la maladie, les cancers nécessitent l'altération de plusieurs gènes. Les cellules de mammifères ont de multiples systèmes de protection pour contrer les effets potentiellement mortels de ces mutations. Toutefois, la présence d'altérations dans les oncogènes, antioncogènes et dans les gènes de stabilité contribue à l'activation du processus de cancérogenèse<sup>323</sup> :

- Les oncogènes (ex. *MYC*, *RAS*, *SRC*) ont la particularité d'être activés constitutivement dans les cellules tumorales. Leur activation via l'apparition d'une mutation dans un seul des deux allèles est suffisante pour conférer un avantage sélectif à la cellule, soit en augmentant la prolifération, soit en empêchant la cellule d'entrer en apoptose.

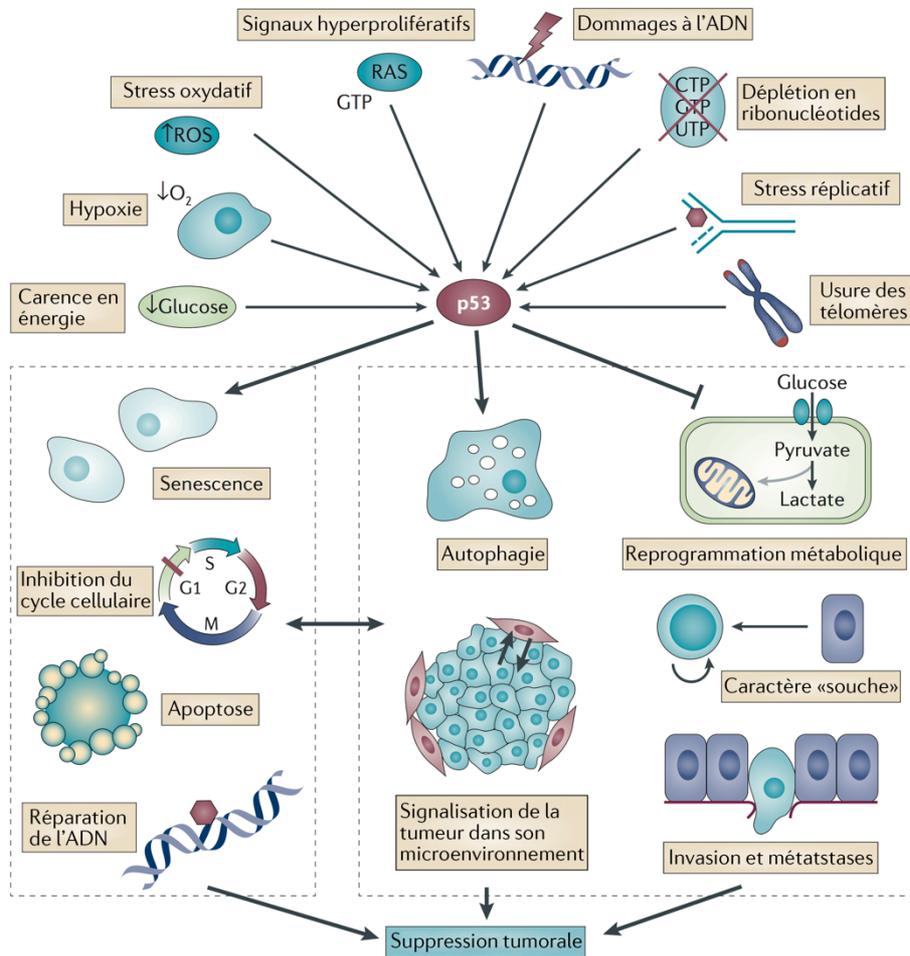
- A l'inverse, les gènes suppresseurs de tumeurs (ex. *TP53*, *APC*, *RB*) contrôlent la prolifération cellulaire et sont généralement impliqués dans l'arrêt du cycle cellulaire ou dans l'induction de l'apoptose (mort cellulaire programmée). Leur inactivation va donc conduire à une prolifération incontrôlée. Toutefois, les mutations des gènes suppresseurs de tumeurs étant dans la plupart des cas récessives, il est nécessaire d'éteindre les deux copies du gène. Dans ce cas-là, la première modification peut correspondre à une mutation d'origine héréditaire et c'est une modification sur l'autre allèle (souvent une délétion) qui va déclencher ou participer au processus de cancérogenèse (perte d'hétérozygotie). Néanmoins, dans le cas particulier du gène *TP53*, la majorité des mutants sont des mutants à effet dominant.

- Une troisième classe de gènes impliqués dans le cancer appelés gènes de stabilité, contrôle les mécanismes de réparations de l'ADN lors de l'étape de réplication (ex. *MSH2*, *-6*, *MLH1*, *MUTYH*), la recombinaison mitotique et la ségrégation des chromosomes (ex. *BRCA1*, *BLM*, *ATM*). Leur inactivation augmente significativement l'apparition de mutations dans tous les autres gènes, mais seules les mutations dans les oncogènes et les antioncogènes affectent la croissance cellulaire et la cancérogenèse. Comme pour les gènes suppresseurs de tumeurs, les deux allèles des gènes de stabilité doivent être inactivés.



**Figure 68 : La protéine p53**

La protéine p53 contient deux domaines d'activation de la transcription en N-terminal (TAD), un domaine riche en proline (PRD), un domaine de liaison à l'ADN (DBD), un domaine de tétramérisation (TET) et une région carboxy-terminale qui est riche en résidus de basiques (Basic). L'inactivation de la p53 dans les tumeurs humaines se produit généralement par des mutations faux-sens dans le DBD de la protéine p53. Six sites de mutations fréquents «points chauds» sont représentés et classés en deux catégories : les mutations de contact qui altèrent des résidus qui sont cruciaux pour l'interaction avec l'ADN et les mutations structurales qui compromettent le repliement du domaine de liaison à l'ADN. Dans les deux cas, la protéine p53 n'est plus capable d'activer ses gènes cibles. *Figure adaptée de Bieging et al., 2014* <sup>324</sup>



**Figure 69 : Signaux d'activation et réponses de p53**

La protéine p53 peut être activée suite à une variété de stress tels que l'usure des télomères, un stress réplcatif ou oxydatif, la déplétion en ribonucléotides, des dommages à l'ADN, des signaux hyperprolifératifs, l'hypoxie ou une carence en énergie. Une fois stabilisée, la protéine p53 va pouvoir activer les processus classiques de sénescence, d'apoptose, de réparation de l'ADN et ou d'inhibition du cycle cellulaire. La protéine p53 peut également inhiber la reprogrammation métabolique, la reprogrammation des cellules différenciées en cellules souches et le caractère invasif et métastatique et activer signalisation de la tumeur dans son microenvironnement et l'autophagie des cellules cancéreuses. *Figure adaptée de Bieging et al., 2014* <sup>324</sup>

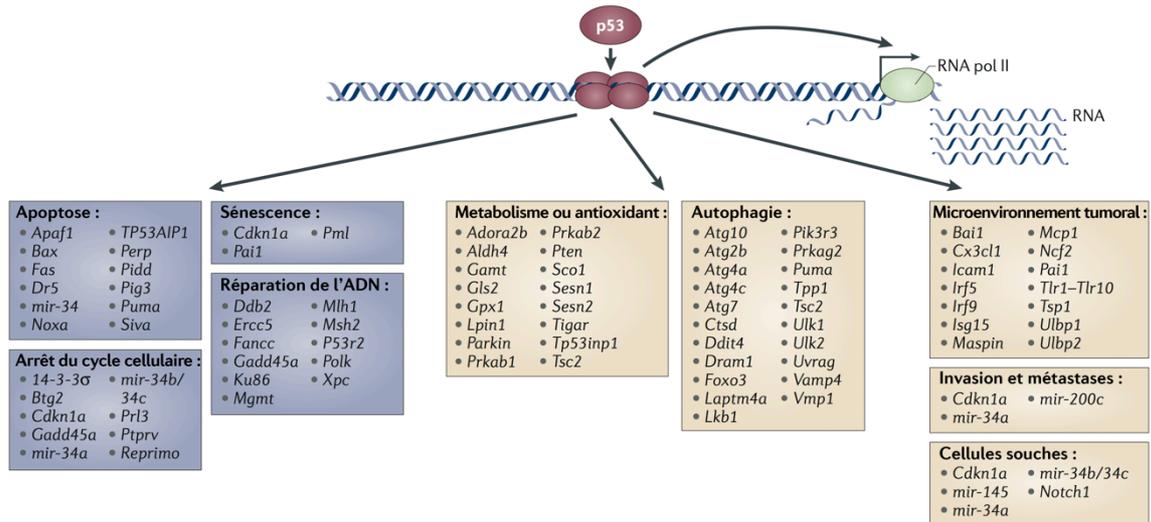
L'apparition de mutations dans ces trois classes de gènes peut se produire dans la lignée germinale, ce qui entraîne des prédispositions héréditaires au cancer, ou dans des cellules somatiques individuelles, ce qui entraîne des tumeurs sporadiques. Dans la lignée germinale, les mutations les plus fréquentes sont ponctuelles (substitutions, délétions ou insertions), tandis que tous les types de mutations peuvent être retrouvés dans les cellules tumorales. En effet, les cancers constituent l'un des rares cas de maladies dans lesquelles des mutations somatiques survenant après la naissance sont pathogènes.

### *b. Le gène suppresseur de tumeur P53*

Le produit de l'antioncogène ou gène suppresseur de tumeur *P53* est un des obstacles majeurs dans ce processus, c'est pourquoi il est impliqué dans plus de 50% des cancers <sup>325</sup>. Le gène *P53* est situé sur le chromosome 17 et contient 11 exons. Contrairement aux autres gènes suppresseurs de tumeurs dont les mutations les plus fréquentes sont des insertions, délétions et mutations non-sens, plus de 80% des altérations du gène *P53* sont des mutations faux-sens dont la plupart sont regroupés sur quelques positions (hot-spot) dans le domaine de fixation à l'ADN <sup>326</sup>. Les mutations non-sens représentant environ 8% de l'ensemble des mutations du gène *P53* <sup>327</sup>.

La protéine p53, appelée ainsi en raison de son poids moléculaire apparent de 53 KDa (Figure 67), est un facteur de transcription de 393 acides aminés, composé de 6 domaines très conservés <sup>322</sup> (Figure 68) : deux domaines de transactivation de la transcription (TAD) situés en N-terminal ; un domaine riche en proline (PRD) ; un domaine de liaison à l'ADN (DBD) qui est la cible de plus de 90% des mutations retrouvées dans les cancers humains ; un domaine de tétramérisation (TET) essentielle à la fonctionnalité qui contient un signal d'export (NES, Nuclear Export Signal) et un signal de localisation au noyau (NLS, Nuclear Localization Domain) ; et enfin, une région carboxy-terminal qui est riche en résidus basiques et permet à p53 de se fixer de manière non spécifique à l'ADN lorsque celui-ci est endommagé.

La protéine p53 agit comme la gardienne du génome en réponse à des stress cellulaires variés tels que la privation de nutriments, l'hypoxie, les signaux hyperprolifératifs, les lésions de l'ADN déclenchés par le stress répliatif, le raccourcissement des télomères, le stress oxydatif, et l'épuisement en ribonucléotides (Figure 69). La transmission de ces signaux conduit à la stabilisation et à l'activation de la protéine par l'intermédiaire de modifications post-traductionnelles et probablement par un changement de conformation de la protéine <sup>328</sup>.



**Figure 70 : Gènes cibles de la protéine p53**

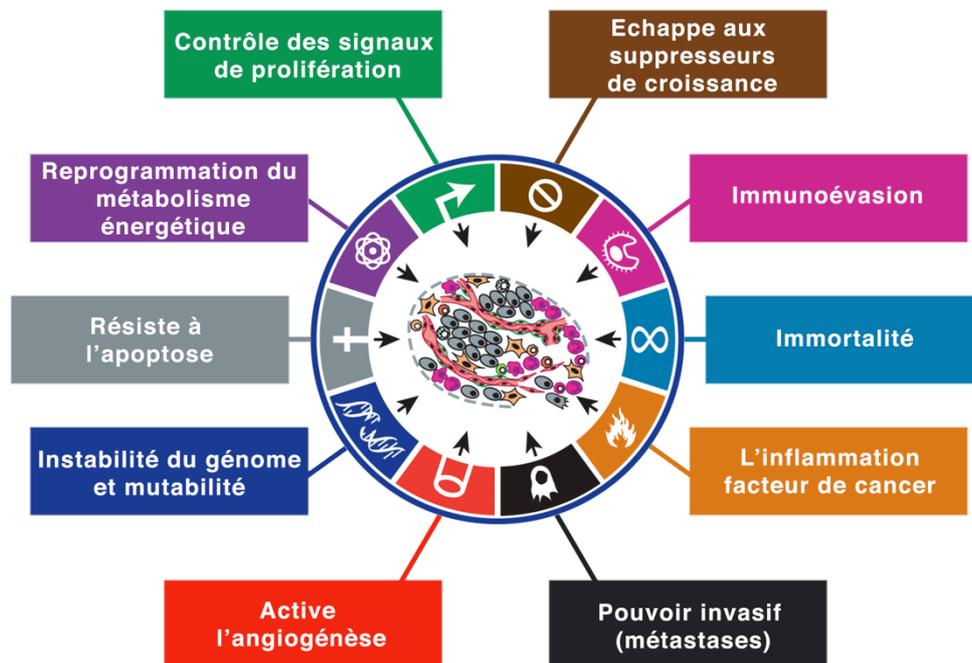
Liste des gènes cibles clés induits par p53 et impliqués dans les processus importants pour la suppression tumorale. Cette liste comprend les réponses canoniques associées à p53 : l'apoptose, l'arrêt du cycle cellulaire, la sénescence et la réparation de l'ADN (violet) et les processus récemment associés à suppression tumorale p53-dépendante comme le contrôle du métabolisme, l'autophagie, les interactions avec le microenvironnement de la tumeur, l'invasion et les métastases, et la biologie des cellules souches (beige). La régulation p53-dépendante de ces gènes a été identifiée chez la souris et ou dans les cellules humaines. Toutefois, quelques gènes ont été actuellement identifiés et ou montré comme étant régulé par p53 que dans l'une de ces deux espèces. *Figure adaptée de Biegging et al., 2014*<sup>324</sup>

La protéine p53 va alors induire l'activation ou la répression transcriptionnelle d'un certain nombre de gènes cibles permettant une réponse adaptée au signal reçu (Figure 70). Les réponses possibles sont les suivantes : l'arrêt du cycle cellulaire, la sénescence, l'apoptose et la réparation d'ADN. En plus de sa capacité à bloquer complètement la progression du cycle cellulaire en réponse à un signal de stress, le niveau basal de p53 suffit à freiner la vitesse de progression à travers le cycle cellulaire. Le facteur de transcription p53 peut également réguler plusieurs processus cellulaires supplémentaires qui permettent de lutter contre le développement tumoral : en bloquant la reprogrammation métabolique, en limitant l'accumulation des espèces réactives de l'oxygène (ROS), en activant l'autophagie qui correspond à l'autolyse d'une partie du cytoplasme par ses propres lysosomes, en rétablissant les signaux antiprolifératifs, en inhibant le renouvellement des cellules-souches et la reprogrammation de cellules différenciées en cellules souches, et en endiguant l'invasion et les métastases. La régulation de ces processus par la protéine p53 lui permet de promouvoir directement la suppression tumorale<sup>324</sup>. En absence de stress, la protéine p53 a une durée de vie très courte, car elle est dégradée par Hdm2, une ubiquitine ligase E3. Celle-ci interagit avec la partie N-terminal de la p53 et ajoute des ubiquitines en C-terminal, permettant de l'adresser au protéasome. Lorsque la p53 est stabilisée, elle active la transcription du gène Hdm2, formant ainsi une boucle de régulation négative contrôlant le niveau de la p53.

### *c. Caractéristiques des cellules tumorales*

Au cours de la cancérogenèse, les cellules normales évoluent progressivement vers un état néoplasique où elles acquièrent successivement un certain nombre de compétences caractéristiques. Ces compétences seraient nécessaires aux cellules pour devenir cancéreuses et finalement malignes. Ainsi, une cellule cancéreuse est caractérisée par l'acquisition de dix traits biologiques<sup>329</sup> (Figure 71). Toutefois, je n'aborderai ici que les plus pertinents pour mes travaux de recherche :

- Contrôle des signaux de prolifération : Les tissus normaux contrôlent soigneusement la production et la libération des signaux qui permettent l'entrée et la progression au travers du cycle de division cellulaire. Ce contrôle permet d'assurer une homéostasie du nombre de cellules et donc du maintien de l'architecture tissulaire et de sa fonction. Ces signaux influencent également d'autres propriétés des cellules, telles que la survie cellulaire et le métabolisme énergétique. Les cellules cancéreuses, en dérégulant ces signaux, acquièrent la capacité de contrôler et maintenir les signaux de prolifération de plusieurs façons.



**Figure 71 : Traits caractéristiques des cellules tumorales**

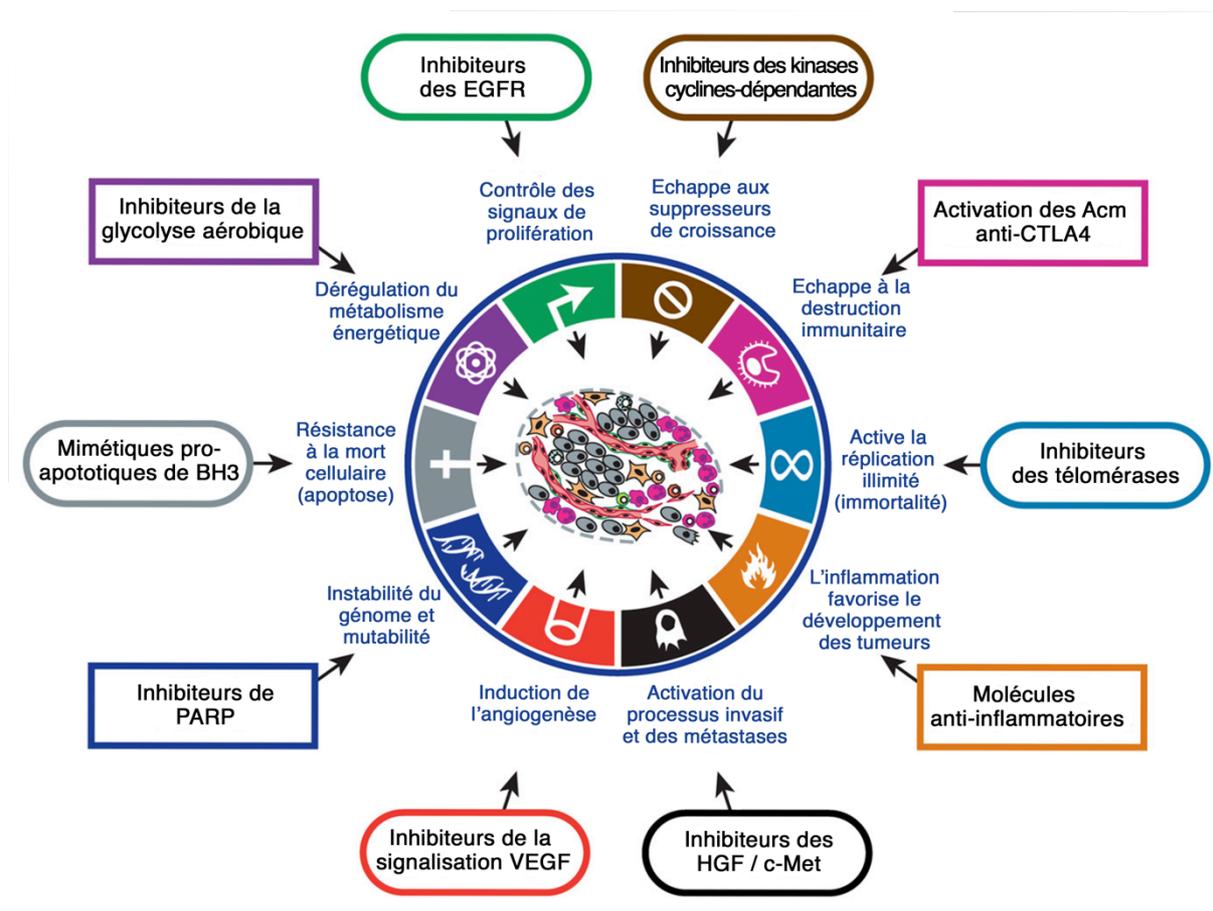
Les cellules cancéreuses ont acquis au cours de leur développement plusieurs caractéristiques dont certaines leur seraient indispensable et permettraient de les caractériser. Ces capacités ou traits caractéristiques sont au nombre de dix : Le contrôle des signaux de prolifération, leur capacité à échapper aux supprimeurs de croissance, leur résistance à l'apoptose, leur nombre de cycle de division illimité (immortalité), leur capacité à activer l'angiogénèse et à pouvoir envahir les tissus et y créer des métastases, leur capacité à reprogrammer le métabolisme énergétique, leur plasticité et instabilité du génome, leur capacité à éviter leur destruction par le système immunitaire et le caractère inflammatoire favorisant les tumeurs et l'instabilité du génomes. *Figure adaptée de Hanahan et al., 2011* <sup>329</sup>

Elles peuvent soit produire leurs propres ligands qui se fixent sur les facteurs de croissance et stimuler la prolifération via un récepteur spécifique, soit émettre un signal vers les autres cellules pour qu'elles produisent des facteurs de croissance, soit activer constitutivement la voie de signalisation opérant en aval des récepteurs aux facteurs de croissance.

- Échappement aux suppresseurs de croissance : En plus de la capacité caractéristique d'induire et de maintenir les signaux de prolifération, les cellules cancéreuses doivent contourner les processus de régulation négative de la prolifération cellulaire. Bon nombre de ces processus dépendent de l'action des gènes suppresseurs de tumeur. Les deux plus connus codent les protéines RB (Rétinoblastome) et p53. Elles jouent un rôle central dans le contrôle des deux circuits clés de régulation qui régissent la prolifération cellulaire, la sénescence et l'apoptose. L'inactivation d'une de ces voies permet d'activer la prolifération cellulaire.

- Résistance à l'apoptose : La mort cellulaire programmée par apoptose sert de barrière naturelle au développement du cancer. Les deux principaux stress responsables de l'induction de l'apoptose proviennent de déséquilibres cellulaires provoqués par les oncogènes et de dommages à l'ADN associés à une hyperprolifération cellulaire. Pour contrer ce processus, les cellules tumorales ont dû développer une variété de stratégies leur permettant de limiter ou contourner l'apoptose. La stratégie la plus courante est l'inactivation du gène suppresseur de tumeur TP53, ce qui empêche la cellule d'activer l'apoptose lors de dommages à l'ADN. Les cellules tumorales peuvent également augmenter l'expression des régulateurs anti-apoptotiques (Bcl-2, Bcl-xL) ou des signaux de survie (IGF1 / 2). Pour cela, les cellules tumorales peuvent soit réguler négativement des facteurs proapoptotiques (Bax, Bim, Puma), soit court-circuiter l'activation de la voie apoptotique induite par certains ligands. La multiplicité des mécanismes permettant de contrer l'activation de l'apoptose reflète bien la diversité des signaux responsables de son activation.

- L'immortalité : Il fut très tôt admis que les cellules cancéreuses avaient besoin d'un potentiel réplicatif illimité pour générer des tumeurs macroscopiques. Le raccourcissement des télomères détermine le potentiel réplicatif limité des cellules normales et doit donc être surmonté par les cellules cancéreuses. Le nombre limité de divisions auxquelles sont soumises les cellules normales humaines a été associé à deux barrières : la sénescence qui correspond à une entrée généralement irréversible dans un état non prolifératif, mais viable, et l'état de crise impliquant l'apoptose. Les cellules normales en culture qui répètent un grand nombre de cycles de division entrent d'abord en sénescence et les cellules qui réussissent à contourner ce processus, entrent alors dans une phase de crise induisant l'apoptose de la grande majorité des cellules.



**Figure 72 : Cibles thérapeutiques des caractéristiques du cancer**

Les médicaments présentés ici ont été développés pour interférer avec les capacités nécessaires à la croissance et à la progression tumorale. Ces traitements sont en cours d'essais cliniques ou sont dans certains cas déjà approuvés pour une utilisation clinique dans le traitement de certaines formes de cancer humain. Les médicaments énumérés ne sont que des exemples illustratifs ; d'autres traitements en cours de développement ont des cibles moléculaires et des modes d'actions différents. EGFR : récepteur du facteur de croissance épidermique ; Acm : anticorps monoclonal ; CTLA4 : lymphocyte T cytotoxique antigène 4 ; HGF / C-Met : récepteur du facteur de croissance hépatocytaire ; VEGF : facteur de croissance de l'endothélium vasculaire ; PARP : poly(ADP-ribose) polymérase ; BH3 : domaine de certaines protéines de la famille des protéines Bcl-2. *Figure adaptée de Hanahan et al., 2011* <sup>329</sup>

Les cellules capables de sortir de cet état de crise présentent un potentiel réplicatif illimité et cette transition est appelée immortalisation. Ces deux barrières anticancéreuses sont contournées par les cellules tumorales grâce à leur capacité à maintenir les télomères suffisamment longs. Pour cela, ces cellules ont une régulation positive de l'expression de la télomérase ou moins fréquemment, et un mécanisme de maintenance des télomères activé.

- Les cellules tumorales ont bien d'autres capacités et peuvent notamment induire l'angiogenèse, provoquant ainsi une importante vascularisation qui permet de soutenir la prolifération cellulaire. Elles peuvent aussi s'étendre et créer des métastases. Plus récemment, il a été montré que l'inflammation contribuerait au développement tumoral et que les cellules tumorales seraient capables de reprogrammer leur métabolisme énergétique et de désactiver les composants du système immunitaire <sup>329</sup>. Au vu du nombre de mutations nécessaires pour acquérir l'ensemble de ces traits caractéristiques et de l'efficacité des systèmes de maintenance et réparation du génome, les cellules cancéreuses augmentent souvent le taux de mutation. Pour cela, plusieurs composantes de la surveillance et de la maintenance du génome peuvent être inactivées, ce qui contribue à la progression tumorale des cellules précancéreuses en augmentant la fréquence d'apparition de mutations favorables.

#### *d. Cibles thérapeutiques*

Certaines cellules cancéreuses présentent une importante quantité d'oncogènes et une hypersensibilité à la réintroduction de gènes suppresseurs de tumeurs. Weinstein IB a nommé cette caractéristique le talon d'Achille du cancer <sup>330,331</sup>. Pourtant si les cellules cancéreuses évoluent simplement par l'acquisition successive de mutations, réparer un seul défaut devrait conduire à un effet modeste sur la croissance (Figure 72). Le processus de cancérogenèse n'est donc pas que la somme de mutations. Ce sont plutôt ces mutations qui ont permis aux cellules cancéreuses de détourner les voies de signalisation utilisées par la cellule. La réintroduction d'un gène suppresseur de tumeur ou la perte d'un oncogène perturbe considérablement l'équilibre construit par les cellules tumorales au cours de la cancérogenèse.

Toutefois, les réponses cliniques qui résultent de la réintroduction d'un gène suppresseur de tumeur ou de la perte d'un oncogène sont généralement transitoires et suivies de rechutes <sup>329</sup>. En réalité, chacune des capacités de base des cellules cancéreuses est régulée par des voies de signalisation partiellement redondantes. L'utilisation d'un traitement spécifique contre une de ces voies provoquera la mort d'une certaine population de cellules tumorales alors que d'autres se seront adaptées et seront devenues résistantes.

Cette adaptation, qui peut être réalisée par mutation ou par reprogrammation épigénétique permet de rétablir la capacité fonctionnelle, ce qui permet de renouveler la croissance tumorale et de provoquer une rechute clinique. La réexpression simultanée de plusieurs protéines clés impliquées dans plusieurs traits caractéristiques des cellules cancéreuses permettrait de prévenir l'apparition de résistances.

#### D. Les autres maladies à codons stop étudiées

Un grand nombre de maladies à codons stop a également fait l'objet de plusieurs études pour évaluer l'effet des supprimeurs de codons stop. Pour revue : <sup>304,332</sup>. Parmi les plus connues et les plus étudiées, sont retrouvées :

- La polykystose rénale de type dominant (ADPKD), la maladie héréditaire monogénique du rein la plus fréquente. Elle se caractérise par l'apparition de kystes au niveau des reins. La présence de mutations dans les gènes *PKD1* et *PKD2* conduit à l'apparition de la maladie avec une sévérité plus ou moins importante selon le type et la localisation de la mutation. Dans 85% des cas, c'est le gène *PKD1* qui est impliqué. Et 25% des mutations impliquées dans cette maladie sont des mutations non-sens <sup>333</sup>

- L'ataxie télangiectasie (A-T) aussi appelé syndrome de Louis-bar est une maladie génétique autosomique récessive. Elle se caractérise par une ataxie (atteinte du système nerveux conduisant à un manque de coordination des mouvements volontaires) ainsi qu'un déficit immunitaire. Les mutations non-sens du gène *ATM* représentant ici 15% des mutations retrouvées chez les patients atteints <sup>334</sup>.

- La thalassémie, aussi appelée maladie des globules rouges, est une maladie génétique atteignant la production d'hémoglobine. Elle se caractérise par l'absence de la chaîne alpha ou bêta. Les mutations non-sens représenteraient 10% des mutations responsables de la maladie <sup>335</sup>.

- Le syndrome de Rett (RTT) est une maladie génétique neurologique qui altère le développement du système nerveux central et touche majoritairement les femmes. Le gène responsable de cette maladie est *MECP2*, localisé sur bras long du chromosome X et est causé dans plus de 60% des cas par des mutations non-sens <sup>336</sup>

- Le syndrome d'Usher de type 1 (USH1) est une maladie génétique qui associe une surdité congénitale et des troubles de la vue. Il existe deux autres formes de cette maladie, type II et III, qui touchent des gènes différents. Le syndrome d'Usher de type I apparaît à la suite de mutations dans au moins 6 gènes, dont le taux de mutation non-sens est d'environ 12% <sup>337</sup>.

- Le syndrome de Hurler (MPS1H) est une maladie génétique lysosomale caractérisée par des altérations musco-squelettiques, un retard de développement et d'importants troubles de la vision et de l'audition. Cette maladie est due à la présence de mutations dans le gène *IDUA* provoquant une déficience enzymatique complète. Dans certaines populations, les mutations non-sens représentent jusqu'à 76% des cas<sup>338</sup>.

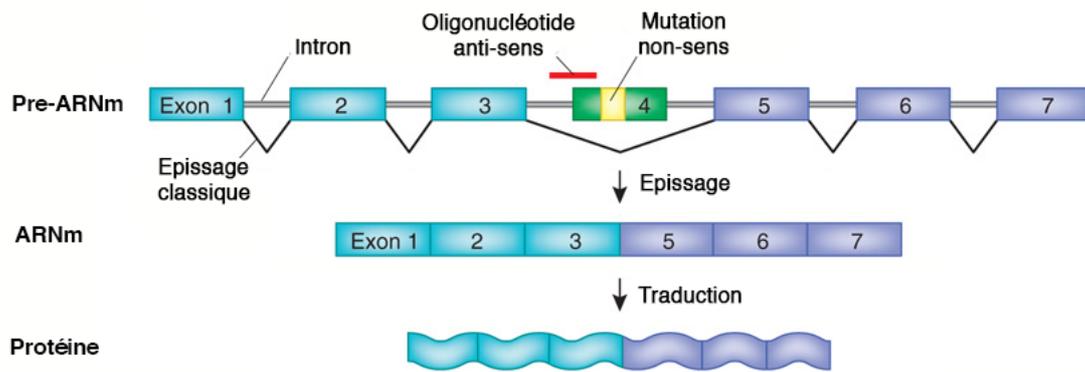
- L'hémophilie est une anomalie hémorragique héréditaire due à l'absence ou au déficit d'un facteur de la coagulation. Deux types d'hémophilie existent selon le gène touché et sont appelés hémophilie A (HA) qui est la forme la plus répandue, et l'hémophilie B (HB). Deux bases de données de mutations ont rapporté qu'environ 9 à 10% des mutations de HA et 8 à 9% de HB sont causés par des mutations non-sens<sup>339,340</sup>.

- L'acidémie méthylmalmonique (AMM) est une maladie autosomale récessive. Les patients atteints ont d'importants troubles digestifs, une perte du tonus musculaire et un retard de développement mental. Cette maladie est due à un déficit héréditaire de l'enzyme méthylmalonyl-CoA mutase qui intervient dans le métabolisme de plusieurs acides aminés retrouvés dans les protéines de l'alimentation. Les mutations non-sens représentent ici 16% des mutations responsables de l'apparition de la maladie<sup>341</sup>.

- L'amyotrophie spinale (SMA) est une maladie neuromusculaire rare caractérisée par une perte des neurones moteurs et une atrophie musculaire. Cette maladie est causée par la présence de mutations dans le gène *SMN1* qui encode la protéine SMN nécessaire à la survie des neurones moteurs. Le nombre et le type de mutations identifiés à ce jour sont insuffisants pour calculer le taux de mutation non-sens retrouvés dans le gène *SMN1*.

## 2. Stratégies thérapeutiques

Les maladies à codon stop prématuré conduisent fréquemment à la perte totale de la fonction de la protéine et à une forte diminution de la quantité d'ARNm en raison de leur dégradation par la voie NMD. En effet, selon la position de la mutation dans le gène, la probabilité de toucher un domaine essentiel de la protéine n'est pas la même. Ainsi, les mutations non-sens situées en C-terminal ont une plus faible probabilité d'éteindre la fonction de la protéine. Trois approches principales ont été développées pour traiter ces maladies : la première consiste à remplacer la mutation non-sens par la séquence sauvage en utilisant des techniques de réparation de gènes. En particulier, le saut d'exon permet de retirer l'exon muté. Le développement de la technique CRISPR/CAS9 pourrait aussi permettre de réparer spécifiquement le gène touché.



**Figure 73 : Le saut d'exon**

L'ARN transcrit à partir d'un gène contenant initialement des séquences correspondant à des exons et des introns. Dans l'exemple représenté, le pré-ARNm contient un codon stop prématuré dans l'exon 4. L'oligonucléotide anti-sens (rouge), en ciblant le site d'épissage 3' de l'intron 3, est capable de bloquer l'épissage de l'exon 4, conduisant à l'exclusion de cet exon lors de l'étape d'épissage. L'ARNm tronqué contient maintenant un cadre de lecture ouvert pouvant être traduit en une protéine tronquée mais potentiellement partiellement fonctionnelle. Figure adaptée de Chamberlain and Chamberlain, 2010<sup>342</sup>

La seconde approche consiste à stimuler la translecture des codons stop en utilisant des ARNt suppresseurs ou en favorisant le recodage grâce à la pseudouridylation de codons stops voir en inhibant la voie NMD. Enfin, la dernière approche est également basée sur la suppression des codons stop, mais la stratégie développée fait appel à des molécules chimiques pour induire la translecture. Cette stratégie ayant fait l'objet d'un de mes projets de thèse, elle sera développée plus en détail dans la partie suivante : 3. *Suppression traductionnelle pharmacologique.*

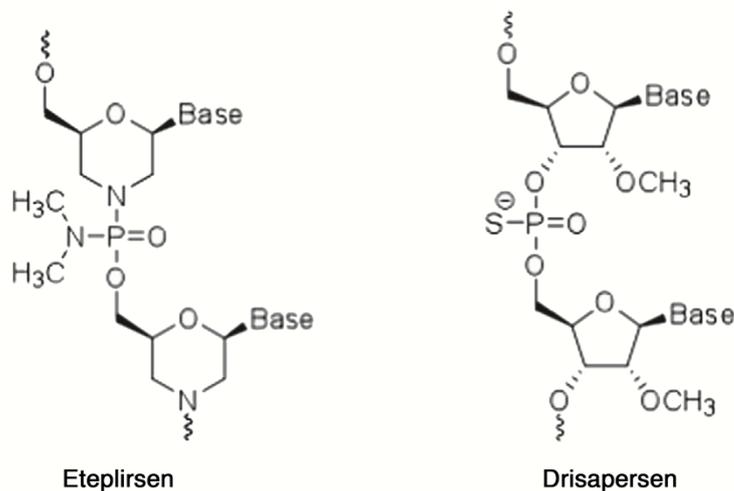
### A. Le saut d'exon

A la fin des années 1970, une approche utilisant des oligonucléotides anti-sens (AON) comme outils de régulation négative de l'expression de gènes spécifiques commence à émerger <sup>343</sup>. Plus tard, il a été découvert que les AON étaient également capables de corriger des défauts d'épissage, ce qui a ouvert un important champ d'application dans le traitement de maladies. La plupart des essais ont été faits dans l'objectif de développer un traitement contre les maladies neuromusculaires, dont la DMD <sup>344</sup>.

Comme mentionné précédemment, la plupart des mutations responsables de la DMD comprennent des délétions (65%), des duplications (6-10%) et des mutations non-sens (13-15%). Ces mutations conduisent à une traduction aberrante et à l'absence de dystrophine dans le muscle. La stratégie du saut d'exon permettrait d'éliminer l'exon muté pour rétablir le cadre de lecture et induire l'expression d'une forme raccourcie de la dystrophine toujours fonctionnelle (Figure 73). Pour cela, les AON interfèrent avec le mécanisme d'épissage classique en masquant les sites d'épissages.

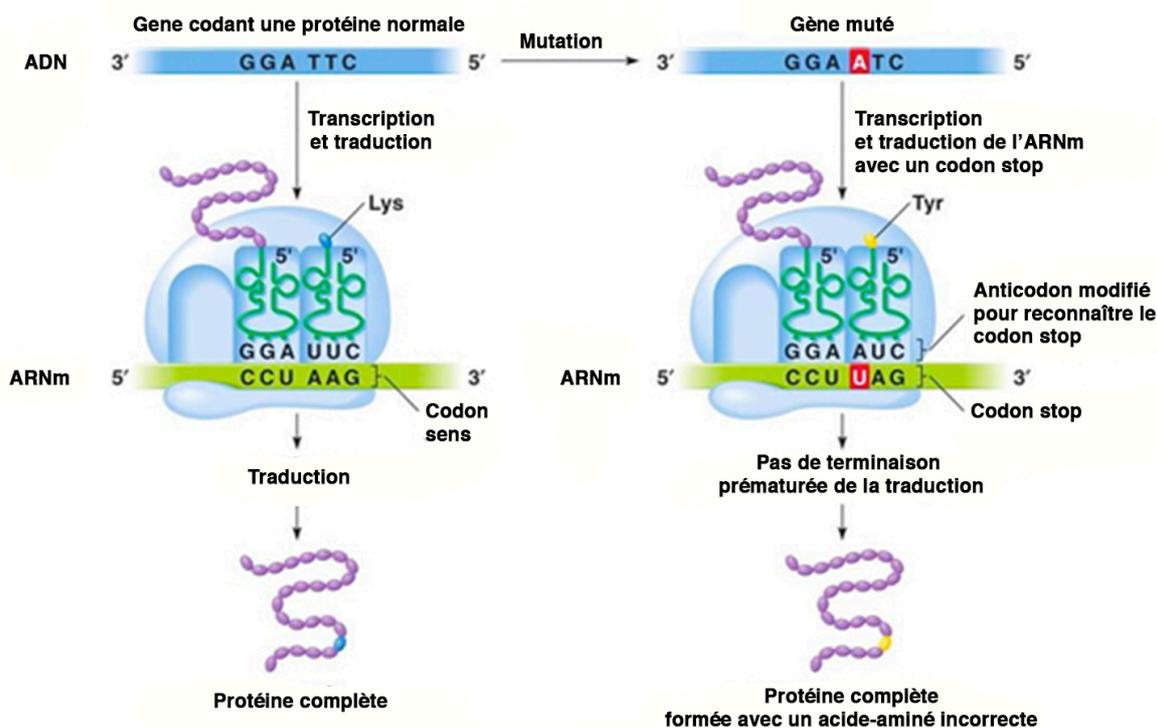
La preuve de principe de l'utilisation du saut d'exon pour la DMD a d'abord été apportée dans les cellules lymphoblastoïdes en 1996 puis dans des cellules de souris en culture *in vitro* <sup>345,346</sup>. Depuis, de nombreuses études *in vivo* ont fourni des preuves précliniques du potentiel thérapeutique de cette stratégie dans plusieurs modèles animaux. En particulier, le modèle de la souris *mdx*, qui porte une mutation non-sens dans l'exon 23, a été largement utilisé pour tester l'efficacité du saut d'exon. Ces essais ont abouti à la restauration généralisée de la dystrophine dans les muscles squelettiques, mais pas dans le cœur <sup>203</sup>. Il a également été montré que cette stratégie permettait de diminuer de manière significative la pathologie musculaire et d'améliorer l'activité locomotrice avec une administration répétée à long terme <sup>347</sup>.

Suite à ces résultats encourageants, le saut d'exon a rapidement évolué vers une évaluation clinique chez des patients atteints de DMD. Toutefois, l'hétérogénéité génétique des patients nécessitait d'adapter cette approche à chaque cas, faisant de cette maladie un exemple de médecine personnalisée.



**Figure 74 : Structure chimique de l'eteplirsen et drisapersen**

L'eteplirsen (phosphorodiamidate morpholino oligomère) et le drisapersen (2'O-méthyl-phosphorothioate-dinucléoside) sont deux molécules développées respectivement par Sarpta Therapeutics Inc. et BioMarin Pharmaceutical Inc. dans le traitement de la DMD. Ces deux molécules sont en cours d'essais cliniques pour les patients dont le gène de la dystrophine contient une mutation non-sens dans l'exon 51. *Figure adaptée de Kole and Krieg, 2015*<sup>305</sup>



**Figure 75 : Les ARNt suppresseurs**

La séquence d'un ADN codant une protéine normale est transcrit en ARNm et traduit en protéine. L'apparition d'une mutation non-sens dans une phase codante conduit à la synthèse d'une protéine tronquée. Les ARNt suppresseurs sont des ARNt modifiés pour reconnaître un codon stop. Dans l'exemple schématisé ici, l'ARNt<sup>Tyr</sup> décode le codon stop UAG, ce qui permet la synthèse d'une protéine complète. Cette protéine contient toutefois un acide-aminé incorrecte qui peut avoir des effets plus ou moins délétères sur son activité biologique. *Figure adaptée de Tortora et al., 2006*<sup>348</sup>

Selon la base de données de la dystrophie musculaire de Leiden (<http://www.dmd.nl/nmdb2/home.php>), le saut d'exon est potentiellement applicable à environ 83% des patients atteints de DMD<sup>349</sup>. La plupart des mutations étant présentes entre les exons 43 et 53, le saut d'un même groupe d'exons permettrait de traiter un important groupe de patients. L'exemple le plus frappant est l'exon 51 qui porte une mutation dans 13% des cas. Pour cette raison, les premiers essais cliniques ont été réalisés sur cet exon.

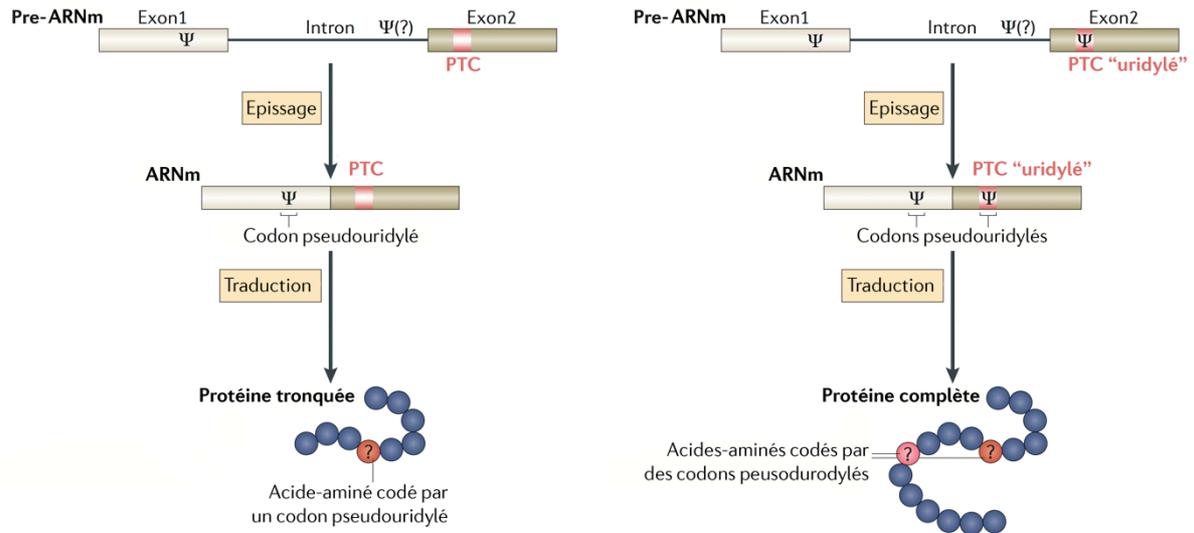
Deux médicaments (AON) ont émergé sur le marché, le drisapersen (Kyndrisa) développé par BioMarin Pharmaceutical Inc. et l'eteplirsén par Sarpta Therapeutics Inc (Figure 74). Le drisapersen a fait l'objet d'un premier essai clinique de phase II qui rapporta la détection de faibles niveaux de réexpression de dystrophine ainsi qu'une stabilisation des fonctions motrices chez les patients. Toutefois, l'essai de phase III n'a pas permis de mettre en évidence une amélioration clinique significative. Un deuxième essai clinique de phase III est en cours (NCT01803412). BioMarin a déposé une demande de mise sur le marché (NDA - New Drug Application) en 2015 que la Food and Drug Administration (FDA) refusa début 2016 par manque de preuves de l'efficacité du médicament. L'essai clinique de phase II de l'eteplirsén a également montré une bonne efficacité du traitement avec une restauration de la dystrophine. Des essais de phases II et III sont en cours (NCT02420379 et NCT0225552). Sarpta Therapeutics a également déposé une demande de mise sur le marché américain l'année dernière qui est toujours en cours d'étude.

En dehors de l'exon 51, cette stratégie est aussi développée par ces deux entreprises sur les exons 44, 45 et 53<sup>350</sup>.

## B. Les ARNt supprimeurs

Les ARNt supprimeurs dérivent d'ARNt dont l'anticodon a été modifié pour reconnaître un codon stop, permettant ainsi l'incorporation d'un acide aminé à cette position (Figure 75). Étant donné que les ARNt supprimeurs miment une interaction correcte codon-anticodon, ils ne sont pas rejetés par le système de proofreading lors de la traduction.

Plusieurs études ont démontré que les ARNt supprimeurs peuvent potentiellement être utilisés en thérapie génique humaine pour le traitement des maladies à codons stop. Dès 1982, un ARNt supprimeur humain fonctionnel a été construit avec succès et a été capable de supprimer une mutation non-sens UAG dans le gène de la  $\beta$ -globine associée à la  $\beta$ -thalassémie<sup>351</sup>. Pour cela, l'anticodon de l'ARNt<sup>Lys</sup> (AAA) a été modifié pour reconnaître le codon UAG.



**Figure 76 : La pseudouridylation comme supprimeur de PTC**

La pseudouridylation correspond à l'isomérisation de l'uridine en pseudouridine. Ces nucléotides modifiés sont décodés différemment par les ARNt car les pseudouridine disposent d'un donneur de liaison hydrogène supplémentaire. La présence d'un PTC dans la séquence codante conduit à la synthèse d'une protéine tronquée (gauche). La modification de l'uridine en première position du PTC permet de réexprimer une protéine de pleine taille grâce au décodage du codon stop par certains ARNt spécifiques. *Figure adaptée de Karijolich et al., 2015*

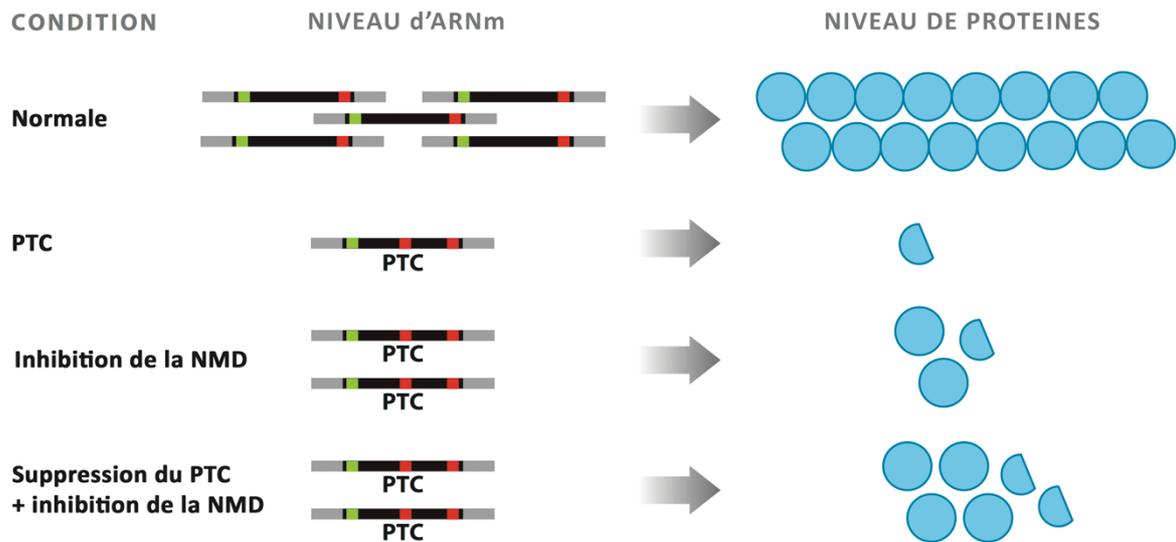
Dans une autre étude, l'ARNt suppresseur (UGA) humain a été transfecté dans des cellules mutantes de xeroderma pigmentosum contenant la mutation non-sens R207X dans le gène XPA. La restauration partielle de l'activité de réparation de l'ADN a été observée, indiquant qu'une réexpression de la protéine XPA a été obtenue <sup>353</sup>.

La première étude *in vivo* utilisant un ARNt suppresseur a été réalisée dans les années 2000 <sup>354</sup>. L'injection directe d'un ARNt suppresseur dans le cœur de souris transgéniques exprimant un gène chloramphénicol acétyltransférase (CAT) contenant une mutation non-sens UAA conduit à la restauration partielle de l'activité de la protéine CAT dans 10% des fibres musculaires. Une approche similaire a été testée chez des souris *mdx* DMD portant la mutation non-sens UAA dans le gène codant la dystrophine. Une semaine après l'injection intramusculaire d'un plasmide exprimant un ARNt suppresseur UAA, la dystrophine pleine taille a été restaurée dans 2,5% des fibres musculaires <sup>355</sup>. Ces résultats suggèrent que les ARNt suppresseurs pourraient être utilisés pour traiter des maladies génétiques humaines provoquées par des mutations non-sens.

Toutefois, cette approche présente plusieurs obstacles qui doivent être surmontés avant de pouvoir être utilisé en thérapie humaine. Une des difficultés majeures de la thérapie génique est d'atteindre la cible cellulaire. L'utilisation d'ARNt suppresseurs complique ce processus, car pour être efficace, ces ARNt doivent être correctement chargés par les aminoacyl synthétases spécifiques et pour cela, ils rentrent en compétition avec les ARNt normaux. De plus, les modifications de la boucle anticodon de l'ARNt permettant la reconnaissance du codon stop peuvent entraîner une baisse de l'efficacité d'aminoacylation et donc de la suppression des codons stop <sup>356</sup>. Enfin, en plus des PTC, les ARNt suppresseurs peuvent reconnaître et décoder les codons stop normaux, ce qui pourrait avoir des effets délétères <sup>357</sup>. Cette stratégie nécessite encore des améliorations pour pouvoir être utilisée dans une approche thérapeutique <sup>358</sup>.

### C. La pseudouridylation

La pseudouridylation chez les mammifères est réalisée par des ribonucléoprotéines (RNP) à boîte H/ACA. Ces RNP comprennent une famille de petits ARN nucléolaires structurellement conservés (snoRNAs) qui permettent de cibler un ARN substrat spécifique. Leur utilisation offre donc, en théorie, la possibilité de pseudouridyler n'importe quel résidu uridine cellulaire <sup>359</sup>.



**Figure 77 - Inhibition de la NMD**

La voie NMD est une voie de surveillance qui reconnaît et dégrade les ARNm aberrants, dont ceux qui contiennent un PTC (Premature Termination Codon). En condition normale, les ARNm ne sont pas dégradés par la voie NMD et sont traduits en protéines. La présence d'un PTC dans la phase codante conduit à la dégradation de l'ARNm par la voie NMD, le niveau de protéines synthétisées est donc particulièrement faible. Certaines molécules sont capables d'inhiber partiellement cette voie, comme le NMDI1 et NMDI14, ce qui permet d'augmenter le niveau d'ARNm et donc de protéines. Cette stratégie thérapeutique est souvent associée à la suppression du codon stop prématuré pour augmenter le niveau de protéines complètes réexprimées. *Figure adaptée de Keeling et al., 2014* <sup>358</sup>

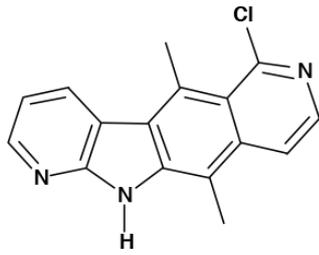
Une étude récente a confirmé cette hypothèse *in vitro* où la translecture d'un gène rapporteur contenant un codon stop  $\Psi$  a généré entre 74 et 100% de protéine rapportrice (Figure 76). En absence de  $\Psi$ , seulement 1% de protéine sauvage a été synthétisé<sup>279</sup>. Un autre ARN guide H/ACA a également été testé dans des cellules de levure en ciblant un PTC dans le gène *CUP1*, conférant la résistance au cuivre. Alors que des cellules exprimant un ARN guide H/ACA contrôle ont été incapables de croître sur du milieu complétement en cuivre, l'expression d'un ARN à boîte H/ACA ciblant le PTC du transcrite *CUP1* muté a permis de restaurer la croissance cellulaire<sup>279</sup>.

Bien que la pseudouridylation de sites spécifiques représente une nouvelle stratégie de suppression des PTC, elle doit également faire face à de nombreux obstacles pour être cliniquement pertinente. Comme pour les ARNt supprimeurs, l'introduction de ce système dans les cellules de telles sortes que leur fonction ne soit ni modifiée ni entravée est un défi majeur qui n'a pas encore été résolu<sup>288</sup>. Ces ARN guides construits artificiellement pourraient aussi cibler les mauvais ARN et entraîner l'introduction de  $\Psi$  à d'autres sites. Davantage d'études seront nécessaires pour déterminer si cette stratégie peut être développée dans une approche thérapeutique.

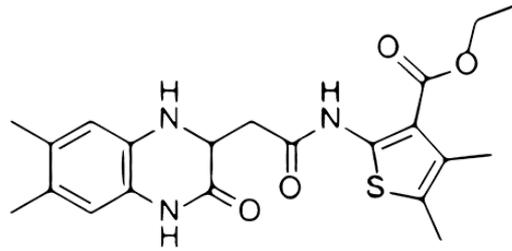
#### D. L'inhibition de la voie NMD

La voie NMD est une voie de surveillance très conservée chez les eucaryotes qui reconnaît et dégrade les ARNm contenant des PTC. L'efficacité de cette voie peut varier d'un individu à un autre<sup>360,361</sup>. Ces différences modulent la sévérité clinique de nombreuses maladies à codon stop<sup>362</sup>, en modifiant probablement l'abondance des transcrits aberrants et donc le niveau de protéines tronquées et de protéines complètes issues de la translecture basale de ces transcrits (Figure 77). En outre, cette variabilité influence également l'efficacité de la thérapie de suppression de codons stop prématurés. Les patients qui expriment des niveaux plus élevés d'ARNm contenant un PTC, du fait d'une moins bonne efficacité de la voie NMD, répondent mieux à ce type de thérapie<sup>363,364</sup>.

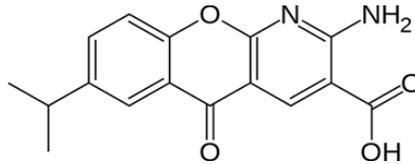
La voie NMD régule l'abondance des ARNm contenant un PTC et représente à ce titre une cible thérapeutique potentielle pour les maladies causées par des mutations non-sens. L'identification de plusieurs cibles thérapeutiques parmi l'ensemble des facteurs intervenants dans cette voie, permettrait de moduler la quantité de transcrits aberrants, et pourrait atténuer certains phénotypes. Par exemple, dans des fibroblastes issus de patients atteints de dystrophie musculaire congénitale d'Ullrich, l'inhibition de la kinase SMG-1 a pour conséquence d'augmenter la quantité de transcrits *COL  $\alpha$ -2 (VI)* et donc de protéines tronquées *COL VI* partiellement fonctionnelles<sup>365</sup>.



NMDI1



NMDI 14



Amlexanox

**Figure 78 - Structure des molécules inhibitrices de la voie de dégradation NMD**

Structure chimique des trois molécules identifiées pour inhiber la voie NMD : NMDI1, NMDI14 et l'amlexanox. NMDI1 et NMDI14 empêchent le clivage des ARNm et l'amlexanox semble bloquer la voie NMD après le recrutement d'UPF1. *Figure adaptée de Durand et al., 2007<sup>366</sup> ; Martin et al., 2014<sup>367</sup> et Gonzalez-Hilarion et al., 2012<sup>368</sup>.*

L'inhibition partielle de la voie NMD augmenterait également la quantité de protéine complète. L'extinction des gènes UPF1 et UPF2 dans des cultures de cellules épithéliales humaines en présence d'un inducteur de translecture (gentamicine) a potentialisé l'effet de cette molécule et augmenté la quantité de protéine CFTR fonctionnelle <sup>360</sup>.

Suite à ces résultats encourageants, plusieurs laboratoires ont recherché des molécules inhibitrices de la voie NMD (Figure 78). La première à avoir été identifiée, NMDI1, agirait en inhibant l'interaction de SMG5 avec UPF1, bloquant ainsi le clivage des ARNm aberrants <sup>366</sup>. Dans une première étude *in vivo*, cette molécule a été coadministrée avec un inducteur de translecture (gentamicine) dans des souris MPS1-H dont le gène *IDUA* est muté par un PTC. Les auteurs ont détecté un niveau élevé d' $\alpha$ -L-iduronidase ainsi qu'une amélioration phénotypique <sup>369</sup>. L'efficacité de cette molécule a aussi été évaluée dans le traitement de cancers où elle pourrait agir de concert avec des agents chimiothérapeutiques <sup>370</sup>. Combiné avec de la doxorubicine, NMDI1 augmente la mort des cellules humaines de cancers du sein (MCF7). La seconde molécule, NMDI14, est issue d'un crible informatique de banques de petites molécules susceptibles de bloquer SMG7. Cette petite molécule inhibe la NMD en empêchant SMG7 d'interagir avec UPF1 et donc de recruter les exonucléases. La potentialisation des inducteurs de translecture a également été observée en traitant des cellules humaines de cancers du sein (HDQ-P1) dont le gène *P53* est muté par un PTC. La forme complète de la protéine p53 étant retrouvée en plus grande quantité lorsque les cellules sont traitées avec du G418 et du NMDI14 <sup>367</sup>. Les résultats obtenus sur le NMDI1 et NMDI14 sont donc très prometteurs, ce qui suggère que l'inhibition partielle de la NMD après le développement embryonnaire serait une approche thérapeutique viable. Son utilisation combinée avec des molécules induisant de la translecture permettrait de traiter plus efficacement les maladies à codon stop.

Toutefois, les molécules NMDI1 et 14 ont un effet global en inhibant la voie NMD sur l'ensemble de ses cibles. Une nouvelle approche a été développée pour ne cibler que le gène contenant le PTC. Cette approche est basée sur l'utilisation d'oligonucléotides anti-sens (ASO), comme dans le saut d'exon, qui empêchent la liaison des EJC à proximité du PTC et stabilise la forme tronquée de l'ARNm <sup>371</sup>. La distance la plus efficace entre le PTC et l'ASO varie entre 24 et 33 nucléotides en aval du PTC selon le gène analysé, *HBB-T39* ( $\beta$ -globine) ou *MECP2-S65X* (protéine de liaison au CpG méthylé) respectivement. Les essais de quantification relative des protéines réalisés sur des cellules U2OS montrent que la synthèse d'ASO augmente la forme tronquée et pleine taille de la  $\beta$ -globine. L'ajout de G418 très concentré ne stimule, de façon surprenante, que faiblement l'expression de la forme pleine taille.

Alors que l'expression de l'ASO dans des cellules U2OS traitées au G418 augmente fortement la forme pleine taille de la protéine.

Bien que l'inhibition de la NMD représente une nouvelle approche de traitement des maladies à codons stop, elle doit être abordée avec prudence en raison des nombreux effets secondaires délétères potentiels. En plus de réguler l'expression des transcrits contenant un PTC, la voie NMD régule l'expression de plus de 10% du transcriptome<sup>145,358</sup>. De plus, les facteurs impliqués dans cette voie de dégradation/régulation jouent un rôle essentiel lors du développement embryonnaire<sup>372,373</sup>. En accord avec ces résultats qui mettent en évidence une absence d'effets indésirables, l'amlexanox, un médicament ayant obtenu une autorisation de mise sur le marché pour le traitement d'ulcères aphteux a récemment été décrit pour inhiber la NMD et induire de la translecture<sup>368</sup> (Figure 78). Il est surprenant d'avoir découvert, lors d'un crible de composés inhibiteurs de la voie NMD, une molécule à la fois capable d'inhiber la voie NMD et d'activer la translecture. De plus, les données présentées ne démontrent pas clairement que l'amlexanox induit de la translecture. Bien que son mécanisme d'action ne soit pas encore identifié, l'amlexanox semble bloquer la voie NMD après le recrutement d'UPF1. Aucune cytotoxicité n'a été observée dans les cellules de mammifères traitées avec ce médicament. Depuis, l'amlexanox a été testé spécifiquement pour sa capacité à induire de la translecture dans un système double rapporteur contenant la mutation G542X (CFTR) dans des cellules de rat (FRT) et pour restaurer l'activité de la dystrophine via la mesure de conductance des ions chlorure<sup>374</sup>. Aucune activité de translecture de l'amlexanox n'a pu être ainsi détectée.

### 3. Suppression traductionnelle pharmacologique

Cette stratégie thérapeutique est basée sur la découverte de molécules, regroupées sous le nom d'inducteurs de translecture, permettant à la machinerie traductionnelle de décoder le codon stop et de poursuivre sa traduction en phase et de synthétiser ainsi une protéine de pleine taille.

La suppression des codons stop prématurés a été décrite pour la première fois en 1996 en tant que traitement potentiel de maladies causées par des mutations non-sens dans le gène CFTR<sup>301</sup>. Cette étude a démontré que le G418, antibiotique de la famille des aminoglycosides, était capable de supprimer des mutations non-sens dans le gène CFTR et de restaurer à la fois des niveaux significatifs de la protéine CFTR et la fonctionnalité de la protéine dans les cellules cultivées. Depuis cette découverte, plus d'une centaine d'études ont analysé la suppression pharmacologique des mutations non-sens comme traitement potentiel pour près de 40 maladies différentes<sup>304</sup>.

La première validation *in vivo* de la faisabilité d'induire la translecture dans un but thérapeutique a été apportée grâce à la souris *mdx* en 1999 par l'équipe de L. Sweeney<sup>375</sup>. Leurs travaux ont montré que l'injection sous-cutanée de gentamicine permet la réexpression de 10 à 20 % de dystrophine complète associée à une amélioration de l'état du muscle. Pourtant le taux de translecture en présence de gentamicine de la mutation *mdx*, mesuré grâce aux systèmes rapporteurs en cultures cellulaires, est très faible (~0.06 %). Il y aurait donc une potentialisation *in vivo*, due à une stabilisation du messager qui échapperait au NMD et une accumulation de la protéine. Néanmoins, l'effet de la gentamicine sur la souris *mdx* n'a pu être reproduit par d'autres équipes<sup>376</sup>. Ces différences ont été expliquées par plusieurs paramètres tels que l'utilisation de différents lots de gentamicine et d'une variation inter-animale dans la façon de métaboliser l'antibiotique.

Néanmoins, cette stratégie comporte plusieurs limitations qui devront être surmontées avant d'envisager un traitement clinique des maladies à codons stop<sup>377</sup>. Une des préoccupations majeures concerne la possibilité qu'une telle stimulation de translecture puisse impacter la terminaison classique de la traduction, conduisant à des dysfonctionnements cellulaires. Toutefois, ni les études *in vivo* sur des modèles animaux ni les essais cliniques sur l'homme n'ont pu mettre en évidence une telle toxicité malgré l'augmentation globale de la translecture<sup>238</sup>. Néanmoins, ces résultats ne démontrent pas le caractère inoffensif de cette stratégie lors d'un traitement clinique à long terme. La seconde limitation concerne l'efficacité de suppression des PTC. En effet, celle-ci est fortement influencée par l'identité du codon stop et par son contexte et varie selon l'inducteur de translecture utilisé. Il est donc nécessaire de déterminer le type de mutation responsable de l'apparition de la maladie. De plus, étant donné l'hétérogénéité génétique des patients, il est indispensable d'évaluer l'efficacité de chaque molécule sur chaque mutation pour estimer les chances de succès de cette thérapie. Enfin, ce type de stratégie nécessitant un traitement clinique à long terme, la toxicité des inducteurs de translecture doit être compatible avec ce type de traitement.

## A. Les éléments de l'ARNm impactant la translecture

### a. La nature du codon stop et son contexte

La capacité des inducteurs de translecture à supprimer divers PTC peut être influencée par un certain nombre de facteurs. En effet, la nature du codon stop et son contexte environnant jouent un rôle tout aussi important dans l'efficacité de la terminaison de la traduction et donc dans la suppression induite par un médicament.

CODON STOP						
UGA	>	UAG	>	UAA	INDUCTEUR DE TRANSLECTURE	REFERENCE
C>U>A>G		C>U~G~C		C>U>G~A	G418	Howard et al., 2000
C>U>A~G		C>U~G~C		C~U~G~C	Gentamicine	
C>U~A>G		C~U~G~C		C~U~G~C	Paramomycine	
C>A~G>U		U>C~G>A		C>U>G>A	G418	Manuvakhova et al., 2000
C>A=G>U		U>C=G>A		C>G>A=U	Gentamicine	
C>A>G>U		C>A=G=U		G=U>A=C	Paramomycine	
C>A=G=U		G>C=U>A		U>C>A=G	Néomycine	
C>A>G>U		U>C>A=G		U>A=C=G	Sisomycine	
C>U>A=G		G>C>A>U		A>U>C>G	Lividomycine	
C>A~G~U		C>A~G~U		C>A~G~U	Gentamicine	Floquet et al., 2012

**Tableau 2 : Influence de la nature du nucléotide +4 sur la translecture induite**

Le codon stop correspond aux bases +1, +2 et +3. Les résultats des travaux obtenus sur l'influence de la nature du nucléotide +4 sur la translecture sont présentés dans ce tableau. Les inducteurs de translecture appartiennent tous à la famille des aminoglycosides. Les nucléotides sont classés par ordre d'influence sur la translecture pour chacun des trois codons stop et selon l'inducteur de translecture utilisé. Figure adaptée de Dabrowski et al., 2015<sup>246</sup>

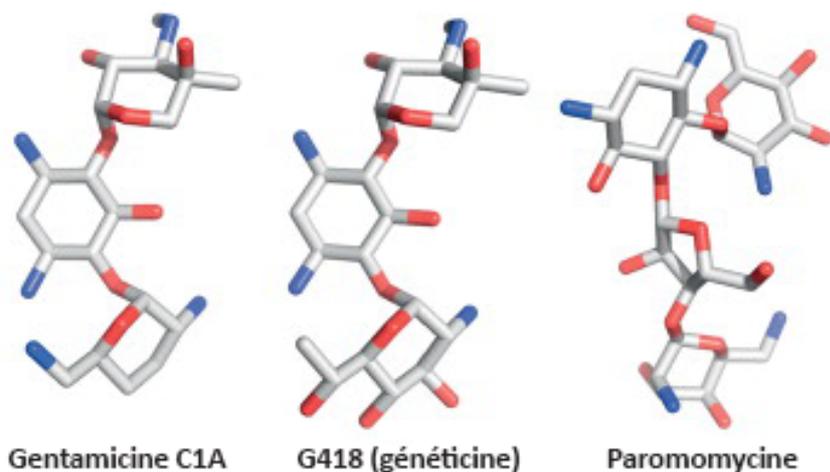
En particulier, notre équipe a montré que la réponse à la gentamicine varie considérablement d'une mutation non-sens à l'autre<sup>378</sup>. De plus, il a été montré que le taux de translecture basal (en absence d'antibiotique) n'était pas corrélé à la réponse de la mutation au traitement par la gentamicine<sup>240</sup>. Ceci indique que les règles de décodage qui régissent le taux de translecture sont différentes en absence ou en présence de gentamicine.

La compréhension actuelle de ces règles ne permet pas de prédire sans ambiguïté l'influence de la nature du codon stop et de son contexte en présence d'un inducteur de translecture. Néanmoins, plusieurs études sur ce sujet indiquent unanimement que UGA est le codon de terminaison associé au plus fort taux de translecture, suivi du codon UAG et UAA<sup>239,240,378,379</sup>.

L'effet du contexte est en revanche plus controversé. En effet, excepté la présence d'une cytosine en position +4 qui est unanimement décrite comme corrélée avec un fort niveau de translecture en présence d'aminoglycosides, aucun consensus n'a été trouvé concernant l'effet des résidus sur les autres positions en amont et en aval du codon stop<sup>246</sup> (Tableau 2). L'analyse statistique du potentiel de translecture de 66 séquences contenant une variété de PTC dans des cellules de souris (NIH3T3), indique que la séquence U-UGA-C a le plus fort potentiel de translecture induit<sup>240</sup>. Plusieurs études indiquent que les nucléotides en positions -6, -5, -1, +4, +8, et +9 influencent aussi l'efficacité de translecture après le traitement à la gentamicine<sup>332</sup>. Par exemple, une guanosine en position +8 et une uridine en +9 ont été rapportées comme étant associées au plus haut niveau de translecture induit<sup>240,380</sup>. Cependant, des résultats controversés ont été obtenus pour la position -5 : alors que Matalonga et ses collaborateurs ont décrit une uridine et une guanidine comme étant les nucléotides les plus appropriés pour une translecture efficace, notre équipe a identifié une adénosine et une cytidine. Les difficultés rencontrées pour parvenir à identifier les règles qui régissent la translecture confirment l'impossibilité de prédire avec certitude l'efficacité de translecture en fonction du contexte nucléotidique.

#### *b. L'acide aminé incorporé*

L'identité de l'acide aminé incorporé au niveau du codon stop prématuré est un élément clé, car il doit être compatible avec la stabilité et l'activité de la protéine réexprimée par translecture. Dans une étude précédente, il a été montré que ni le traitement avec la négamycine, un antibiotique non-aminoglycoside, ni le traitement avec la gentamicine n'ont abouti à la réexpression de la chaîne  $\alpha 2$  laminine complète, malgré un effet important sur la translecture<sup>381</sup>.



**Figure 79 : Principaux inducteurs de translecture de la famille des aminoglycosides**

Structure chimique de la gentamicine (PDB : LLL), généticine (PDB : GET) et paromomycine (PDB : PAR) représentés sous forme de bâtonnets, sans les hydrogènes. Les atomes d'azote, oxygène et carbone sont représentés respectivement en bleu, rouge et blanc. *Figure adaptée de Bidou et al., 2012*<sup>222</sup>

Les auteurs supposent que l'acide-aminé incorporé pourrait être délétère, la protéine serait alors instable et non détectable. Les ARNt proche-cognat capables de reconnaître les codons stop chez la levure ont récemment pu être identifiés par notre équipe <sup>245</sup>. Cependant, l'influence des inducteurs de translecture sur la nature de l'acide aminé incorporé au niveau d'un codon stop est peu connue. L'unique étude ayant analysé les ARNt incorporés chez la levure en présence et en absence de gentamicine a montré que ce sont les mêmes ARNt proche-cognat qui sont incorporés, mais que la proportion de chacun d'entre eux est faiblement modifiée. Ainsi, le codon UGA est décodé par le tryptophane, la cystéine ou l'arginine, alors que les codons UAG et UAA sont décodés par la glutamine, la tyrosine ou la lysine en présence de gentamicine <sup>247</sup>. La gentamicine ne modifierait donc pas la nature de l'ARNt proche-cognat incorporé.

## B. Les molécules inductrices de translecture

### a. Les molécules ayant un effet avéré

#### ★ Les aminoglycosides

Les antibiotiques de la famille des aminoglycosides sont la première classe de molécules identifiées comme inducteur de translecture. Toutefois, tous les aminoglycosides ne possèdent pas cette caractéristique. D'une manière générale, les aminoglycosides contenant un groupement 6'-OH sur leur premier cycle, comme la paromomycine et le G418, ont une affinité supérieure pour le centre de décodage du ribosome que les molécules ayant un groupement NH<sub>2</sub> à cette position (néomycine) <sup>239,382</sup>.

Dés 1985, la translecture d'une mutation non-sens dans le gène de la chloramphénicol acétyltransférase (CAT) a été réalisée avec la paromomycine et la généticine (G418). Depuis, la gentamicine est l'aminoglycoside le plus couramment utilisé pour induire la translecture dans les modèles de maladies (Figure 79). En 1999, la première expérience *in vivo* dans la souris *mdx* pour la DMD démontrait l'efficacité de la gentamicine pour réexprimer une protéine complète <sup>375</sup>. Ces résultats ont servi de stimulus pour plusieurs essais cliniques au cours desquels la gentamicine a été injectée par voie intraveineuse chez des patients atteints de DMD et de BMD. Un premier essai clinique de 2 semaines avec injection journalière a rapporté une absence de bénéfices cliniques (NCT00005574) <sup>383</sup>. En effet, aucune réexpression de la dystrophine n'a pu être détectée après traitement. Deux autres études réalisées avec un temps de traitement allongé à 2 fois 6 jours d'injection intraveineuse de gentamicine séparé par un intervalle de 7 semaines ont rapporté une augmentation de l'expression de la dystrophine <sup>384</sup>.

Maladie	Molécule	Gène cible	Mutation	Codon stop	Modèle	Système	Réponse	Référence
DMD	Gentamicine	<i>DMD</i>	mdx (3185C->T)	UAA A	Ex vivo In vivo	myotubes primaires Souris <i>mdx</i>	Oui Oui	Barton-Davis et al., 1999
Mucoviscidose	Gentamicine	<i>CFTR</i>	G542X	UGA G	In vitro	cDNA dans les cellules HeLa	Oui	Howard et al., 1996
			R553X	UGA G			Oui	
			ΔF508/W1282X	UGA A	Ex vivo	Cellules IB3-1	Oui	Bedwell et al., 1997
			G542X	UGA G	In vivo	Souris <i>Cftr</i> <sup>-/-</sup>	Oui	Du et al., 2002
Cancer	Gentamicine Amikacine Tobramycine	<i>TP53</i>	R210X	UGA C	In vitro	Cellules REF	Oui	Floquet et al., 2012
	Gentamicine Amikacine		W53X	UAG U	In vitro	pAC99 dans les cellules NIH3T3	Oui	
			Q144X	UAG C			Oui	
			W146X	UGA G			Oui	
			Q192X	UAG C			Oui	
			R196X	UGA G			Oui	
			R213X	UGA C			Oui	
			E298X	UAG C			Oui	
			R306X	UGA G			Oui	
			Q317X	UAG C			Oui	
			Y327X	UAA U			Oui	
	R342X		UGA G	Oui				
	G418		Q192X	UAG C	In vitro	pAC99 dans les cellules H1299	Oui	
			R213X	UGA C			Oui	
	Gentamicine G418 Amikacine G418 Paromomycine Gentamicine G418 Gentamicine Tylosine G418 Gentamicine Paromomycine Tylosine Tylosine		R213X/R213X	UGA C	In vitro	Cellules HDQ-P1	Oui	
Gentamicine G418		R1450X	UGA G	In vitro	Double rapporteur dans les cellules HCT116	Oui		
						Double rapporteur dans les cellules HEK293	Oui	
Gentamicine G418		APC	Cellules SW1417	Oui				
				Tylosine Tylosine	E853X	UAG A	In vivo	Cellules HT-29
Tylosine		L850X	UAG G					
				Niemman-Pick A/B	Gentamicine	<i>SMDP1</i>	G503A/G503A	UAG G
A939C/C939A		UAA G	Non					
Gangliosidose I		Gentamicine	<i>GLB1</i>	C1370T/T791C	UGA A			Oui
				1159delC/G1128A	UGA U			Non
Sanfilippo B	Gentamicine	<i>NAGLU</i>	G503A/C700T	UAG A	Ex vivo	Fibroblastes primaires de patients	Oui	
Fabry	Gentamicine	<i>GLA</i>	G503A/C1696T	UAG A/UAG C			Oui	
Syndrome d'Hurler type I	Gentamicine	<i>IDUA</i>	C679T	UGA A			Non	
			A1205G/G1205A	UAG G			Non	
			A1205G/G1205A	UAG G			Non	
Syndrome d'Hunter type II	Gentamicine	<i>IDS</i>	C1327T	UGA U			Oui	

**Tableau 3 : Synthèse des essais réalisés avec les aminoglycosides**

Ce tableau regroupe l'ensemble des informations abordées dans le manuscrit concernant les études *in vitro*, *in vivo* et *ex vivo* réalisées avec des aminoglycosides inducteurs de translecture.

Enfin, une étude de 6 mois avec une administration hebdomadaire ou de deux fois par semaine de gentamicine a confirmé l'efficacité de cette molécule à induire de la translecture chez les patients atteints de la DMD et a conduit à une faible amélioration clinique. (NCT00451074) <sup>385</sup>.

La gentamicine a également été largement testée pour des cas de mucoviscidose. Les premières études sur des modèles cellulaires ont montré que la gentamicine était capable d'induire la translecture de mutations non-sens dans le gène *CFTR* entraînant la synthèse d'une protéine CFTR complète et fonctionnelle <sup>301,386</sup>. Chez la souris transgénique G542X-hCFTR, la gentamicine pourrait partiellement restaurer l'expression de la protéine humaine (hCFTR) <sup>387</sup>. Les essais cliniques réalisés chez les patients atteints par cette maladie ont testé l'effet de la gentamicine administrée soit par voie nasale, soit par intraveineuse. Les résultats varient considérablement. Dans les études de Wilschanski et ses collaborateurs <sup>364,388</sup> et de Sermet-Gaudelus et ses collaborateurs <sup>389</sup>, le traitement par la gentamicine a été associé à des améliorations cliniques alors qu'aucun bénéfice clinique n'a été observé dans l'étude publiée par Clancy et ses collaborateurs <sup>390</sup>. Une des raisons de l'échec de cette dernière étude pourrait provenir des génotypes des patients qui étaient majoritairement hétérozygotes avec des mutations non-sens différentes. Ces observations montrent bien à quel point il est nécessaire de caractériser en culture cellulaire l'efficacité des inducteurs de translecture sur les différentes mutations avant de s'engager dans un processus thérapeutique.

Très peu d'études décrivent cette stratégie pour les gènes suppresseurs de tumeurs. (Pour revue : Bordeira-Carrico et al., 2012 <sup>302</sup>). Bedwell et ses collaborateurs ont montré que les aminoglycosides peuvent supprimer les mutations non-sens dans le gène *P53* <sup>391</sup>. Toutefois, bien que les résultats de cette étude soient encourageants, une seule mutation a été testée en culture cellulaire et la fonctionnalité de la protéine réexprimée n'a pas été analysée. Dans une autre étude, il a été montré que plusieurs aminoglycosides ont amélioré les symptômes cliniques de la tumorigenèse dans un modèle de souris portant une mutation non-sens dans le gène *APC*, et dans des xénogreffes de cellules cancéreuses humaines avec la tylosine <sup>392</sup>. Plus récemment, une analyse complète de l'effet des aminoglycosides a été réalisée sur une mutation non-sens endogène du gène *P53* dans la lignée cellulaire cancéreuse humaine (HDQ-P1) et sur deux autres mutations fréquemment retrouvées dans les cancers humains. Le traitement par la gentamicine, le G418 et l'amikacine a permis de supprimer ces mutations non-sens et de restaurer l'expression d'une protéine fonctionnelle <sup>327</sup>.

Maladie	Molécule	Gène cible	Essai clinique	Réponse	Résultats	Référence NCT	Ref. bibliographique
BMD	Gentamicine		Phase 1	0/2	Pas d'amélioration clinique	NCT00005574	Wagner et al., 2001
DMD	Gentamicine	DMD	Phase 1	3/4	Pas d'amélioration clinique Reexpression de la dystrophine		Politano et al., 2003
			Phase 1	6/12	Faible amélioration clinique Réexpression de la dystrophine	NCT00451074	Malik et al., 2010
			Phase 1	7/9	Faible amélioration clinique. Amélioration de l'activité du canal chlore		Wilschanski et al., 2000
			Phase 1	4/5	Bénéfice clinique pas clair Amélioration de l'activité du canal chlore		Clancy et al., 2001
Mucoviscidose	Gentamicine	CFTR	Phase 1	17/19	Amélioration clinique variable expliqué par les niveaux de transcrits Réexpression de la dystrophine Amélioration de l'activité du canal chlore		Wilschanski et al., 2003
			Phase 2	6/13	Amélioration clinique Réexpression de la dystrophine chez certains patients Y122X Amélioration de l'activité du canal chlore	NCT00376428	Sermet-Gaudelus et al., 2007
Hémophilie	Gentamicine	FVIII	Phase 1	1/3	Faible amélioration clinique Réexpression de FVIII		James, 2005
		FIX	Phase 1	1/2	Faible amélioration clinique Réexpression de FIX		

**Tableau 4 : Synthèse des essais cliniques réalisés avec les aminoglycosides**

Ce tableau regroupe l'ensemble des informations abordées dans le manuscrit concernant les études cliniques réalisées avec des aminoglycosides inducteurs de translecture. Les informations correspondants aux NCT sont consultables sur le site : <https://clinicaltrials.gov/>.

Dans une étude récente publiée par Matalonga et ses collaborateurs, les fibroblastes de 11 patients présentant un PTC impliqué dans six différentes maladies lysosomales ont été traités avec de la gentamicine<sup>380</sup>. Une augmentation de l'activité enzymatique d'un facteur supérieur à 3 a été observée dans 5 des 11 lignées cellulaires après traitement à la gentamicine. Cependant, bien que cette augmentation ne se soit révélée suffisante pour restaurer la fonction du lysosome que dans les fibroblastes d'un seul patient, elle pourrait suffire dans d'autres maladies lysosomales<sup>380</sup>.

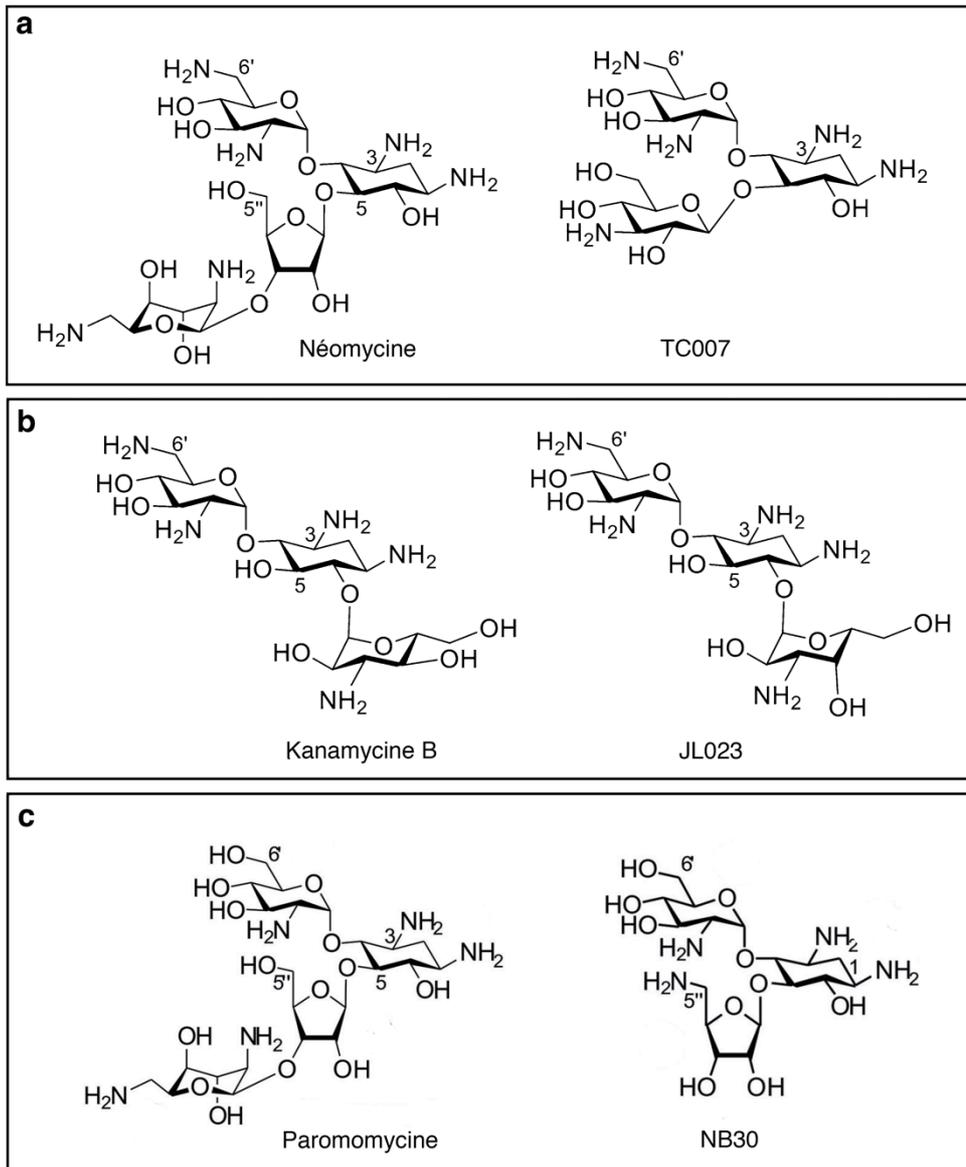
Chez des patients atteints de l'hémophilie, un traitement avec de la gentamicine a montré une faible amélioration clinique de 2 des 5 patients. Il est donc peu probable que la gentamicine soit l'inducteur de translecture le plus approprié pour un traitement de l'hémophilie sévère<sup>340</sup>.

Les études abordées précédemment sont résumées dans les tableaux 3 et 4.

Un grand nombre de travaux a également été réalisé, *in vitro*, *ex vivo* et *in vivo* sur des maladies comme l'ataxie télangiectasie (A-T), le syndrome de Rett (RTT), le syndrome d'Hurler et l'amyotrophie spinale (SMA)<sup>304,332</sup>. La majorité de ces études a conclu que les cellules traitées par des aminoglycosides (Gentamicine, G418, paromomycine et amikacine) ont répondu positivement à ce type de traitement, avec une efficacité dépendant de la mutation analysée, de la molécule, de sa concentration et de la lignée cellulaire utilisée.

Le défaut majeur des aminoglycosides, qui complique leur utilisation clinique, provient de l'étroite gamme de concentration à laquelle ces molécules peuvent être utilisées. En effet, la dose nécessaire pour induire suffisamment de translecture est souvent associée à une importante toxicité, principalement une ototoxicité et une néphrotoxicité<sup>393</sup>. La néphrotoxicité induite par les aminoglycosides est réversible et peut être prise en charge cliniquement avec une thérapie d'hydratation, ce qui permet aux patients de récupérer une fonction rénale normale une fois le traitement interrompu<sup>394</sup>. Ce n'est toutefois pas le cas de leur ototoxicité qui provoque une perte auditive neurosensorielle irréversible<sup>395</sup>. En outre, des études plus récentes ont démontré la présence d'une toxicité rétinienne provoquée par ces antibiotiques, qui pourraient conduire à la perte de la vision<sup>337,396</sup>. Par ailleurs, les aminoglycosides traversent difficilement la barrière intestinale, ce qui contraint à les administrer par voie intraveineuse ou intramusculaire.

Le mécanisme qui induit sélectivement une toxicité au niveau des reins, de l'oreille interne et de la rétine n'est pas encore bien défini à ce jour. Plusieurs études ont suggéré que la mégaline, un récepteur d'endocytose multiligand abondant dans les reins et l'oreille interne provoquerait l'absorption et l'accumulation des aminoglycosides dans ces organes.



**Figure 80 : Les trois groupes de dérivés d'aminoglycosides**

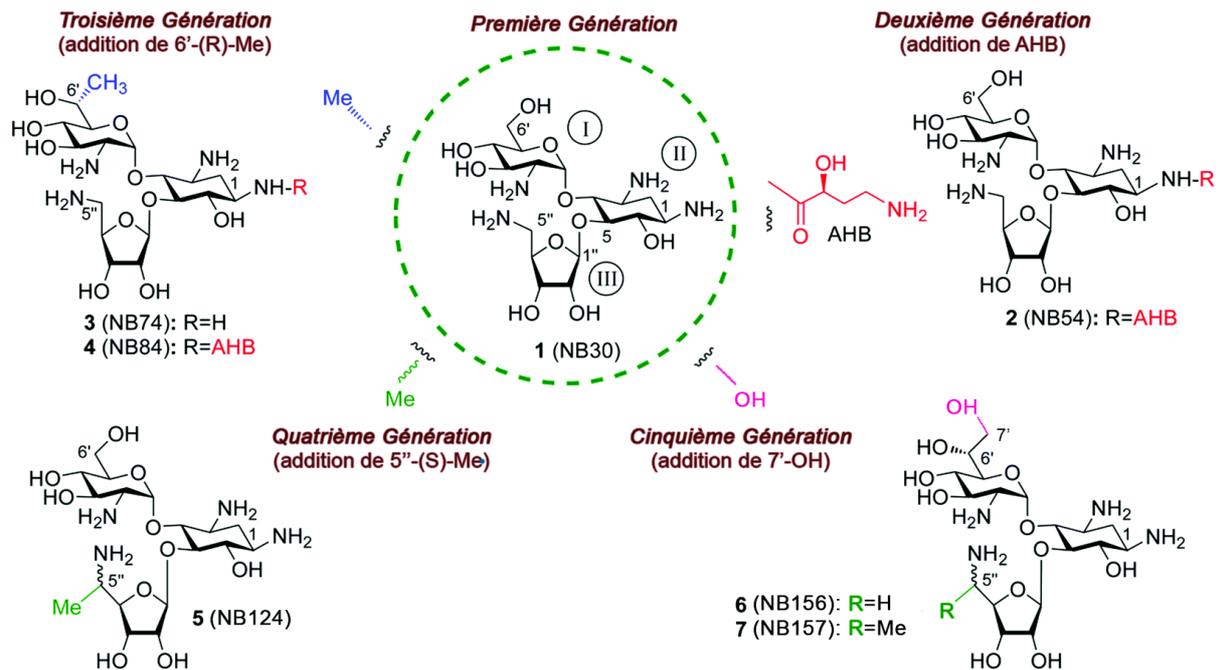
**a** - Structure chimique de la néomycine et de l'un de ses dérivés TC007 (TC). **b** - Structure chimique de la kanamycine B et de l'un de ses dérivés JL023 <sup>97</sup>. **c** - Structure chimique de la paromomycine et de l'un de ses dérivés NB30 (NB). Ces dérivés ont été synthétisés pour améliorer l'efficacité de translecture et diminuer la toxicité des aminoglycosides. La modification de la structure chimique est l'une des trois principales stratégies développées dans les maladies à codons stop. *Figure adaptée de Nagel-Wolfrum et al., 2016* <sup>332</sup>

Une fois à l'intérieur des cellules, leur toxicité serait principalement attribuable à leur charge positive, ce qui les rend sujets à interagir avec divers composants cellulaires chargés négativement tels que les phospholipides ou les phospholipases. Ces interactions seraient responsables de la génération d'espèces réactives à l'oxygène (ROS) et donc à la toxicité cellulaire <sup>395</sup>. En outre, l'effet des aminoglycosides sur la machinerie traductionnelle mitochondriale eucaryote est également considéré comme un facteur important de toxicité. Ce modèle suggère que les aminoglycosides inhiberaient la synthèse des protéines mitochondriales, créant un stress oxydatif et conduisant à la mort cellulaire <sup>222,397</sup>. Les effets secondaires délétères des aminoglycosides empêchant leur utilisation à long terme, cumulés à une efficacité de translecture modérée, montrent à quel point il est important d'identifier d'autres molécules alternatives apparentées ou non aux aminoglycosides ou de potentialiser leur effet.

### ★ Les dérivés d'aminoglycosides

Au cours de ces dernières années, de nombreux travaux de recherche ont été menés dans le but d'atténuer la toxicité associée aux aminoglycosides et d'améliorer l'activité de translecture en synthétisant des analogues des aminoglycosides <sup>304</sup>. Plusieurs stratégies ont été développées dont la modification d'inducteurs de translecture déjà connus. Cette approche est basée sur l'hypothèse selon laquelle les éléments structurels des aminoglycosides responsables de leur toxicité peuvent être séparés de ceux impliqués dans la translecture. Sur la base de cette hypothèse, les éléments structurels qui ont été identifiés pour permettre la liaison des aminoglycosides aux ribosomes cytoplasmiques (améliorant la translecture) tout en réduisant leur affinité pour les ribosomes mitochondriaux (réduction de la toxicité) ont été installés sur un squelette d'aminoglycoside. A ce jour, trois groupes de dérivés d'aminoglycosides ont été synthétisés : ceux basés sur la modification de la néomycine et appelés TC, ceux basés sur la kanamycine B et appelés JL et les dérivés de la paromomycine ou généticine nommés NB (Figure 80) <sup>398</sup>.

Pour étudier l'efficacité des composés TC et JL, leur capacité à induire de la translecture a été analysée dans des fibroblastes de patients atteints d'amyotrophie spinale. Pour l'ensemble des dérivés, le niveau de protéine SMN a augmenté de manière significative et principalement avec TC007 <sup>398</sup>. La structure mère de TC007 ne traversant pas facilement la barrière hémato-encéphalique, TC007 a été administré directement dans le système ventriculaire du cerveau chez des souris exprimant la forme tronquée de la protéine SMN.



**Figure 81 : Molécules dérivées de la paromomycine (NB)**

A ce jour, 5 générations de dérivés ont été synthétisées. Le NB30, molécule de première génération est issue du squelette de la paromomycine. La seconde génération a été synthétisée en ajoutant le groupement AHB sur le cycle II, décrit dans la littérature pour améliorer l'efficacité de translecture et diminuer la toxicité, sur le NB30 pour former le NB54. Puis, l'ajout du 6'-(R)-Me, provenant du G418, sur le cycle I du NB30 et NB54 a conduit respectivement aux composés NB74 et NB84. Le groupement AHB a ensuite été abandonné au profit du groupement 5''-(S)-Me qui a été ajouté sur le cycle III et former la molécule NB124. Enfin, une dernière génération de molécules NB a été synthétisée en ajoutant un 7'-OH afin d'augmenter l'affinité des composés NB156 et NB157 pour le ribosome eucaryote. *Figure adaptée de Shalev et al., 2014*<sup>382</sup>

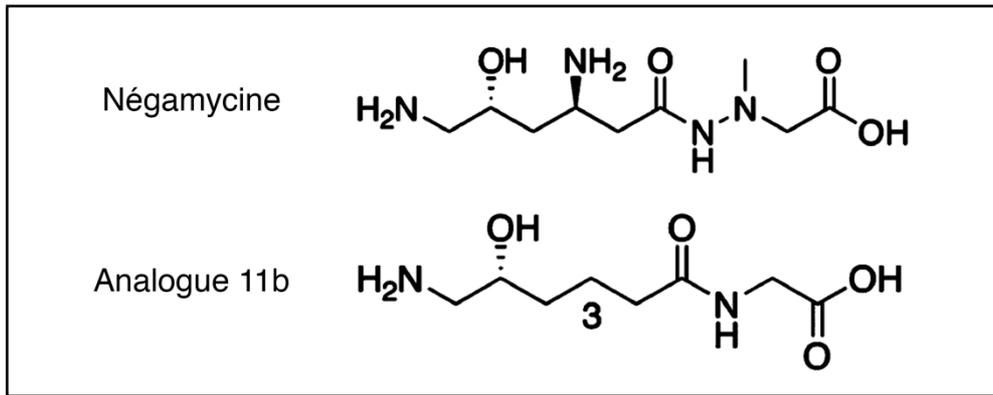
En comparaison avec les souris dont les deux copies sont inactivées et dont la mort survient entre le 4<sup>ème</sup> et 6<sup>ème</sup> jour postnatal, les souris traitées avec TC007 ont survécu jusqu'au 13<sup>ème</sup> jour<sup>399</sup>. De plus, le traitement régulier des souris n'a conduit à aucun effet néfaste sur la durée de vie ou de poids des souris, ce qui démontre une biocompatibilité améliorée par rapport aux autres aminoglycosides classiques tels que le G418<sup>400</sup>. Des injections sous-cutanées ont également été testées, mais aucune amélioration n'a été décelée<sup>401</sup>. Ces résultats indiquent que l'absorption et la distribution de la molécule jusqu'à l'organe cible sont essentielles pour améliorer les résultats cliniques d'une étude. Par conséquent, de nouvelles stratégies doivent aussi être développées pour faciliter la délivrance de la molécule jusqu'à sa cible.

La synthèse d'analogues a également conduit à l'élaboration de dérivés de la paromomycine (Figure 81). La première molécule, appelée NB30, contient un groupement 6''OH sur le premier cycle et un groupement 5'-NH<sub>2</sub>. Ces structures chimiques déterminent l'affinité de la molécule pour le ribosome cytoplasmique plutôt que le mitochondrial<sup>402</sup>. Le NB30 a montré une biocompatibilité élevée et une activité de translecture dose-dépendant, mais son efficacité est inférieure à celle de la paromomycine dont elle est issue et de la gentamicine<sup>403</sup>.

Dans une tentative d'amélioration de l'efficacité du NB30, le groupement (S)-4-aminp-2-hydroxybutanoyl plus communément appelé AHB a été ajouté en position N-1 du NB30 pour former le NB54<sup>404</sup>. L'ajout de ce groupement se justifie pour deux raisons : tout d'abord parce qu'il a été montré que son addition en N-1 sur plusieurs aminoglycosides diminue leur toxicité, et parce que ce groupement serait capable d'augmenter l'efficacité de translecture des aminoglycosides en améliorant leur affinité pour le site A du ribosome cytoplasmique<sup>405,406</sup>. Comme attendu, ce nouveau dérivé est encore moins toxique que le NB30 et est capable de supprimer, plus efficacement que le NB30 et la gentamicine, des mutations non-sens associées à la mucoviscidose, la DMD, et aux syndromes de Rett, d'Usher et d'Urler en culture cellulaire<sup>396,403,407</sup>. Il a également été démontré que le NB54 était capable de supprimer les PTC dans des modèles murins de la mucoviscidose et du syndrome d'Urler, avec plus d'efficacité que la gentamicine<sup>408,409</sup>. Plusieurs études sur les gènes responsables du syndrome d'Usher ont aussi démontré que les aminoglycosides et composés « NB » induisaient de la translecture *in vitro* en culture cellulaire et permettaient de restaurer l'expression et l'activité de la protéine harmonine (*USH1C*). De plus, cette protéine complète a également été retrouvée *in vivo* après injection du NB54 dans des souris transgéniques<sup>396,410</sup>.

Par la suite, le groupement (R)-6'-méthyle a été ajouté sur le cycle glucosamine (cycle I) du NB54 pour donner le dérivé de troisième génération NB84<sup>411</sup>. Parmi l'ensemble des molécules de la famille des aminoglycosides, le G418 est le seul à posséder ce groupement. Le G418 est aussi le plus fort inducteur de translecture de cette famille et également le plus toxique<sup>239</sup>. Malheureusement, il n'existe aucune donnée provenant d'études de relation structure-activité permettant d'identifier quel groupement serait responsable de l'efficacité et de la toxicité. La séparation de ces éléments structurels permettrait pourtant d'améliorer grandement les aminoglycosides. Que ce soit en culture cellulaire ou sur des souris (*Idua*<sup>tm1K<sub>mke</sub></sup>) MPSIH (syndrome d'Hurler), le NB84 supprime les mutations non-sens associées à cette maladie plus efficacement que la gentamicine et que les générations précédentes de dérivés<sup>369,409</sup>. Le traitement des souris a même montré une restauration significative de l'activité de l' $\alpha$ -L-iduronidase. Les résultats obtenus dans les tissus du système nerveux central suggèrent que le NB84 est capable de passer la barrière hémato-encéphalique, indiquant une bonne disponibilité du composé même dans les tissus protégés par des barrières physiologiques. Un traitement à plus long terme a aussi été réalisé et a montré une progression modérée de la maladie et aucun effet toxique lié au traitement<sup>412</sup>.

Une quatrième génération de composés a été synthétisée en ajoutant un groupement 5''-méthyle en configuration (R) ou (S), en présence et en absence du groupement AHB. Le dérivé décrit pour avoir la meilleure efficacité de translecture porte les structures chimiques AHB et 5''(R)-Me. Étonnamment, le dérivé qui a été sélectionné par la suite (NB124) a perdu le groupement AHB sans que la raison n'en soit mentionnée<sup>408</sup>. Les auteurs ont aussi mis en évidence dans ces travaux que l'augmentation de translecture était clairement corrélée à l'augmentation de l'inhibition de la traduction mitochondriale eucaryote, ce qui démontre que la spécificité des aminoglycosides pour le ribosome eucaryote est liée à l'activité de suppression des PTC. Le NB124 améliorerait d'un facteur 10 l'index thérapeutique (ratio efficacité/toxicité) par rapport à la gentamicine dans le modèle de souris de la mucoviscidose (G542X-hCFTR), soit une restauration de 7% de l'activité sauvage du canal chlore<sup>413</sup>. Des études pharmacocinétiques ont démontré que le NB124 atteint une concentration sérique supérieure à la gentamicine peu de temps après son administration, mais sa concentration diminue aussi plus rapidement par la suite. Ce profil pharmacocinétique est favorable, car la translecture peut être limitée par la pénétration de la molécule dans les cellules épithéliales, qui dépend de cette concentration. De plus, le NB124 présente une ototoxicité inférieure à la gentamicine.



**Figure 82 : Structure de la négamycine et de l'un de ses analogues**

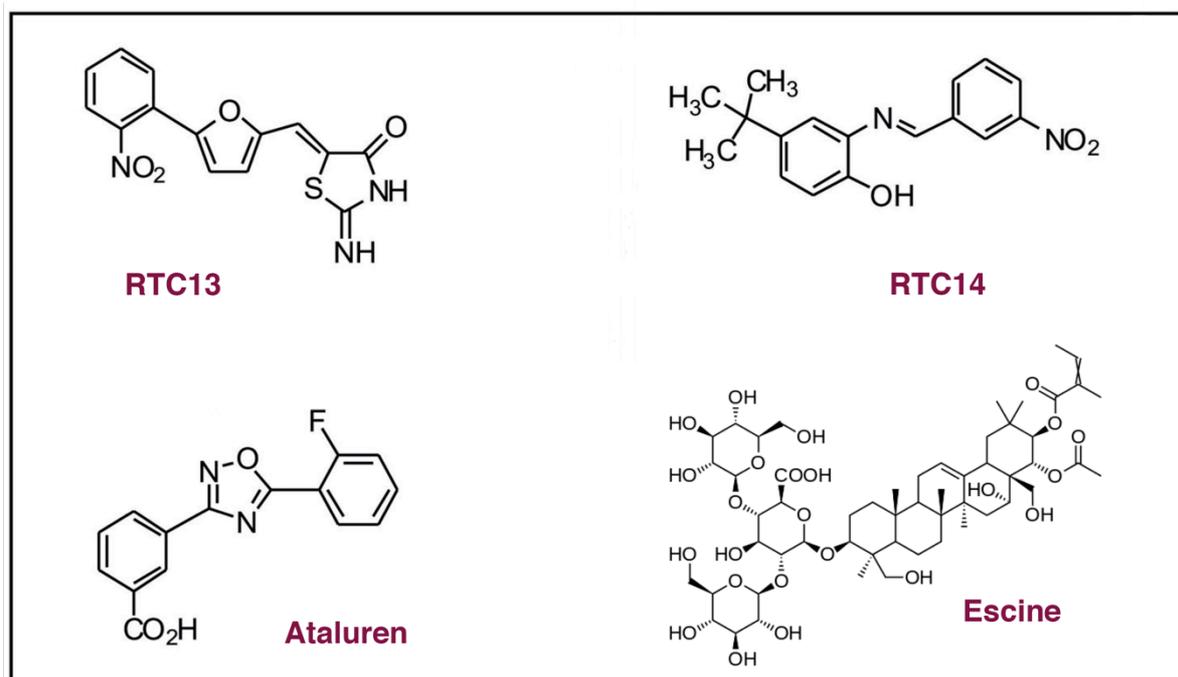
Structure de la négamycine, qui est un inducteur de translecture non apparenté aux aminoglycosides. Plusieurs analogues ont été développés pour améliorer l'efficacité sur la translecture et diminuer la toxicité de la négamycine. L'analogue 11b est la molécule qui répond le mieux à ces deux critères. *Figure adaptée de Taguchi et al., 2012*<sup>414</sup>

En effet, il a aussi été observé que plus la spécificité des aminoglycosides pour le ribosome procaryote était faible, moins la traduction mitochondriale était impactée, ce qui démontre qu'une faible affinité pour le ribosome procaryote est corrélée avec une faible ototoxicité <sup>412</sup>.

Enfin, une dernière génération a été récemment synthétisée en ajoutant au (R)-6'-méthyle des NB74 et NB124 un groupement hydroxyle (7'-OH) sur la chaîne latérale du cycle I, formant ainsi les molécules NB156 et NB157 respectivement <sup>410</sup>. La modification de ce groupement chimique fait suite à l'obtention des structures de plusieurs aminoglycosides dans le ribosome de levure <sup>291</sup>. Dans cette structure, le cycle I du G418 forme peu de liaisons et est faiblement lié au ribosome. Le cycle I du G418 est lié au ribosome par deux liaisons, une première entre le 6'-O du cycle I avec le 1-N du résidu G1645, et une liaison hydrogène entre le 4'-O et le groupement oxygène phosphate des bases A1755-A1756. Le (R)-6'-Me est relativement éloigné du résidu C1646, et l'ajout d'un hydroxyle supplémentaire (7-OH) permettrait de créer une nouvelle liaison avec C1646 tout en conservant la même configuration absolue au niveau de la position 6' (Figure 82) <sup>410</sup>. Testés sur des mutations associées au syndrome d'Hurler et à la mucoviscidose, ces deux nouveaux dérivés induisent davantage de translecture que les molécules dont elles sont issues sans augmenter leur toxicité. Ces deux composés sont actuellement à l'étude et les effets indésirables liés à un traitement prolongé doivent encore être évalués. L'amélioration de la biocompatibilité et de l'activité de translecture des aminoglycosides met en évidence les avantages de cette stratégie.

### ★ La négamycine

L'activité de translecture de la négamycine a été identifiée lors de son injection sous-cutanée dans des souris *mdx* sur une période de 2 semaines. L'analyse du muscle squelettique des souris traitées a révélé la présence d'environ 10% du niveau normal de la dystrophine. Cependant, la restauration de la dystrophine n'a pas été détectée dans les muscles cardiaques. Lorsque la durée du traitement a été prolongée jusqu'à 4 semaines, moins de 10% de la dystrophine normale a été observée <sup>295</sup>. La négamycine a été testée dans une autre expérience *in vivo*, où un système double rapporteur contenant la mutation UGA/319d issue de la dystrophine humaine a été transfecté dans le muscle squelettique de souris C56BL/6. Par la suite, les souris ont reçu une injection de négamycine par voie intramusculaire une fois par jour pendant 3 jours. La translecture a été augmentée d'un facteur 38 suite au traitement avec cette molécule <sup>381</sup>, ce qui démontre le potentiel de translecture de la négamycine pour une utilisation *in vivo*.



**Figure 83 : Molécules non-aminoglycosides dont l'effet sur la translecture est supposé**

Structure chimique de composés, identifiés par criblage haut débit, qui seraient actifs sur la translecture. RTC13 et RTC14 sont issus d'un même crible sur une mutation non sens associée à l'ataxie télangiectasie (A-T). L'ataluren est la molécule la plus connue et est déjà utilisée dans des essais cliniques sur la myopathie de Duchenne <sup>22</sup> et la mucoviscidose <sup>300</sup>. L'escine est une molécule qui vient d'être identifiée lors d'un crible sur des molécules naturelles et qui induit la suppression de mutations non-sens associés à la mucoviscidose. *Figure adaptée de Nagel-Wolfrum et al., 2016 et Mutyam et al., 2016* <sup>332,374</sup>

La négamycine, la gentamicine et l'amikacine ont également été testées sur quatre mutations non-sens du gène *LAMA2* (C1546X, R1326X, Q1437X, Q1240X) impliqué dans la dystrophie musculaire congénitale (CMD). Ces mutations ont été insérées dans un gène rapporteur et ont été testées dans des fibroblastes de souris (NIH3T3). L'efficacité de translecture induite par la négamycine est égale ou supérieure à celle de la gentamicine pour toutes les mutations, en particulier avec la mutation C1546X où elle induit 3% de translecture contre 0,6% et 0,4% pour la gentamicine et l'amikacine respectivement <sup>381</sup>. Malheureusement, aucune de ces drogues ne permet la réexpression de la chaîne  $\alpha 2$  laminine de pleine taille dans des cellules possédant la mutation endogène.

La négamycine semble avoir un effet moins toxique que la gentamicine *in vivo* malgré une activité antibactérienne comparable. L'injection sous-cutanée de doses variables de négamycine dans des souris mdx de 7 semaines montre que le poids corporel des souris à la dose la plus faible augmente continuellement au même titre que les souris non traitées, tandis que le poids corporel diminue pour les souris traitées avec une forte dose sans pour autant être mortelle. A l'inverse, toutes les souris sont mortes dans les 4h après un traitement à forte dose de gentamicine <sup>295</sup>. De plus, l'ototoxicité induite par la gentamicine n'a pas été retrouvée lors du traitement avec la négamycine. Les différences de structure pourraient expliquer cette plus faible toxicité, toutefois, d'autres effets secondaires liés à l'utilisation de la négamycine peuvent exister et devraient être étudiées. L'analyse pharmacocinétique a conclu que cette molécule très polaire possède une faible biodisponibilité, une faible liaison aux protéines présentes dans le plasma (PPB) ce qui lui permet de diffuser et de traverser efficacement les membranes. De plus, elle présente une faible clairance hépatique qui se traduit par de faibles interactions médicamenteuses en cas de co-administration avec une autre molécule <sup>415</sup>.

Dernièrement, Taguchi et ses collaborateurs ont synthétisé plusieurs analogues et ont mesuré leur efficacité de translecture chez des souris READ (Readthrough Evaluation and Assessment by Dual reporter), qui expriment un gène double rapporteur interrompu avec un PTC <sup>414</sup>. Les analogues de la négamycine ont été injectés dans la région abdominale des souris pendant 7 jours. Les auteurs ont montré que seule la molécule 11b induit davantage de translecture que la gentamicine et que la négamycine (Figure 83). L'effet biologique de cette molécule a ensuite été vérifié par des essais *in vivo* d'immunohistochimie mettant en évidence une réexpression de la dystrophine dans les muscles squelettiques des souris mdx. Enfin, l'analogue 11b présente une toxicité plus faible que la négamycine, ce qui permettrait le traitement à long terme de patients atteints de la DMD.

*b. Les molécules ayant un effet supposé*

Au cours de la dernière décennie, les recherches pour identifier des composés nouveaux se sont intensifiées. En effet, les molécules actuelles ne permettent pas d'induire efficacement de la translecture sur l'ensemble des PTC identifiés dans les maladies à codon stop, et leur toxicité est un frein majeur à leur utilisation en thérapie humaine. De plus, leur diffusion à travers les membranes et leur pénétration dans les cellules peuvent altérer de façon importante l'efficacité du traitement.

Dans le cadre de cette stratégie de suppression des codons stop par des molécules chimiques, l'identification de nouveaux composés répondant à ces critères et ayant un spectre d'action différent de celui des antibiotiques connus permettrait de répondre à une demande croissante de traitements de ce type de maladies. Parmi les diverses stratégies qui ont été développées, la plus connue et la plus utilisée est le criblage haut débit<sup>334,416,417</sup>. Cette approche sans a priori vise à tester plusieurs milliers de molécules provenant de chimiothèques (banques de molécules) sur un système rapporteur permettant ainsi l'identification de molécules actives sur la translecture. La cible et le mode d'action des molécules ainsi découvertes ces huit dernières années ne sont toujours pas connus et bien que ces composés semblent très prometteurs, leur efficacité est parfois encore à démontrer.

**★ Les RTCs**

Certains composés non apparentés aux aminoglycosides, appelés petites molécules translectrices (SMRT – Small Molecule ReadThrough) ont été identifiés dans un système utilisant un plasmide contenant le gène *ATM* muté par un PTC couplé à la technique PTT-ELISA (Protein-Truncation-Test Enzyme-Linked-Immunesorbent Assay)<sup>334</sup>. Environ 34 000 composés ont été testés pour leur capacité à induire de la translecture, conduisant à l'identification de 12 molécules de faible poids moléculaire (RTCs – Readthrough Compounds). Les molécules, RTC13 et RTC14 (Figure 82), ont montré une activité de translecture en culture cellulaire sur des mutations non-sens associées à l'ataxie télangiectasie (A-T). Dans les cellules dérivées de patients et traitées par ces deux molécules, les auteurs observent une augmentation significative des niveaux de protéines ATM en comparaison avec des cellules non traitées. En outre, RTC13 restaure l'expression de la dystrophine *in vivo* chez la souris *mdx*<sup>417,418</sup>. Dans une autre étude, l'efficacité de RTC13 a été analysée dans des cellules neuronales générées à partir de cellules souches pluripotentes induites (CSPi) de patients atteints de l'A-T et portant des mutations non-sens. Le traitement de ces cellules avec le RTC13 a restauré la fonction de réparation de l'ADN<sup>419</sup>.

En outre, le RTC13 a induit la translecture sur plus de 20 mutations non-sens rencontrées dans le gène *COL7A1* responsable de différentes maladies<sup>417</sup>. La molécule RTC13 étant moins cytotoxique que RTC14, les efforts se sont concentrés sur le développement d'analogues de RTC13 (Jung et al., 2011). Au cours de l'étude de relation structure-activité, plusieurs analogues ont montré une bonne activité de translecture.

Plus récemment, plusieurs composés de type aminoglycosides et non-aminoglycosides, dont les RTC13 et RTC14 ont été testés dans quatre maladies lysosomales<sup>420</sup>. Dans les fibroblastes de patients atteints du syndrome Sanfilippo, aucune activité enzymatique n'a été retrouvée après traitement avec les RTC alors qu'une stabilisation de l'ARNm a été observée dans les cellules d'un des deux patients. Ces résultats ne remettent pas en cause l'efficacité de ces composés, car les mutations testées n'avaient pas de contextes favorables à la translecture.

D'autres molécules SMRT comme, GJ071 et GJ072 ont été identifiées dans un second criblage de 36 000 molécules avec le même système biologique. Ces composés ont montré des efficacités de translecture comparables, voire meilleures que RTC13 pour des mutations non-sens dans le gène *ATM*<sup>421</sup>.

## ★ PTC124

Au cours d'un programme de recherche de médicaments, 800 000 molécules ont été criblées par l'entreprise PTC Therapeutics sur le gène rapporteur de la luciférase et ont conduit à l'identification du PTC124 (également connu sous le nom d'ataluren ou Translarna™)<sup>416</sup>. L'ataluren (3-[5-(2-fluorophenyl)-[1,2,4]oxadiazole-3-yl]-benzoïque acide) est un composé achiral de 284 Da n'ayant aucune similarité structurale avec aminoglycosides ou d'autres médicaments utilisés en clinique (Figure 83)<sup>422</sup>. Sélectionné pour induire la translecture d'un PTC dans le gène de la luciférase, le PTC124 fut l'objet de nombreuses critiques, car il stabilise la luciférase et augmente ainsi les mesures d'activité de cette enzyme indépendamment de la translecture<sup>423</sup>. Petlz et ses collaborateurs ont répondu que les résultats d'Auld et ses collaborateurs ont été obtenus avec un kit de détection de la luciférase différent de celui utilisé lors du crible et que leurs essais ont été réalisés sur des lignées stables et non des transfections transitoires<sup>424</sup>. Pour autant, alors que la gentamicine induit de la translecture, le PTC124 n'a montré aucune efficacité sur un système rapporteur utilisant la luciférase de renilla<sup>425</sup>. Plusieurs travaux ont également mis en évidence l'inefficacité du PTC124 à induire de la translecture dans différents modèles<sup>84,420,426-428</sup> alors que d'autres ont observé l'inverse<sup>337,380,429-431</sup>. Tant que le mécanisme d'action du PTC124 n'aura pas été découvert, la question de l'efficacité du PTC124 restera ouverte.

Maladie	Molécule	Gène cible	Essai clinique	Réponse	Résultats	Référence NCT	Ref. bibliographique
Mucoviscidose	PTC124	CFTR	2*Phase 1		Pas de translecture non-spécifique, ni de toxicité (2*50mg/kg), ni d'accumulation de la drogue sur 62 personnes volontaires		Hirawat et al., 2007
			Phase 2	0/24	Pas d'amélioration clinique	NCT00234663	Résultats non publiés
			Phase 2a	17/23	Réduction des anomalies épithéliales électrophysiologiques Bénéfices cliniques	NCT00237380	Kerem et al., 2008
			Phase 2	15/30	Augmentation de l'activité du canal chlore Réexpression de la CFTR Bénéfices cliniques	NCT00458341	Sermet-Gaudelus et al., 2010
			Phase 2b	11/18	Amélioration des fonctions pulmonaires Bénéfices cliniques	NCT00351078	Wischanski et al., 2011
			Phase 3	/100	Pas d'amélioration des fonctions pulmonaires Pas de bénéfices cliniques	NCT00803205	Kerem et al., 2014
			Phase 3	?	<i>Essai clinique en cours</i>	NCT02139306	
DMD	PTC124	DMD	Phase 2a	13/36	Pas d'amélioration de la force et des fonctions musculaires Reexpression de la dystrophine dans le muscle (in vivo) Pas de bénéfices cliniques	NCT00264888	Finkel et al., 2013
			Phase 2b	/117	Ralentissement de la progression de la maladie à faible dose. Aucune amélioration à dose plus élevée. Pas de bénéfices cliniques	NCT00592553	Bushby et al., 2014
			Phase 3	/228	Pas d'amélioration au test de 6 minutes de marche Faible amélioration au test de 47m de marche Pas de bénéfices cliniques	NCT01826487	<i>Pas encore publié</i>

**Tableau 5 : Synthèse des essais cliniques réalisés avec le PTC124**

Ce tableau regroupe l'ensemble des informations abordées dans le manuscrit concernant les études cliniques réalisées avec le PTC124 <sup>432</sup>. Les informations correspondants aux NCT sont consultables sur le site : <https://clinicaltrials.gov/>.

A ce jour, plus d'une trentaine d'articles ont été publiés sur l'activité de l'ataluren dans plus d'une vingtaine de modèles *in vitro* et *in vivo* de maladies causées par des mutations non-sens. Toutefois, je ne détaillerai ici que les deux principales : la mucoviscidose et la DMD. Tous les essais cliniques réalisés sur l'ataluren administré par voie orale ont démontré que cette molécule était bien tolérée et n'ont révélé aucun effet indésirable grave liée à la molécule chez les adultes et les enfants <sup>22,432,433</sup>.

L'efficacité de l'ataluren a été analysée dans le traitement de la mucoviscidose. L'administration orale et l'injection intrapéritonéale de l'ataluren ont permis de restaurer partiellement l'expression de la protéine hCFTR chez la souris G542X-hCFTR <sup>429</sup>. En outre, des essais fonctionnels ont démontré chez les souris que le traitement avec l'ataluren restaure environ 27% de la capacité de transport des ions chlorures du canal chlore stimulé par l'AMPc <sup>429</sup>. Au vu de ces résultats et de ceux obtenus lors de l'essai clinique de phase 1 <sup>434</sup>, un premier essai clinique de phase 2 (NCT00234663) a été initié sur 24 patients atteints de mucoviscidose et n'a pas mis en évidence de réponse au traitement (résultats non publiés) <sup>302</sup>. Trois autres essais de phase 2 ont été réalisés par la suite (NCT00237380 ; NCT00458341 ; NCT00351078). L'administration orale de l'ataluren a montré des effets bénéfiques chez ces patients en augmentant significativement l'activité du transport des ions chlorure dans les cellules de l'épithélium respiratoire, et en augmentant l'expression de la protéine CFTR au niveau de la membrane apicale dans les cellules de l'épithélium nasal <sup>363,435,436</sup>. Toutefois, des réponses variables ont été observées notamment, plusieurs groupes de patients qui n'ont pas répondu du tout au PTC124 <sup>436</sup>. Cette variabilité peut être liée à des différences dans le degré d'efficacité de la NMD sur l'ARNm CFTR <sup>437</sup>. En outre, les résultats démontrent la nécessité d'une meilleure caractérisation préclinique de l'efficacité de l'ataluren dans le contexte cellulaire spécifique du patient. Un essai clinique de phase 3 (NCT00803205) a alors été entamé sur plus de 200 patients, dans lequel l'ataluren n'a pas amélioré la fonction pulmonaire des patients atteints de la mucoviscidose <sup>432</sup>. Toutefois, selon les auteurs, ces résultats seraient à pondérer, car certains des patients avaient été traités par de la tobramycine qui inhiberait l'effet du PTC124. Dans le groupe de patients non traités à la tobramycine, une amélioration significative de la fonction pulmonaire serait observée. Un deuxième essai clinique de phase 3 a été relancé, en excluant les patients traités récemment par la tobramycine ou tout autre aminoglycoside. Cet essai clinique est toujours en cours (NCT02139306).

L'ataluren a également été appliqué à des modèles de la DMD tels que des cultures de myotubes de souris mdx et a conduit à la restauration de l'expression de la dystrophine complète. Les souris mdx traitées oralement par de l'ataluren avaient un niveau de dystrophine environ 20% supérieur à celui des souris mdx non traitées.

Une amélioration de la fonction musculaire a également pu être observée et a été corrélée à la diminution du taux de créatine kinase (CK), qui est un marqueur des dommages musculaires <sup>418</sup>. Les résultats obtenus lors d'un essai clinique de phase 2a (NCT00264888) ont montré la réexpression de la dystrophine chez certains patients, mais sans améliorations cliniques <sup>438</sup>. Un essai clinique de phase 2b de l'ataluren a montré lors d'un essai de 6 minutes de marche, un déclin plus faible dans la cohorte traitée avec une faible dose que dans le groupe placebo. Toutefois, aucune amélioration lors de ce test n'a été observée dans la cohorte recevant une dose plus élevée <sup>439</sup>. L'analyse de ces données suggère que la courbe dose-réponse serait inversée ou en forme de cloche <sup>22</sup>. Malgré ces résultats modérés, un essai clinique de phase 3 a été initié (NCT01826487) et a montré qu'aucun des 47 sur 228 patients traités avec le Translarna™ n'a perdu de capacité de marche après 48 semaines contrairement à aux 4 des 52 patients ayant reçu le placebo (Prnewswire – 15/10/2015). PTC Therapeutics a reçu un avis favorable de l'Agence européenne du médicament (EMA) pour une autorisation de mise sur le marché conditionnelle de l'ataluren (Translarna™) en mai 2014. Alors que sur le marché américain, PTC Therapeutics a soumis une demande d'autorisation de mise sur le marché qui lui a été refusée en février dernier, car les essais cliniques de phase 2b et 3 sont négatifs et ne fournissent pas de preuve substantielle de l'efficacité de l'ataluren. L'EMA devra désormais statuer sur la prolongation de l'autorisation provisoire du Translarna™ dans le traitement de la DMD.

Les études cliniques abordées précédemment sont résumées dans le tableau 5.

### ★ Escine

L'escine, vient d'être identifiée lors d'un crible à grande échelle visant à identifier des molécules inductrices de translecture (Figure 83) <sup>374</sup>. Le crible de 1600 molécules ayant déjà obtenu leur autorisation de mise sur le marché a été réalisé via deux systèmes indépendants : un système double rapporteur comportant une mutation non-sens de la CFTR (G542X) stablement intégrée dans les cellules thyroïdiennes de rat (FRT) et un essai de mesure de la conductance des ions chlorides dans des cellules FRT portant la mutation CFTR-G542X ou W182X. Les auteurs ont montré que l'escine est capable d'induire de la translecture, et de restaurer l'expression d'une CFTR fonctionnelle. Cette molécule est constituée de saponines qui est produit naturellement pas les plantes. Initialement utilisée comme anti-inflammatoire, anti-œdémateux et vasoprotecteurs <sup>440</sup> l'escine est facilement disponible et n'a pas d'effets secondaires connus, ce qui en fait un très bon candidat dans le traitement des maladies à codons stop.

L'ensemble des maladies à codons stop identifiées met en évidence le large champ d'application des molécules inductrices de translecture dans le traitement des maladies causées par des mutations non-sens. Toutefois, en dépit d'importants progrès réalisés dans la compréhension du mécanisme de translecture, il n'est pas encore possible de prédire avec certitude quelle sera la sensibilité d'un PTC à un traitement donné. C'est pourquoi il est nécessaire d'analyser en culture cellulaire la capacité de translecture des différentes molécules sur les différents PTC associés à ces maladies. Néanmoins, les stratégies d'identification de nouveaux composés et de modifications d'inducteurs de translecture déjà connus ont largement contribué à l'amélioration de l'efficacité de ce type de thérapie et à l'enrichissement de l'arsenal thérapeutique dans le développement des médecines personnalisées. En dehors des molécules dérivées des aminoglycosides, la caractérisation de ces molécules, y compris leur mode d'action, leur pharmacodynamique, leur pharmacocinétique ainsi que leurs effets physiologiques est souvent limitée. Ces connaissances permettraient pourtant d'accélérer le développement de nouvelles molécules.

## V. Objectifs de thèse

Mes travaux de thèse ont porté sur deux projets centrés sur la suppression traductionnelle des codons stop chez les mammifères. Mes deux projets ont pour objectif de mieux comprendre les mécanismes moléculaires impliqués dans les événements de recodage et de contribuer au développement et à l'amélioration des stratégies thérapeutiques dans le traitement des maladies à codon stop.

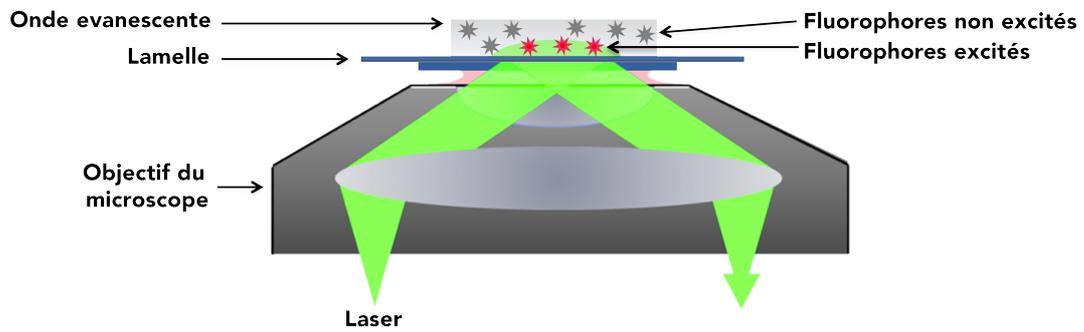
Le premier projet a été réalisé en étroite collaboration avec Nathalie Barbier et Karen Perronet dans l'équipe de Natahalie Westbrook (Laboratoire Charles Fabry UMR8501 - Institut d'Optique). Ce projet concerne la mise au point d'un système de traduction *in vitro* permettant de suivre la traduction eucaryote en particule unique. Mon objectif était à la fois de mettre au point un système de traduction *in vitro* en molécule unique permettant d'obtenir individuellement pour chaque ribosome des informations sur la cinétique de la traduction et d'observer pour la première fois la dynamique d'un événement de recodage en molécule unique. Dans un premier temps, j'ai finalisé la mise au point du protocole expérimental, ce qui m'a permis d'acquérir des données de cinétiques de la traduction. Puis, j'ai analysé la vitesse de traduction et les premières étapes d'élongation de la traduction spécifique d'une initiation sur une entrée interne de ribosome<sup>441</sup>. Ces cinétiques de la traduction m'ont également permis révéler que les premiers cycles d'élongation de la traduction sont ralentis lorsque les ribosomes initient sur l'IRES du virus CrPV (Cricket Paralysis Virus).

Mon deuxième projet consiste en l'identification de nouvelles molécules actives sur la translecture. Pour cela, deux approches ont été réalisées. La première est basée sur la modification de la structure chimique de molécules inductrices de translecture déjà connues afin d'optimiser leur capacité à supprimer les codons stop prématurés et à rendre leur utilisation compatible avec des traitements cliniques. Cette approche est réalisée en collaboration avec le laboratoire de chimie du Dr Timor Baasov (Institut de technologie Technion - Haïffa, Israël). J'ai caractérisé plusieurs molécules de synthèse et analysé leur capacité à induire la translecture. J'ai ensuite déterminé la capacité de la molécule la plus active à permettre la réexpression de la protéine p53 fonctionnelle et à induire l'apoptose d'une lignée tumorale humaine. La seconde approche a consisté à cribler des chimiothèques de molécules et a été réalisée en étroite collaboration avec Goulven Merer et Jean-Christophe Cintrat de l'équipe chimie combinatoire et criblage à haut débit (CEA - Saclay). J'ai tout d'abord criblé plus de 17 000 composés issus de deux banques de molécules. Puis j'ai cherché à éliminer les faux positifs et à quantifier précisément l'efficacité des molécules « hits ». J'ai ensuite caractérisé la molécule la plus prometteuse.

# ***Résultats***

***Traduction en molécule unique***

***Le plus grand risque,  
c'est de n'en prendre aucun  
(Greena Davis)***



**Figure 1 : Microscopie à onde évanescente**

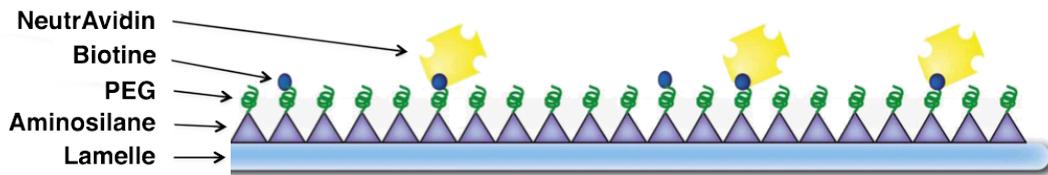
Le faisceau incident du laser est envoyé sur la lamelle de manière à être totalement réfléchi. Seuls les fluorophores présents dans l'onde évanescente, d'environ 100 nm de hauteur, (en vert) sont excités (étoiles rouges). *Figure adaptée de la thèse d'Antoine Le Gall, 2011.*

## Résultats : Traduction en molécule unique

### 1. Résumé de l'article : Kinetics of IRES-mediated eukaryotic using single molecule fluorescence microscopy

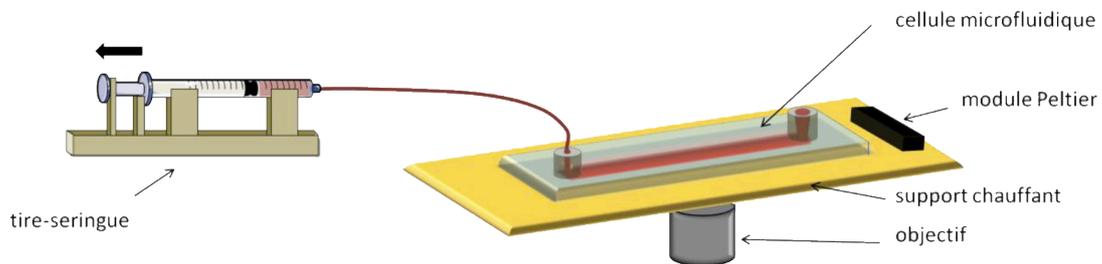
Ces dix dernières années, la cristallographie et la cryo-EM ont permis d'affiner nos connaissances des différentes étapes de la traduction. Ces études structurales ont ainsi permis de mieux comprendre le fonctionnement du ribosome procaryote et eucaryote en révélant l'existence de mouvements de grande amplitude lors de la traduction. Plus récemment, la structure du ribosome eucaryote et de ses deux sous unités (40S et 60S) a été élucidée pour la première fois. Toutefois, ces techniques ne saisissent que des instantanés d'un processus dynamique et ne permettent que difficilement d'appréhender la nature, l'ampleur et la vitesse des mouvements. Les approches en molécule unique sont en revanche parfaitement adaptées à l'étude de ces phénomènes dynamiques. Elles permettent en effet d'étudier des processus asynchrones et de mesurer des événements rares sans en moyenniser le comportement individuel. Ceci a permis des avancées majeures dans la compréhension de la dynamique du ribosome procaryote. La plupart de ces études utilisent un marquage du ribosome procaryote ou des protéines ribosomales telles qu'eS25 sans affecter la capacité d'assemblage du ribosome. Cependant, le marquage du ribosome ou des facteurs d'élongations peut modifier la traduction de nombreuses façons. Une alternative consiste à étudier la traduction en molécule unique par FRET (Forster Resonance Energy Transfer) qui permet de mesurer des variations de distance de 1 à 10 nm entre 2 molécules fluorescentes compatibles, mais encore une fois, il est nécessaire d'attacher un fluorophore sur les ARNt ou sur le ribosome.

L'objectif de ce projet est d'étudier la cinétique d'élongation de la traduction eucaryote en microscopie à fluorescence par réflexion totale interne (TIRF - Figure 1) en molécule unique avec une machinerie traductionnelle de mammifère non modifiée. Pour cela, nous avons développé un système de traduction où l'initiation est réalisée grâce à la présence d'une IRES (Internal Ribosome Entry Site) et où l'ARNm est marqué spécifiquement par des oligonucléotides couplé à un fluorophore. Nous avons ainsi pu mesurer la vitesse d'élongation de la traduction eucaryote d'un ribosome non marqué et mettre en évidence que l'initiation traductionnelle dépendante d'une IRES ralentit fortement la vitesse des premiers cycles d'élongation.



**Figure 2 : Chimie de surface**

Les lamelles sont traitées pour fixer spécifiquement les molécules à étudier. Elles sont recouvertes d'une couche d'aminosilane qui permet l'ancrage d'une couche de PEG. Certaines molécules de PEG sont couplées à une biotine qui interagit fortement avec des protéines NeutrAvidin. Ces protéines possédant 4 sites de fixation pour des biotines, elles pourront aussi accrocher la molécule d'intérêt par l'intermédiaire d'une biotine. *Figure adaptée de la thèse d'Hélène Chommy.*



**Figure 3 : Cellule microfluidique, dispositif chauffant et système d'injection**

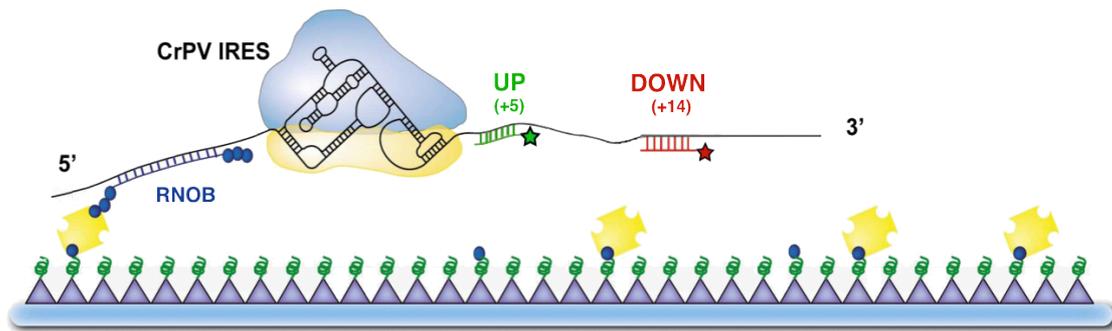
Les échantillons sont déposés dans le réservoir de la cellule microfluidique. Ils sont aspirés dans la chambre d'aspiration grâce au tire-seringue. La cellule microfluidique est thermostatée par le module Peltier qui est fixé à l'extrémité du support en cuivre. L'objectif est également chauffé pour éviter les pertes thermiques de l'échantillon à son contact. *Figure issue de la thèse d'Hélène Chommy.*

Le type de microscopie que nous avons utilisé, dite « à onde évanescente » impose que le système biologique étudié soit fixé à la surface de la lamelle. Une chimie de surface a donc été développée pour fixer spécifiquement les ARNm étudiés (Figure 2). Cette chimie est constituée d'une couche de polyéthylène glycol<sup>302</sup> qui est électriquement neutre et permet de passiver la surface. Les PEG sont couplés à une biotine, sur laquelle se fixe avec une haute affinité la NeutrAvidin. Cette protéine possède 4 sites d'interactions avec les biotines, ce qui permet d'y accrocher l'ARNm via son extrémité 5' couplée à un oligonucléotide biotinylé.

L'accroche de ce système biologique à la surface de la lamelle se fait dans une chambre microfluidique (Figure 3). La traduction démarre lorsque le lysat de réticuloocyte de lapin (RRL) est injecté, apportant ainsi l'ensemble des facteurs nécessaires à la traduction. Les mesures de cinétiques sont obtenues grâce au marquage de l'ARNm par deux oligonucléotides couplés spécifiquement à deux fluorophores distincts (ATTO647N et ATTO565). Lors de la traduction, le ribosome, via son activité hélicase, décroche la première sonde (R-UP) après 5 cycles d'élongation puis la deuxième sonde (G-DOWN) située 9 codons plus loin (Figure 4).

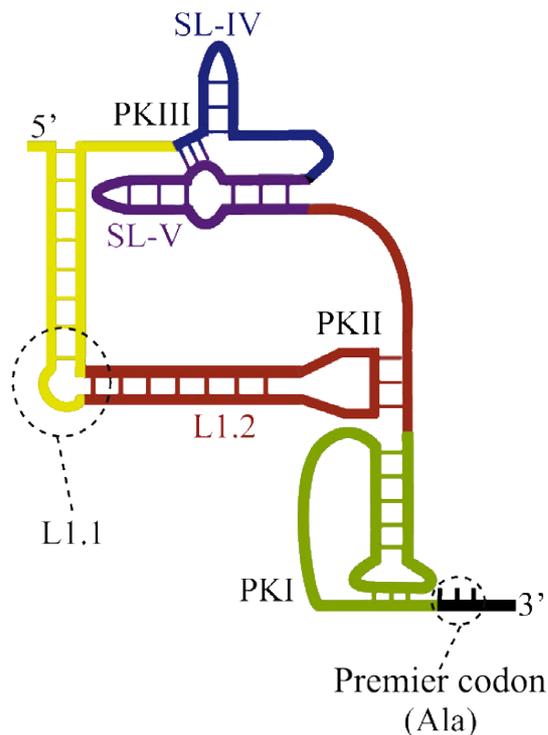
L'étape d'initiation de la traduction eucaryote est particulièrement complexe et fait intervenir plus de 40 facteurs. Plutôt que de reconstituer cette étape *in vitro*, nous avons utilisé l'IRES CrPV (Cricket Paralysis Virus) sur lequel le ribosome 80S purifié initie sa traduction en absence de l'ensemble des facteurs d'initiation et de l'ARNt initiateur (Figure 5). Afin de vérifier que le ribosome est capable de se fixer sur l'IRES en présence des sondes marquées, des essais de retard sur gel ont été effectués et ont confirmés que cette IRES permet la fixation des ribosomes. De plus, la proximité de la séquence de fixation de la première sonde ne semble pas affecter le recrutement des ribosomes.

La cinétique de traduction est réalisée en utilisant un microscope à fluorescence par réflexion totale interne permettant de localiser en temps réel les deux sondes marquées et fixées aux ARNm. Dans nos conditions expérimentales, nous avons pu observer que les deux sondes colocalisent sur 80% des ARNm. Nous avons alors analysé la cinétique de traduction de ces ARNm en quantifiant le départ de la sonde R-UP sur 250 secondes. Les résultats montrent que 50s après injection du RRL,  $51 \pm 7$  % des sondes ont été détachées de l'ARNm. Nous avons vérifiés que ces départs n'étaient ni liés au photoblanchiment des fluorophores, ni à la présence d'autres activités biologiques telles que les RNase en réalisant des contrôles après traduction et sur un ARNm sans IRES. Ces résultats suggèrent donc que les ribosomes sont capables d'initier sur l'IRES et traduire la séquence d'ARNm en détachant la première sonde marquée.



**Figure 4 : Initiation du ribosome sur l'ARNm, accroche à la surface et traduction**

Les ARNm utilisés pour les expériences en molécule unique possèdent une IRES CrPV qui permet le recrutement du ribosome 80S indépendamment de l'ensemble des facteurs d'initiations et de l'ARNt initiateur. Les ARNm sont accrochés par l'intermédiaire de l'oligonucléotide biotinylé RNOB. Les ARNm sont visualisés grâce à deux sondes fluorescentes hybridées en aval de l'IRES (UP et DOWN). La traduction est initiée par l'ajout de lysat de réticulocyte de lapin (RRL). Le ribosome traduit alors l'ARNm et décroche successivement via son activité hélicase les deux sondes fluorescentes. *Figure adaptée de la thèse d'Hélène Chommy.*



**Figure 5 : Structure de l'IRES CrPV**

Structure secondaire de l'IRES de type 4 CrPv (Cricket Paralysis Virus). Cette IRES permet le recrutement du ribosome 80S indépendamment des facteurs d'initiations et de l'ARN initiateur. Elle est composée de 3 pseudonoeuds (PKI, PKII et PKIII) et de deux tige-boucle (SL-IV et SL-V) et deux régions décrites pour ne pas interagir avec la 40S (LA.A et LA.2). Le pseudonoeud I mime l'ARNt initiateur en se logeant au site P du ribosome. *Figure adaptée de Fernandez et al., 2014* <sup>442</sup>.

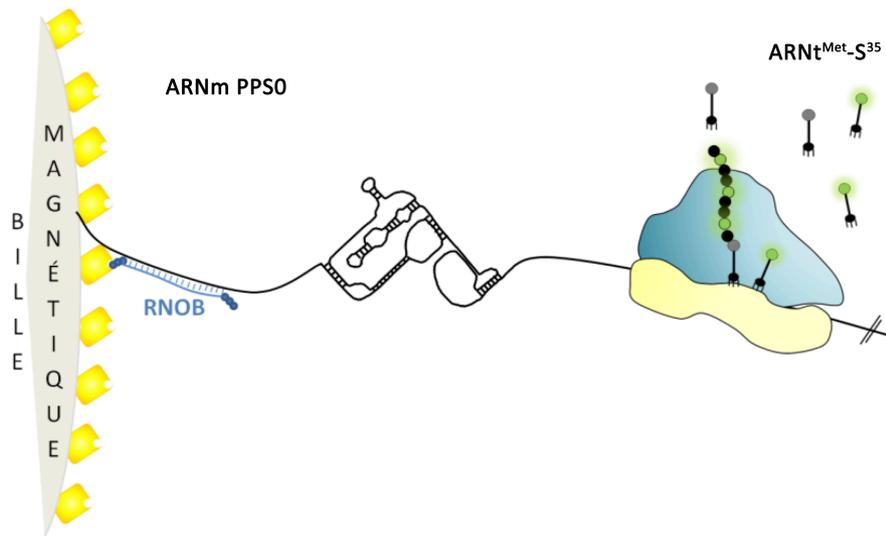
Nous avons ensuite cherché à déterminer la vitesse de traduction du ribosome au cours de l'élongation. Pour cela, nous avons analysé le temps nécessaire aux ribosomes pour traduire la séquence d'ARNm située entre les deux sondes. Ainsi, la différence de temps entre le départ de G-DOWN et de R-UP rapporté à la distance séparant les deux sondes permet de calculer la vitesse de traduction. Ces résultats ont été normalisés par le contrôle sans IRES et nous avons fait correspondre une gaussienne (fit). Nous avons ainsi pu déterminer que  $12,5 \pm 2,1$ s sont nécessaires aux ribosomes pour traduire 9 codons, soit environ  $1,4 \pm 0,2$ s par cycle de translocation.

Sachant que le RRL contient des ribosomes, nous avons cherché à identifier si la traduction observée était réalisée par les ribosomes 80S purifiés et préincubés avec l'ARNm ou par les ribosomes du RRL. La probabilité que la sonde R-UP se détache a été déterminée à partir de traductions réalisées avec et sans ribosomes préincubés. Les résultats montrent que l'efficacité traductionnelle est plus importante lorsque les ribosomes sont préincubés avec l'ARNm (50%) que sans (10%), ce qui confirme que la traduction est réalisée majoritairement par les ribosomes purifiés et non par ceux du RRL.

L'analyse des données de traduction a également permis de mettre en évidence qu'il n'était pas possible de faire correspondre la courbe représentant la probabilité de départs de chacune des sondes dans le temps avec une loi mathématique comprenant un seul paramètre temps. En effet, si l'ensemble des cycles d'élongations étaient identiques, ce paramètre correspondrait au temps caractéristique d'un cycle d'élongation. Or, il est nécessaire d'intégrer un second paramètre temps pour obtenir cette courbe et ces deux paramètres correspondent à un temps court et un temps long. De plus, la distance séparant l'IRES du R-UP a été augmentée à 9 codons et le temps long est resté identique. Ces résultats suggèrent donc que cette première phase longue décrit le temps nécessaire au ribosome pour réaliser ses deux premiers cycles de translocation et se détacher de l'IRES et que le temps dit rapide correspond à la vitesse d'élongation du ribosome une fois détachée de l'IRES.

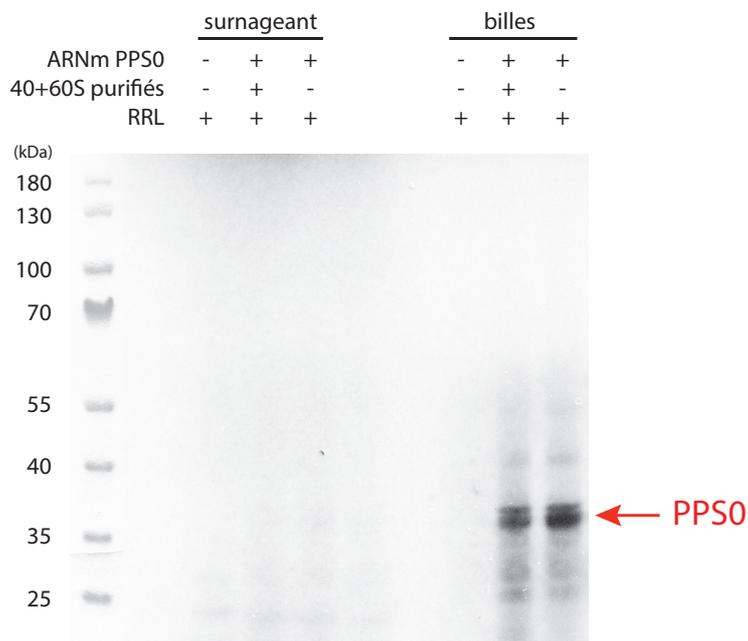
Nous avons donc pu mettre au point un système de traduction eucaryote in vitro en molécule unique sans modifier ni l'ARNm ni les facteurs de traduction où des ribosomes 80S purifiés sont capables d'initier une traduction sur une IRES CrPV et de traduire une séquence d'ARNm tout en détachant des sondes marquées. Grâce à ce système, la vitesse d'élongation des ribosomes purifiés a pu être déterminée et nous avons pu montrer que les deux premiers cycles d'élongation sont plus lents probablement du à l'encombrement de l'IRES dans le ribosome.

**2. Article : Kinetics of IRES-mediated eukaryotic using single molecule fluorescence microscopy**



**Figure 6 : Principe de l'expérience de traduction *in vitro* sur billes magnétiques en mesure d'ensemble**

Les ARNm PPS0 sont fixés grâce à l'oligonucléotide RNOB (bleu) sur des billes magnétiques couvertes de streptavidines (jaune). Les protéines PPS0 synthétisées sont marquées par l'incorporation d'acides aminés Methionine  $S^{35}$ . Comme l'ARNm PPS0 est dépourvu de codon stop, les ribosomes atteignant la fin de l'ARNm ne sont pas dissociés et les peptides ne sont pas relargués. *Figure adaptée de la thèse d'Hélène Chommy.*



**Figure 7 : Traduction *in vitro* sur une surface en mesure d'ensemble**

Les protéines des surnageants de traduction et celles associées aux billes magnétiques sont analysées dans un gel d'électrophorèse SDS-PAGE. Les protéines sont ensuite détectées sur un film à autoradiographie. Trois conditions ont été testées : un contrôle négatif sans ARNm PPS0 ; une traduction en présence de ribosome purifiés et préincubés (40S+60S purifiés) sur l'IRES et une traduction où seuls les ribosomes du RRL peuvent traduire (sans 40S+60S purifiés). La protéine PPS0 est indiquée par une flèche rouge.

### 3. Traduction in vitro sur billes magnétiques en mesure d'ensemble

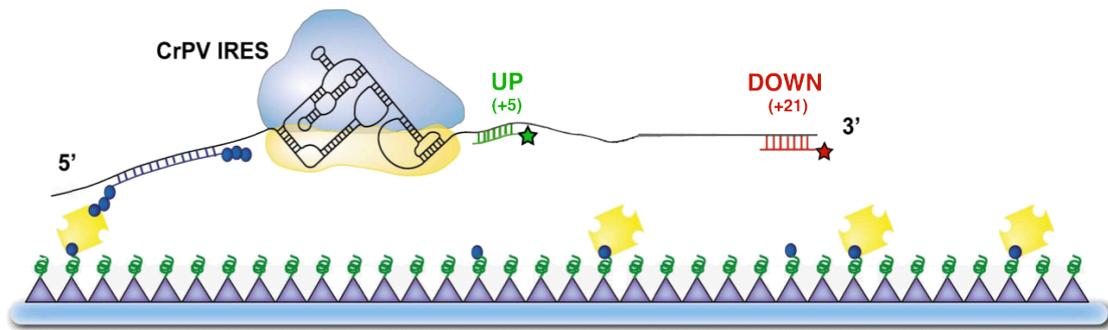
Nous avons ensuite analysé la capacité des ribosomes purifiés et des ribosomes issus du RRL à traduire un ARNm ancré à une surface en initiant sur l'IRES CrPV. L'ARNm utilisé contient la phase codante de PPS0 qui code la polymérase 1 de Influenza Virus. La taille de la PPS0 est de 261 acides aminés dont 11 sont des méthionines. Cet ARNm ne contenant pas de stop, les ribosomes atteignant la fin de l'ARNm ne sont pas dissociés et les peptides ne sont pas relargués. Lors de l'expérience, j'ai testé trois conditions : (1) un témoin négatif ne contenant pas d'ARNm. (2) une condition où les ribosomes purifiés sont préincubés sur les appariements ARNm PPS0/RNOB. Le premier cycle de traduction est donc réalisé par les ribosomes purifiés et préincubés. Toutefois, si les ribosomes présents dans le RRL sont capables d'initier une traduction sur l'IRES, ils pourront traduire les cycles suivants. (3) une troisième condition où les ribosomes sont apportés par le RRL lors de son ajout à la réaction de traduction. Les complexes ARNm PPS0/RNOB/ribosomes ou ARNm PPS0/RNOB sans ribosomes sont ensuite accrochés à des billes magnétiques recouvertes de streptavidines. Tous les facteurs nécessaires à l'élongation sont apportés lors de l'ajout du RRL. Afin de détecter les protéines PPS0 synthétisées, nous avons ajouté de la méthionine  $S^{35}$  (Figure 6). L'ARNm PPS0 ne contient pas de codon stop. En fin de traduction, les billes magnétiques permettent d'isoler les complexes comprenant les ribosomes avec les peptides naissants marqués avec les acides aminés méthionine  $S^{35}$ . Les protéines du surnageant (RRL) et celles associées aux billes sont analysées par SDS-PAGE. Le gel est ensuite séché et déposé dans une cassette sur un Hyperfilm avant d'être révélé.

Dans les puits correspondants aux surnageants (Figure 7, puits 1 à 3), aucune bande n'est observée, ce qui est attendu, car les protéines sont associées aux ribosomes accrochés aux billes. En revanche, dans les puits correspondants aux billes (Figure 2, puits 4 à 6), une protéine autour de 37kDa est détectée et correspond à la protéine PPS0. L'intensité de la bande est légèrement plus importante dans le puits correspondant à la traduction par les ribosomes du RRL. Ces derniers sont donc bien capables d'initier sur une IRES et traduire un ARNm accroché à une surface dans le RRL.

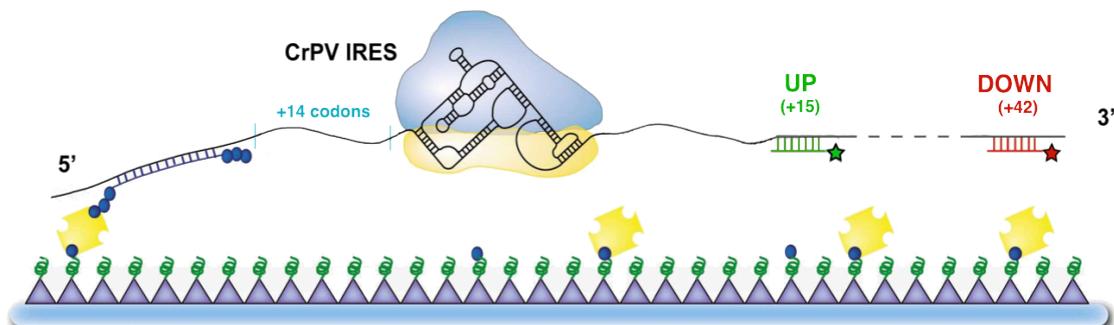
### 4. Acquisition de cinétiques de traduction complémentaires

Nous avons souhaité dans un premier temps déterminer si les seuls ribosomes apportés par le RRL sont capables de traduire l'ARNm en molécule unique. Comme le premier ARNm conçu pour les expériences de traduction en molécule unique, cet ARNm n'est pas coiffé.

### A - Condition 1



### B - Condition 2



**Figure 8 : Les différentes conditions analysées**

La traduction des ribosomes apportés par le RRL a été étudiée sur deux séquences d'ARNm schématisés ici. **A** – L'ARNm est fixé à la surface de la lamelle par l'intermédiaire d'un oligonucléotide biotinylé, appelé RNOB (bleu). Le ribosome initie sa traduction sur l'IRES CrPV et la cinétique de traduction est observée par les départs des sondes UP et DOWN, situées respectivement à 5 et 21 codons de l'IRES. **B** – La distance séparant l'IRES du site d'hybridation du RNOB a été augmentée de 14 codons. Les sites d'hybridation des sondes UP et DOWN ont également été éloignés de l'IRES et l'un de l'autre. Ils sont situés respectivement à 15 et 42 codons de l'IRES.

	Condition 1		Condition 2	
	Ribosomes du RRL		Ribosomes du RRL	
Position des sondes	UP (+5)	DOWN (+21)	UP (+15)	DOWN (+42)
Proportion d'ARNm traduits (%)	11 ± 1	8 ± 1	41	60
Vitesse d'élongation (s/codons)	0,5 ± 0,2	0,3 ± 0,1	0,5	0,2
Temps long (s)	19 ± 5	19 ± 7	25	12
Nombre de points	1100	966	1235	1178

**Figure 9 : Cinétique de traduction dans les 2 conditions testées**

Ce tableau regroupe les résultats de l'analyse des cinétiques de traduction des ribosomes du RRL pour les deux fluorophores (UP et DOWN) et dans les deux conditions testées. La distance séparant les sondes fluorescentes de l'IRES est indiquée entre parenthèse en nucléotides. Plusieurs paramètres ont été évalués : la proportion d'ARNm traduits, la vitesse d'élongation et le temps long. Les résultats de la condition 2 sont des données préliminaires dont la variabilité n'a pas encore pu être calculé. Le nombre de point utilisé pour l'analyse est précisé.

Il possède, de l'extrémité 5' vers l'extrémité 3', une séquence d'appariement pour l'oligonucléotide d'accroche RNOB, une IRES CrPV et un site d'accroche pour l'oligonucléotide fluorescent UP et un pour l'oligonucléotide fluorescent DOWN (Figure 3). De plus, l'ARNm est dépourvu de codon stop dans la phase de lecture définie par le codon d'initiation de l'IRES. Les essais préliminaires réalisés sur l'ARNm présenté dans l'article semblaient montrer que la traduction était plus rapide lorsqu'elle était réalisée par les ribosomes apportés par le RRL. Par conséquent, j'ai modifié la séquence d'ARNm précédemment construite de telle sorte que le site d'hybridation de la sonde DOWN soit situé à une distance de 21 codons de l'IRES (Figure 8-A), contre 14 codons dans nos essais précédents.

Cet ARNm est hybridé avec les oligonucléotides RNOB et les sondes UP et DOWN. Le complexe est incubé 10' à 37°C pour décrocher les sondes mal hybridées puis il est injecté dans la cellule microfluidique pour être fixé à la lamelle. La chambre est ensuite rincée pour enlever les molécules non accrochées à la surface et les fluorophores en excès. La traduction est déclenchée par l'injection de RRL juste après avoir lancé l'acquisition des images. La réaction est réalisée à 30°C et une image est prise toute les 5s durant 250s puis les cinétiques de traduction sont analysées. Les résultats obtenus dans cet essai (Figure 9 – Condition 1) montrent qu'environ 10% des ARNm sont traduits par les ribosomes apportés par le RRL. De plus, leur vitesse de traduction est très rapide, entre 0,3 et 0,5s par codon et le temps long est de 19s. La proportion d'ARNm traduit par ces ribosomes reste faible malgré l'absence des ribosomes purifiés et préincubés. En revanche, les ribosomes du RRL traduisent bien plus rapidement l'ARNm que ceux purifiés.

Nous avons alors cherché à savoir si la proximité de l'IRES de la surface de la lamelle pouvait provoquer une gêne stérique et être responsable de la faible proportion d'ARNm traduits. J'ai donc allongé la distance entre l'IRES et le site d'accroche de l'ARNm à la lamelle en ajoutant 14 codons (Figure 8-B). La traduction étant particulièrement rapide, nous avons également décidé d'éloigner les sites d'hybridations des sondes fluorophores de l'IRES. Cette séquence est traduite en molécule unique dans les mêmes conditions que le test précédent. Les premiers résultats montrent que la proportion d'ARNm traduit est au moins 4 fois plus importante que dans la condition précédente (Figure 9 – Condition 2). La vitesse de la traduction et le temps long sont quant à eux approximativement identiques. Il existerait donc une corrélation entre la distance séparant l'IRES de la lamelle et la proportion d'ARNm traduits.

# ***Résultats***

***Caractérisation de nouveaux  
inducteurs de translecture  
développés par design  
rationnel***

***Visez la lune,  
même si vous cherchez vous serez parmi les étoiles  
(Winston Churchill)***

## Résultats : Caractérisation de nouveaux inducteurs de translecture développés par design rationnel

### 1. Résumé de l'article : Cancer cell apoptosis induced by translational suppression of p53 nonsense mutation

La majorité des cancers est liée à l'apparition de mutations inactivant les gènes suppresseurs de tumeurs. Entre 10 et 30% des mutations impliquées dans ces gènes sont des mutations non-sens qui conduisent à l'arrêt prématuré de la traduction et à la synthèse d'une protéine tronquée non fonctionnelle. La réexpression de la copie sauvage d'un gène suppresseur de tumeur dans une cellule déficiente pour celui-ci permet de limiter la prolifération cellulaire en induisant l'arrêt du cycle ou l'apoptose. Une des stratégies thérapeutiques possibles consiste à utiliser des molécules qui favorisent l'événement d'incorporation d'un acide aminé au niveau de ce codon stop prématuré (translecture) pour permettre la synthèse d'une protéine complète. Cette stratégie comporte toutefois deux limitations majeures. La première concerne la faible efficacité des molécules identifiées à ce jour pour induire la translecture des codons stop prématurés <sup>22</sup>. En effet, l'efficacité de translecture est fortement influencée par l'identité du codon stop et son contexte, les molécules telles que le PTC124 et les RTCs sont spécifiques de certaines mutations. La seconde limitation est liée à la toxicité des molécules, comme la gentamicine qui appartient à la famille des aminoglycosides. L'utilisation de cette molécule dans un traitement clinique à long terme provoquerait une perte auditive irréversible associée à une potentielle perte de la vision. Toutefois, cette toxicité n'étant pas liée à sa capacité à induire la translecture des PTC, il est envisageable de modifier sa structure chimique afin de diminuer sa toxicité tout en améliorant son efficacité.

L'objectif de ce projet est d'étudier la possibilité d'utiliser des inducteurs de translecture dérivés d'aminoglycosides afin de traiter les cancers dans lesquels un gène suppresseur de tumeur est inactivé par une mutation non-sens. Pour cela, nous avons testé plusieurs molécules de synthèse et caractérisé l'une d'entre elles pour sa capacité à induire la réactivation du gène *P53*, le gène suppresseur de tumeur le plus fréquemment muté dans les cancers humains.

Sur la base de la structure chimique de la paromomycine et du G418, cinq dérivés ont été synthétisés, le NB74, NB84, NB122, NB124 et NB128. Le but était d'améliorer leur activité de suppression traductionnelle et de baisser leur toxicité. Dans un premier temps, nous avons quantifié précisément leur efficacité de translecture sur la mutation R213X, qui est la mutation non-sens la plus fréquemment retrouvée dans le gène suppresseur de tumeur *P53*. Cette mutation et son contexte ont été intégrés dans un système double rapporteur puis transfectés dans la lignée cellulaire NIH3T3. Le traitement des cellules par les 5 dérivés a montré que l'ensemble des molécules de synthèse a une meilleure efficacité de translecture que la gentamicine. De plus, le composé NB124 présente le plus fort taux de translecture des cinq nouvelles molécules.

Toutefois cette efficacité dépendant de la nature du codon stop et de son contexte, nous avons cherché à déterminer si le NB124 était capable d'induire de la translecture sur les 5 autres mutations les plus fréquemment retrouvées dans les gènes suppresseurs de tumeurs *P53* et *APC*. Les taux de translecture obtenus montrent que le NB124 est systématiquement meilleur que la gentamicine, quelle que soit la mutation testée. A la concentration la plus élevée, le NB124 induit une augmentation du taux de translecture d'un facteur compris entre 7 et 34 par rapport au taux de base, selon la mutation analysée. Cette molécule est donc capable de supprimer les 3 codons stop et s'avère active dans plusieurs contextes nucléotidiques différents.

La capacité du NB124 à stabiliser l'ARNm *P53* mutant et à restaurer l'expression d'une protéine p53 a été analysée sur la lignée tumorale humaine HDQ-P1 qui possède de façon endogène la mutation la plus fréquente, R213X. Des expériences de PCR quantitatives ont permis de démontrer que l'ARNm mutant est soumis à la voie de dégradation NMD et que le traitement par le NB124 permet de stabiliser les transcrits *P53* de façon dose-dépendante. Le traitement des cellules HDQ-P1 avec le NB124 a également permis de détecter par western-blot la réexpression de la forme complète de la protéine p53.

Pour déterminer si cette protéine est fonctionnelle, nous avons mesuré sa capacité de transactivation sur un système rapporteur dont l'expression est dépendante de la fixation de la protéine p53 sur ses séquences promotrices. Le traitement par le NB124 a permis d'activer l'expression du gène rapporteur d'un facteur 80 à la dose la plus élevée. De plus, comparativement au G418, le NB124 est clairement plus efficace et permet de réexprimer davantage de protéine p53 active. L'analyse par PCR quantitative des taux d'ARNm de gènes cibles cellulaires de la protéine p53 a montré que le traitement par le NB124 permettait de réguler correctement l'expression de deux gènes cibles majeurs : le gène *p21*, impliqué dans l'arrêt du cycle cellulaire est régulé positivement d'un facteur 150, le gène *Bax*, impliqué dans l'apoptose d'un facteur 3.

Pour être utilisée en thérapie antitumorale, la molécule NB124 doit pouvoir activer l'apoptose de cellules cancéreuses. L'activation de la voie apoptotique a été vérifiée par western blot sur la lignée tumorale humaine H1299 dans laquelle l'ADNc P53 muté à la position 213 a été stablement intégré (H1299-p53R213X). Nous avons analysé la capacité de la caspase 3 à cliver la protéine PARP en présence de NB124. Les résultats montrent que la forme clivée de PARP devient clairement détectable après 30 heures de traitement avec le NB124 et que la totalité est clivée à 50 heures. L'activation de l'apoptose a été confirmée en cytométrie de flux qui a permis de démontrer que l'ajout du NB124 dans les cellules H1299-p53R213X induisait l'apoptose dans 38% des cellules.

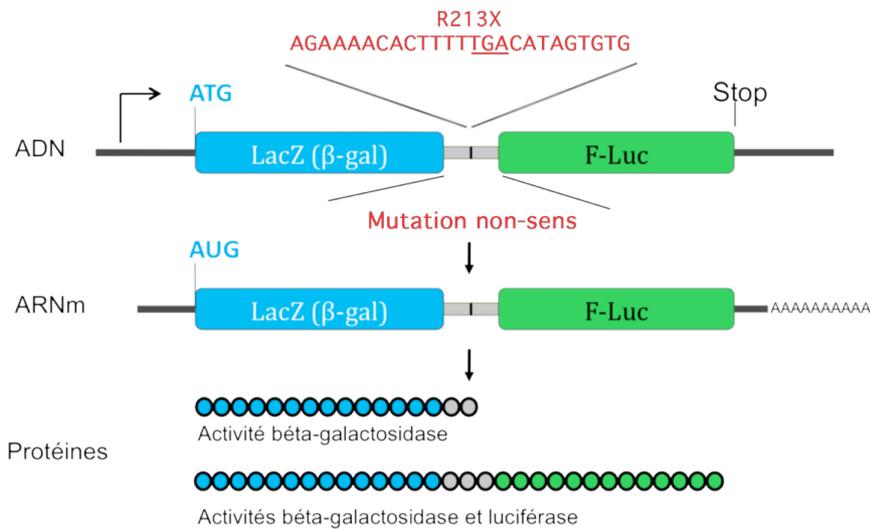
La molécule NB124 permet donc de réexprimer par translecture une protéine p53 entière et fonctionnelle, capable de transactiver l'expression de deux de ses gènes cibles impliqués dans l'arrêt du cycle cellulaire et dans l'apoptose. De plus, ces résultats démontrent clairement la capacité du NB124 à contrer le processus de tumorigenèse en activant l'apoptose des cellules tumorales. L'optimisation de molécules chimiques permettant d'induire de la translecture est donc une stratégie prometteuse dans le traitement des cancers.

## **2. Article : Cancer cell apoptosis induced by translational suppression of p53 nonsense mutation**

# ***Résultats***

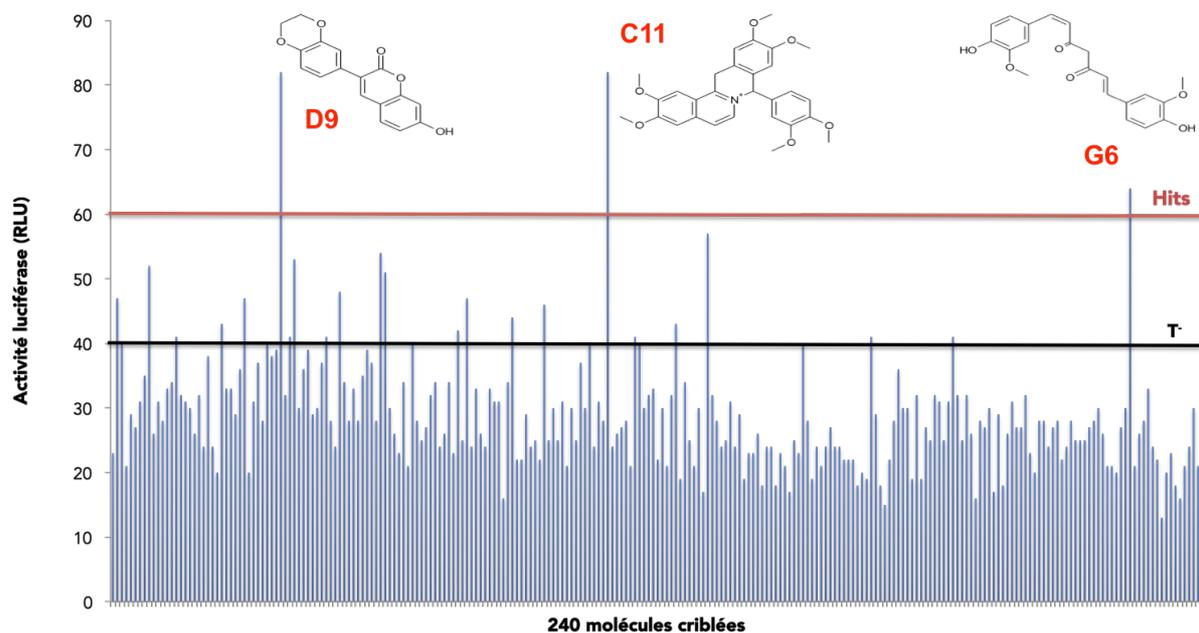
***Identification de nouveaux  
inducteurs de translecture par  
criblage à grande échelle***

***Le génie est fait d'1% d'inspiration  
et de 99% de transpiration  
(Thomas Edison)***



**Figure 1 : Système rapporteur pAC99-R213X**

La mutation non-sens R213X entourée de son contexte nucléotidique naturel est insérée en phase entre les séquences codantes (ORFs) des gènes *LacZ* et *luc* (pAC R213X). Le ratio des activités enzymatiques luciférase sur  $\beta$ -galactosidase permet de calculer le taux de translecture de la mutation analysée. Ce ratio est normalisé par rapport à celui d'un vecteur témoin qui n'a pas de codon stop entre les ORFs *LacZ* et *luc*. Ce rapport représente le 100 % d'activité luciférase.



**Figure 2 : Sélection des molécules de la chimiothèque Greenpharma sur leur capacité à induire la translecture.**

Le système double rapporteur pAC99-R213X a été utilisé pour transfecter les cellules NIH3T3. Les cellules ont été traitées ou non par chacune des 240 molécules naturelles à 200  $\mu$ M pendant 24h (n=1) et l'activité luciférase a été mesurée. La ligne noire (T) correspond à la valeur de l'activité luciférase obtenue en absence de molécules. La ligne rouge<sup>387</sup> indique le seuil d'activité choisi pour sélectionner les molécules comme inducteur potentiel de translecture. Les trois molécules ayant une activité luciférase supérieure au seuil fixé sont : D9 (3-(2,3-dihydro-benzo[1,4]dioxin-6-yl)-7-hydroxy-chromen-2-one), C11 (8-(3,4-dimethoxy-phenyl)-2,3,10,11-tetramethoxy-5,6,13a-tridehydro-berbinium) et G6 (curcumine). La structure chimique de ces composés est représentée sur le graphique.

## Résultats : Identification de nouvelles molécules inductrices de translecture par criblage haut débit

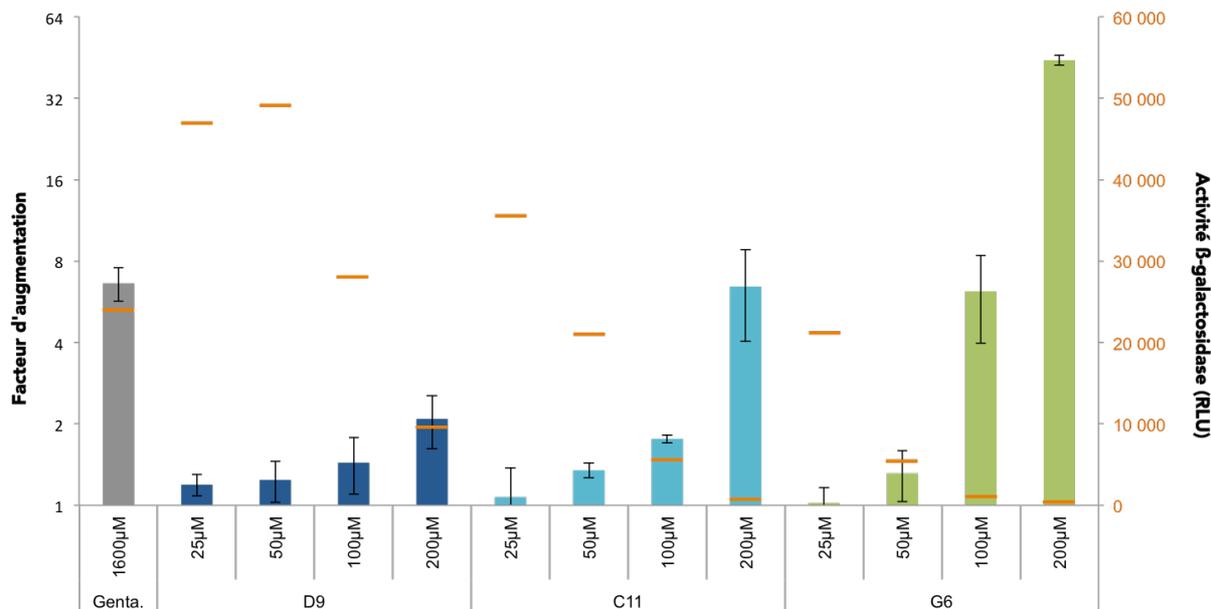
Malgré les résultats très encourageants obtenus avec le NB124, cette molécule est toujours apparentée à la famille des aminoglycosides. Ces composés agissent sur le ribosome en se fixant au site de décodage, provoquant une augmentation des erreurs traductionnelles<sup>238</sup>. De plus, leur spectre d'action est limité à certaines mutations. C'est pourquoi il est important d'identifier de nouvelles familles de molécules qui permettront la suppression traductionnelle des mutations non-sens par d'autres mécanismes. Pour cette raison nous avons réalisé deux cribles à moyenne échelle puis un crible à grande échelle sans a priori sur le mode d'action des molécules recherchées.

### 1. Criblage de la chimiothèque Greenpharma

Dans un premier temps, nous avons criblé la chimiothèque GreenPharma (Prestwick Chemical) constituée de 240 composés naturels de structures très diverses et sélectionné pour avoir un fort potentiel d'activité biologique. Les molécules sont fournies à une concentration de 5 mM dans du DMSO et ont été criblées à une concentration finale de 200  $\mu$ M et 100  $\mu$ M.

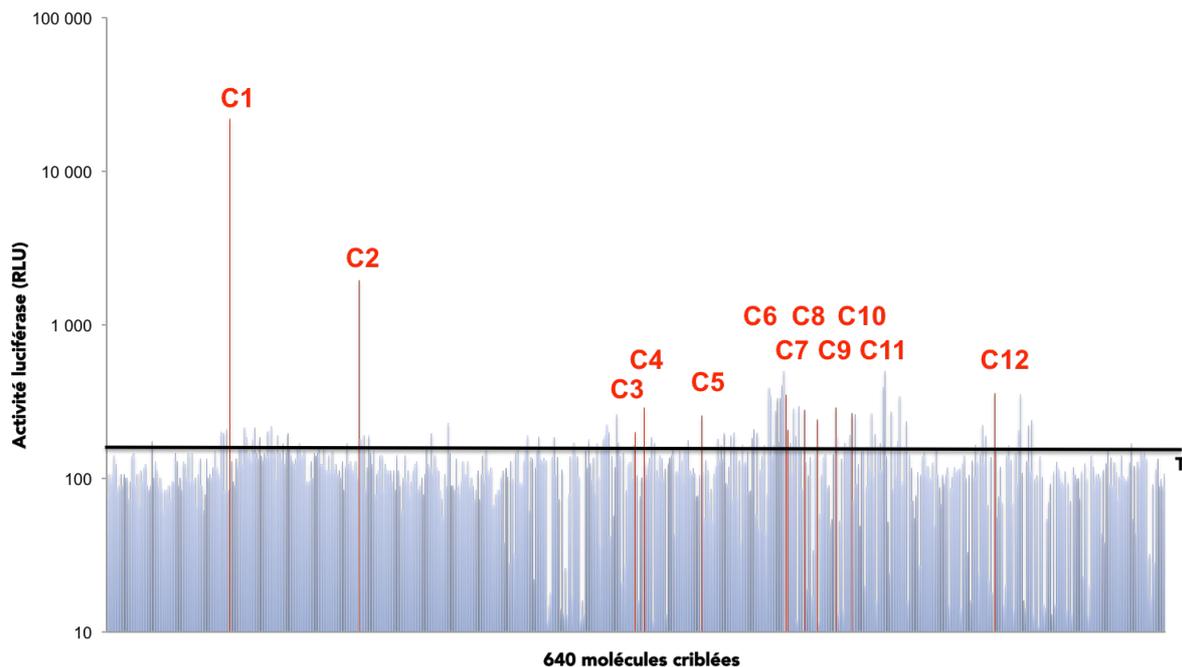
Afin de quantifier précisément l'efficacité de translecture de ces molécules, nous avons utilisé la mutation R213X, retrouvée dans le gène *P53* et déjà décrite comme favorable à la translecture (T TGA C)<sup>240,327</sup>. Cette mutation, entourée de son contexte (4 codons en amont et 3 codons en aval du codon stop) a été clonée dans un système double rapporteur développé au laboratoire (vecteur pAC99), entre les séquences codantes de la bêta-galactosidase et de la luciférase *firefly* (Figure 1). Cette construction est utilisée pour transférer de façon transitoire des fibroblastes de souris, NIH3T3 et les cellules sont traitées avec les 240 molécules de la Greenpharma (n=1). Après 24h de traitement, les extraits protéiques sont réalisés et les activités luciférase et  $\beta$ -galactosidase sont quantifiées par luminométrie dans un lecteur microplaques.

Les résultats des dosages montrent que pour la plupart des molécules, l'activité luciférase mesurée est inférieure à celle retrouvée dans les cellules non traitées. Les valeurs d'activité  $\beta$ -galactosidase sont majoritairement inférieures au seuil de détection probablement due à une cytotoxicité trop importante et ne sont donc pas prises en compte dans les résultats présentés (Figure 2).



**Figure 3 : Effet des molécules D9, C11 et G6 sur la transfecture**

Nous avons quantifié précisément la capacité des molécules à induire la transfecture grâce au plasmide double rapporteur pAC99-R213X. Les cellules ont été traitées durant 24h avec le contrôle gentamicine et les trois molécules issues du crible de la chimiothèque Greenpharma, D9, C11 et G6 (n=3). Ces trois composés sont testés à une gamme de concentration allant de 25 μM à 200 μM. Les mesures d'activités en présence et en absence de molécules permettent de déterminer le facteur d'augmentation. L'activité β-galactosidase qui est un indice de la viabilité cellulaire observée dans chacune des conditions testées est représentée en orange.



**Figure 4 : Sélection des molécules de la chimiothèque nationale essentielle sur leur capacité à induire la transfecture**

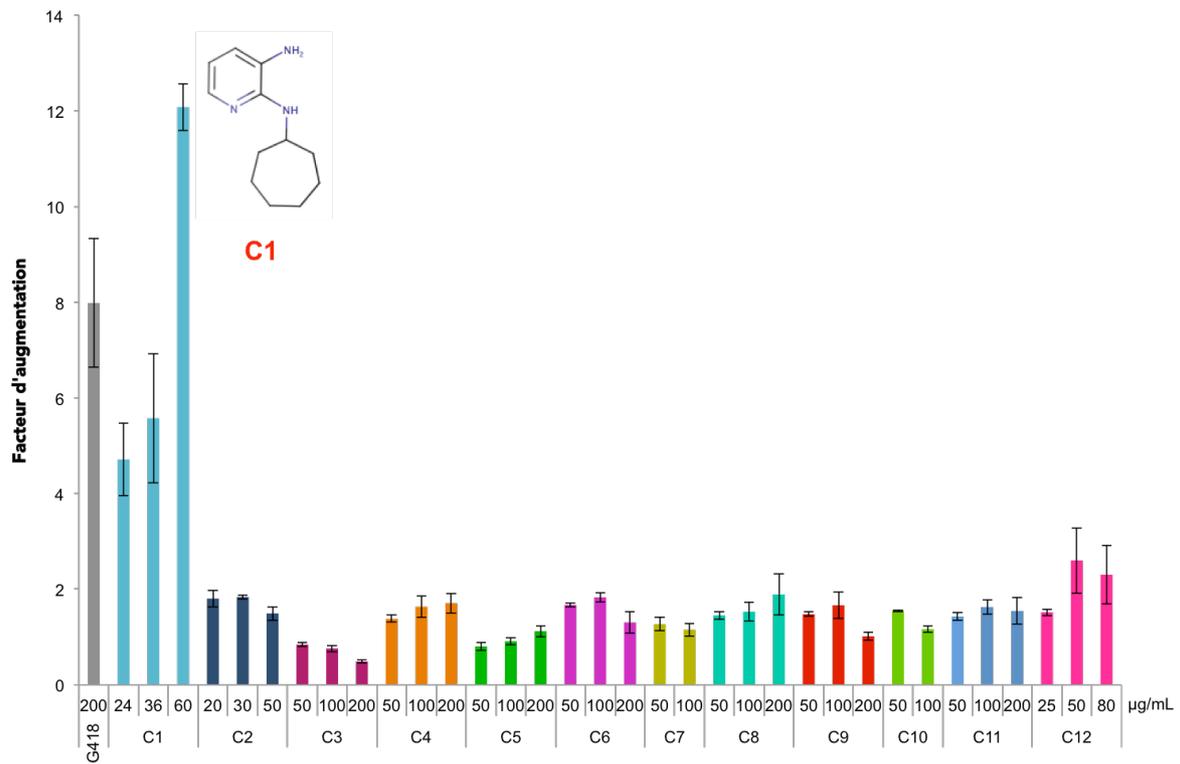
Après transfection des cellules par le système double rapporteur pAC99-R213X, chacune des 640 molécules est additionnées à 50 μg/mL pendant 24h et l'activité luciférase est mesurée (n=1). La ligne noire (T) correspond à la valeur de l'activité luciférase lorsqu'elle est traduite en absence de molécules. Les molécules induisant un niveau d'activité luciférase et un rapport d'activité luciférase/β-galactosidase de 1,5 fois supérieur à celle des contrôles négatifs sont représentées en rouge. Ces molécules ont été sélectionnées à l'issu de ce crible et ont été nommés de C1 à C12. Les concentrations sont exprimées ici en μg/mL et non en μM.

Nous avons identifié et sélectionné trois molécules qui induisent une activité luciférase 1,5 fois supérieure à celles des cellules contrôle. Ces molécules, de structure chimique différente sont la (3-(2,3-dihydro-benzo[1,4]dioxin-6-yl)-7-hydroxy-chromen-2-one), la (8-(3,4-dimethoxy-phenyl)-2,3,10,11-tetramethoxy-5,6,13a-tridehydro-berbinium) et la curcumine. Les activités  $\beta$ -galactosidase mesurées en présence des molécules sont particulièrement faibles et dénotent une toxicité importante de la majorité d'entre elles. Nous avons donc criblé une seconde fois la banque Greenpharma, à 100  $\mu$ M contre 200  $\mu$ M précédemment, sans pour autant identifier de nouvelles molécules induisant une activité luciférase supérieure au seuil fixé. Pour confirmer et préciser l'effet des trois molécules d'intérêt sur la translecture, de nouveaux dosages ont été réalisés avec une gamme de concentration allant de 25  $\mu$ M à 200  $\mu$ M (n=3) (Figure 3). Quelque soit la molécule testée, le facteur d'augmentation entre l'activité luciférase des cellules traitées et celle des cellules contrôles, croit de façon dose-dépendante. Cette réponse est associée à une forte diminution de l'activité  $\beta$ -galactosidase et donc à une toxicité croissante. Aux doses les plus élevées, 100  $\mu$ M et 200  $\mu$ M, les molécules C11 et G6 induisent un taux de translecture équivalent ou supérieur à l'aminoglycoside de référence, la gentamicine. Néanmoins, leur toxicité est 20 fois plus élevée. La molécule D9 est moins toxique que C11 et G6, mais sa capacité à induire l'activité luciférase est 2 à 4 fois plus faible. Aucune des molécules identifiées ne répond donc aux deux critères de sélection choisis : induction de translecture et faible toxicité.

## 2. Criblage de la Chimiothèque Nationale Essentielle (CNE)

Nous nous sommes donc tournés vers une banque plus importante, la Chimiothèque Nationale Essentielle <sup>443</sup> qui représente un sous-groupe de 640 molécules sélectionnées selon des critères de diversité maximale afin de constituer un échantillon représentatif de la chimiothèque nationale. Cette banque regroupe des collections de produits de synthèse, de composés naturels et d'extraits naturels existants dans les laboratoires publics français. Les molécules nous ont été transmises à 1 mg/mL et ont été criblées à 50  $\mu$ g/mL sur le même système rapporteur que lors du crible de la chimiothèque Greenpharma.

Le criblage de la CNE a mis en évidence une centaine de molécules capables d'induire une augmentation de l'activité luciférase d'au moins un facteur 1,5 (n=1) (Figure 4). Toutefois, la valeur de la bêta-galactosidase n'a pas pu être mesurée sur l'ensemble des molécules, car certaines sont trop toxiques, c'est pourquoi nous avons choisi deux critères pour sélectionner les molécules d'intérêt :



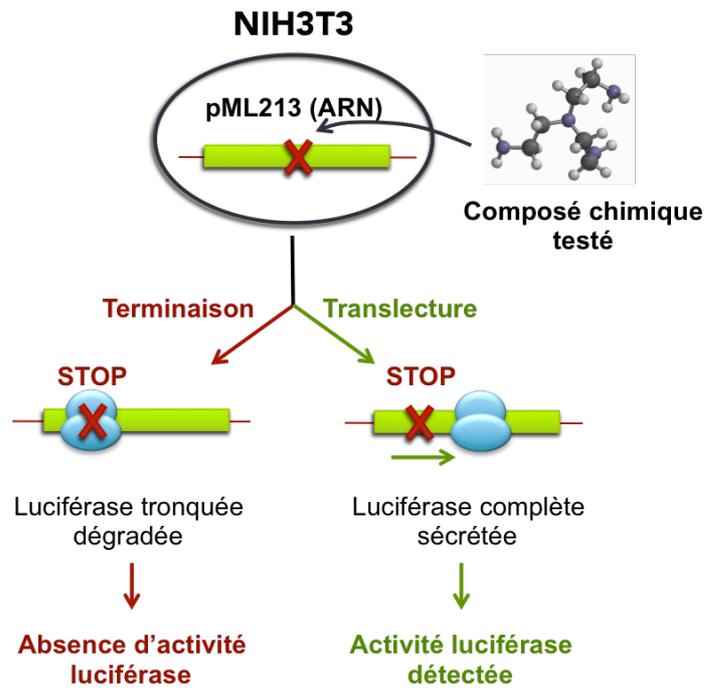
**Figure 5 : Effet des molécules sélectionnées de la CNE sur la translecture**

Nous avons quantifié précisément la capacité des molécules à induire la translecture grâce au plasmide double rapporteur pAC99-R213X ; Les cellules NIH3T3 ont été traitées durant 24h avec le contrôle gentamicine et les 12 molécules issues du cribble de la CNE, C1 à C12 (n=3). Ces 12 composés ont été testés à une gamme de concentration variable et sans dépasser 5% de DMSO. Les mesures d'activités en présence et en absence de molécules permettent de déterminer le facteur d'augmentation. La structure chimique de la molécule C1 est représentée sur le graphique. Les concentrations sont exprimées ici en µg/mL et non en µM.

l'activité luciférase avant normalisation doit être 1,5 fois supérieure à celle des cellules non traitées car des valeurs inférieures correspondraient soit à une capacité d'induction de translecture trop faible soit à une toxicité trop importante, et une activité luciférase après normalisation (rapport d'activité luciférase/ $\beta$ -galactosidase) 1,5 fois supérieur à celui des contrôles négatifs. Ces critères nous ont permis de sélectionner 12 potentiels inducteurs de translecture (C1-C12) que nous avons commandés.

Pour confirmer et préciser l'effet potentiel sur la translecture de ces molécules, un nouveau dosage a été réalisé sur une gamme de concentration de 50  $\mu\text{g/mL}$  à 200  $\mu\text{g/mL}$  (Figure 5). En revanche, le DMSO étant toxique pour les cellules au-delà d'une certaine concentration, nous n'avons pas pu tester les doses les plus élevées pour toutes les molécules, et notamment celles qui ont du être reprises à faible concentration en raison de leur mauvaise solubilité. Les résultats ont montré que les facteurs d'augmentation entre l'activité luciférase des cellules traitées et celle des cellules contrôles, aux différentes doses testées ne sont pas statistiquement différents avec les molécules C2 à C12. De plus, aucun effet dose-réponse n'est observé avec ces composés. A contrario, le facteur d'augmentation obtenu en traitant les cellules avec la molécule C1 croit de façon dose-dépendante. Cette molécule est également la seule à augmenter l'activité luciférase mesurée de façon plus importante que l'aminoglycoside de référence, le G418, dont la concentration est trois fois plus élevée que C1. Nous avons donc poursuivi nos essais uniquement avec C1 (N2-cycloheptyl-2,3-pyridinediamine).

Les systèmes rapporteurs basés sur le gène de la luciférase étant sujets à l'identification de faux positifs qui interagissent directement avec le substrat (luciférine) ou qui stabilisent la luciférase, nous avons traité les cellules NIH3T3 non transfectées pendant 24h en présence de C1 ou d'un aminoglycoside contrôle, le G418. Le dosage des activités luciférase et  $\beta$ -galactosidase a mis en évidence que la molécule C1 était capable d'induire une augmentation de signal de luminescence malgré l'absence du gène codant la luciférase. Cela signifierait que C1 pourrait stimuler un facteur cellulaire capable de réaliser la réaction d'oxydation de la luciférine, provoquant la libération de photons mimant ainsi l'activité luciférase. Par conséquent, aucune des molécules criblées de la CNE, dans nos conditions expérimentales, ne permet d'augmenter le taux de translecture.



**Figure 6 : Test biologique développé pour le criblé à grande échelle**

Ce système rapporteur est basé sur l'utilisation d'un gène rapporteur de luciférase de *Metridia* interrompu par une mutation non-sens et intégré stablement dans la lignée cellulaire NIH3T3. Le traitement par une molécule inactive sur la translecture conduira à la synthèse d'une protéine tronquée qui sera dégradée dans la cellule et non détectée lors du dosage de l'activité luciférase. Alors qu'en présence d'une molécule inductrice de translecture, le ribosome continuera sa traduction malgré la présence d'un codon stop prématuré et synthétisera une luciférase complète. Cette dernière comporte une séquence signal d'export extra-cellulaire, ce qui permet de doser son activité directement dans le milieu de culture.

### 3. Criblage à haut débit de la Prestwick et d'une fraction de la Chembridge

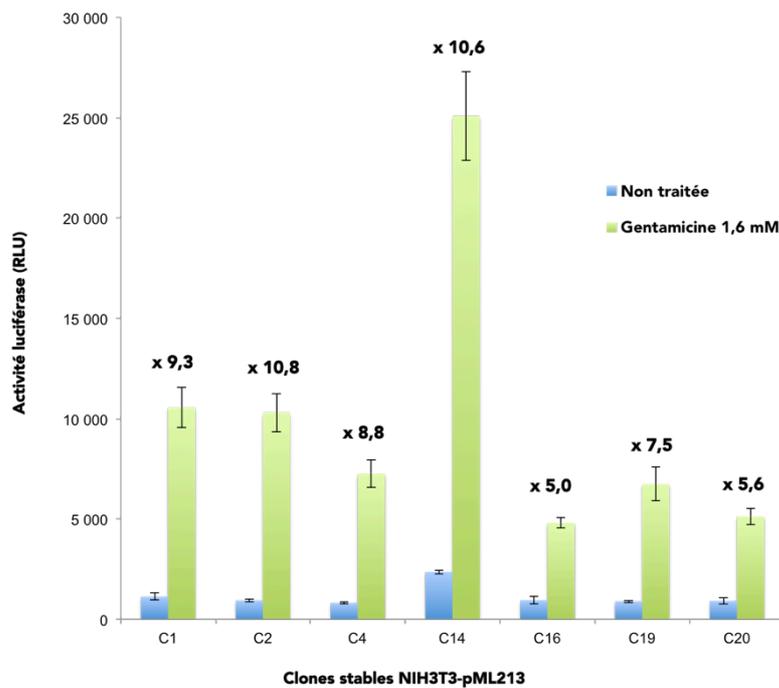
Ces résultats et ceux obtenus précédemment par d'autres laboratoires ayant réalisé des cribles de chimiothèques illustrent bien la difficulté d'identifier des inducteurs de translecture. Les cribles à grande échelle (HTS) réalisés par les laboratoires pharmaceutiques se focalisent de plus en plus sur des banques de taille réduite, mais définies selon des critères bien spécifiques<sup>368,374</sup>. Le développement d'un test biologique pertinent contribue également largement à la réussite du crible à haut débit en augmentant la probabilité d'identifier un vrai « Hit » (« touche » en français) et en limitant la quantité de faux-positifs et faux-négatifs. Nous avons donc développé notre propre système biologique et criblé plus de 17 000 composés sur leur capacité potentielle à induire de la translecture d'un codon stop prématuré, en collaboration avec l'équipe de J-C. Cintrat, du CEA de Saclay.

#### A. Construction du test biologique

Afin de mettre en évidence de nouvelles molécules inductrices de translecture, nous avons cherché à développer un test biologique compatible avec la réalisation d'un crible à grande échelle de chimiothèques. Pour cela, ce système doit répondre à plusieurs critères : il doit être facile à mettre en œuvre, il doit permettre la détection spécifique des molécules induisant de la translecture et doit permettre de répondre à l'ensemble des normes de contrôle qualité sur lesquelles je reviendrai plus tard.

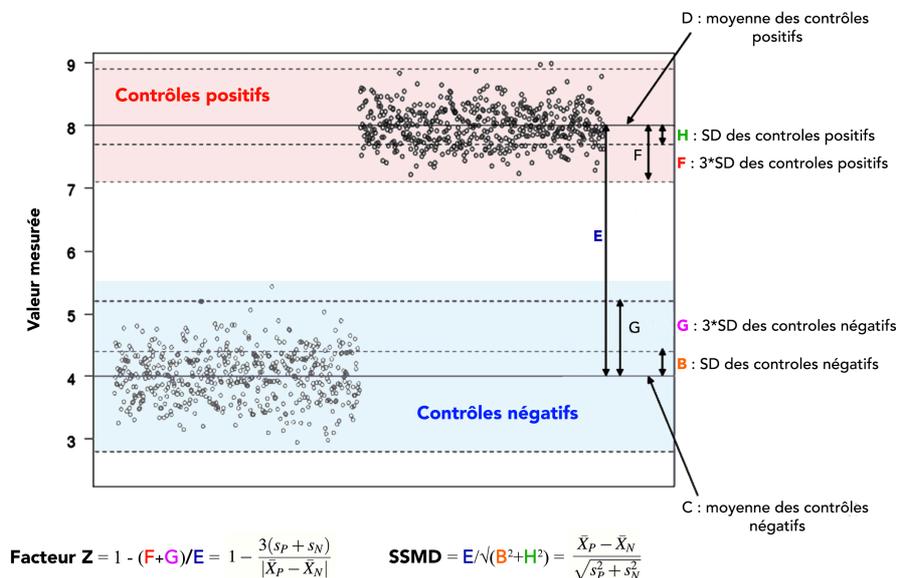
Nous avons donc développé un système reposant sur le vecteur pMetLuc (pML)<sup>444</sup> qui contient un gène codant une luciférase de *Metridia* (*metluc*), produisant une bioluminescence bleue ( $\lambda_{\max} = 480\text{nm}$ ), sous contrôle d'un promoteur fort et constitutif, pCMV. Cette luciférase a la particularité d'être sécrétée dans le milieu de culture, ce qui permet de réaliser les dosages enzymatiques en s'affranchissant des étapes de récoltes des cellules et d'extraction protéiques<sup>445</sup>. Son activité est mesurée par l'ajout de son substrat, la coelenterazine, qui après oxydation, génère une émission de photons. Ce vecteur contient aussi le gène de résistance à la néomycine qui permet de sélectionner les cellules l'ayant intégré de façon stable.

Le principe du crible consiste à interrompre le gène *metluc* par une séquence contenant un codon stop afin de conditionner son expression à un événement de translecture (Figure 6). Nous avons cloné la mutation R213X dans la région N-terminale du gène *metluc*. Après avoir vérifié que la transfection transitoire du système rapporteur



**Figure 7 : Sélection d'un des clones ayant intégré le vecteur pML213**

L'activité luciférase de 7 clones de cellules NIH3T3 ayant stablement intégré le vecteur pML213 a été mesurée en absence et en présence de gentamicine 1,6 mM durant 24h (n=4). Le signal de luminescence est quantifié après ajout de coelenterazine (Clontech). Le facteur d'augmentation entre la condition traitée et contrôle est indiqué au dessus de chaque histogramme. Le clone stable sélectionné est le C14 car il a non seulement un facteur d'augmentation très élevé, mais son signal est en moyenne 2 à 3 fois plus élevé que les autres clones testés.



**Figure 8 : Indices statistiques de contrôle qualité des mesures du crible à grande échelle**

Schéma représentant le résultat de la mesure des contrôles positifs et négatifs d'une plaque. La déviation standard (SD) permet d'évaluer la dispersion des mesures autour de la moyenne. Le facteur Z est un indice statistique qui tient compte à la fois de la dispersion des points (F et G) et de l'écart entre les contrôles positifs et négatifs (E). Le SSMD est un indice alternatif de qualité d'un crible HTS, qui est capable de prendre en compte davantage de biais que le facteur Z. Les formules de calcul du facteur Z et du SSMD sont indiquées où « s » correspond à la déviation standard, «  $\bar{X}$  » la moyenne, « p » les contrôles positifs et « n » les contrôles négatifs.

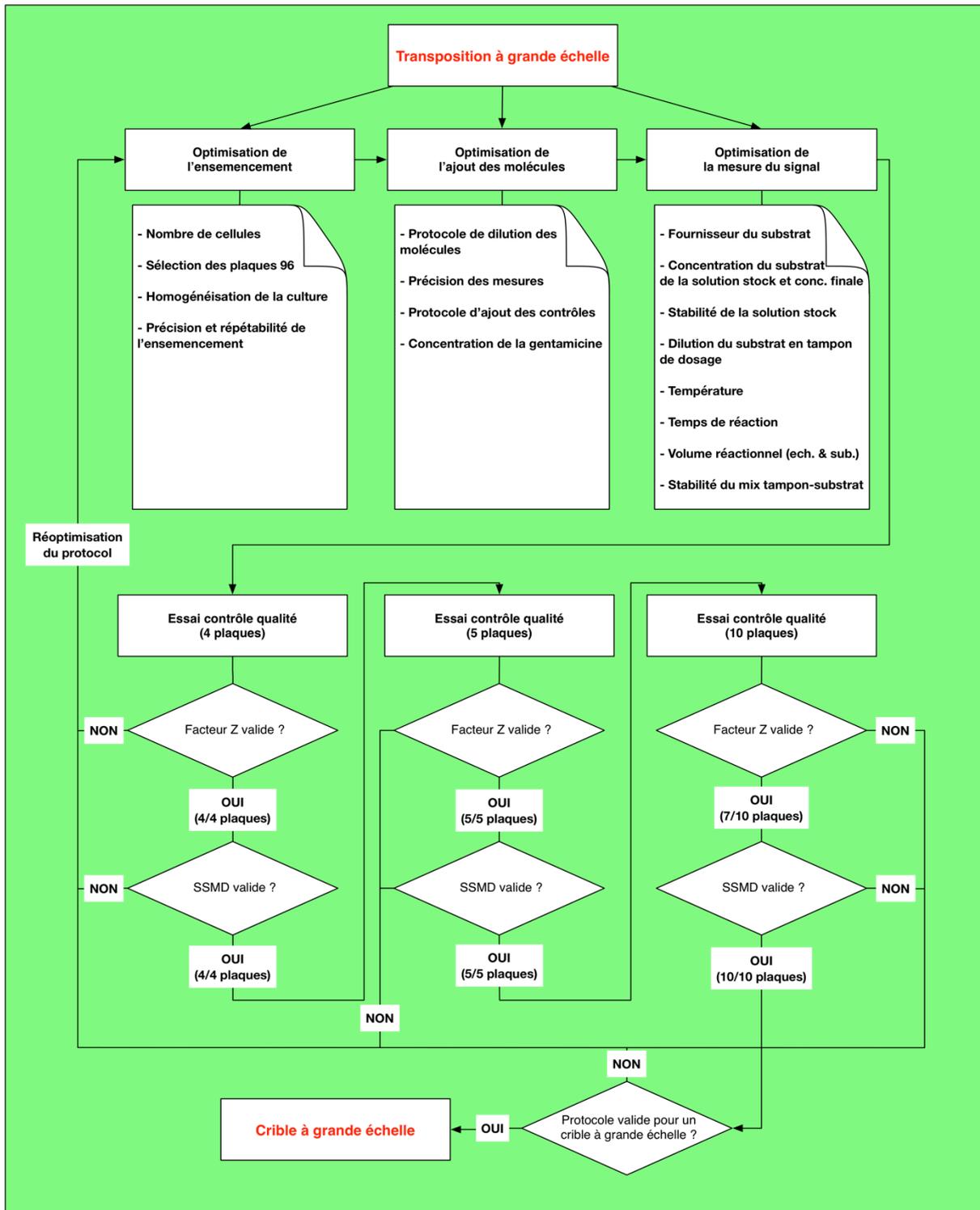
dans les cellules NIH3T3 répondait bien au traitement par un inducteur de translecture connu, la gentamicine, nous l'avons intégré de façon stable dans les cellules NIH3T3. Des essais de translecture ont été effectués dans une vingtaine de clones et nous avons sélectionné le clone C14 dont l'induction d'activité luciférase est parmi les plus importantes en présence de gentamicine et dont le signal de luminescence est le plus élevé (Figure 7). Nous avons ensuite vérifié la spécificité de notre système en traitant les cellules C14 (3T3pML213) avec soit un inducteur de translecture connu, la gentamicine (1,6 mM), soit avec l'apramycine (1,85 mM) qui appartient également à la famille des aminoglycosides et qui n'induit pas de translecture malgré sa liaison au site A du ribosome <sup>446</sup>. Les dosages ont montré que seul le traitement des cellules 3T3pML213 avec la gentamicine induisait une augmentation de l'activité luciférase par rapport aux cellules non traitées. Ces résultats démontrent que l'activité luciférase de *Metridia* est induite par les molécules stimulant la translecture.

## B. Contrôle qualité lors d'un crible à grande échelle

Lors d'un crible à grande échelle, de nombreuses variables environnementales, instrumentales et biologiques interviennent et peuvent affecter sa qualité. En effet, la plupart des essais conçus pour les HTS possèdent un degré de variabilité et d'erreur relativement élevé. Plusieurs indices statistiques ont été développés pour s'assurer que les données recueillies répondent aux exigences de qualité requises. Ces indices, attribuent un score à chaque plaque et valident ou invalident ses résultats de mesure. Pour cela, des contrôles positifs et négatifs sont introduits dans toutes les plaques et sont analysés après normalisation et correction des erreurs systématiques en utilisant les deux principaux outils de mesure statistique de taille d'effets, le facteur Z et le SSMD (Figure 8) <sup>447</sup> :

Le facteur Z (ou Z') est l'indice statistique le plus utilisé dans les HTS et permet d'analyser la variabilité entre deux populations, qui sont ici les contrôles positifs et négatifs. Il prend en compte à la fois la dispersion des mesures (déviation standard), plus communément appelés « points » et l'écart entre les contrôles positifs ( $C_{pos}$ ) et négatifs ( $C_{neg}$ ). Ainsi, une faible dispersion et un fort écart entre les contrôles impacteront positivement le score du facteur Z. La qualité d'un crible HTS est alors catégorisée selon la valeur du facteur Z obtenue :

- pour  $1 > Z > 0,9$  : le criblage est d'excellente qualité
- pour  $0,9 > Z > 0,7$  : le criblage est bon
- pour  $0,7 > Z > 0,5$  : l'amélioration du processus de mesure rendra la sélection des « hits » plus exacte.
- $Z = 0,5$  : le minimum acceptable pour un crible à grande échelle



**Figure 9 : Mise au point du dosage pour un criblé à grande échelle**

Arbre décisionnel représentant les différentes optimisations apportées au test biologique pour la mise en œuvre du criblé à grande échelle. A l'issue de cette phase de mise au point, trois essais en conditions réelles de criblage ont permis d'évaluer le niveau de contrôle qualité des plaques criblées et de valider le protocole pour sa réalisation à grande échelle.

Le SSMD, où Différence des Moyennes Strictement Standardisées, a été plus récemment intégré aux HTS et analyse l'ampleur de la différence entre deux populations. Ce test prend en compte les mêmes paramètres que le facteur Z, mais il pondère la variabilité des deux populations contrôles différemment.

Cela permet de relier directement le score du SSMD à la probabilité que les deux populations contrôles soient statistiquement différentes (Figure 8) <sup>447,448</sup>.

La grille des scores du SSMD varie selon le type de crible et l'importance de l'effet du contrôle positif. Nous avons choisi ici d'utiliser la gentamicine qui est l'inducteur de translecture le mieux caractérisé à l'heure actuelle et qui ne présente pas de toxicité cellulaire. La grille des scores s'appliquant à notre crible est donc :

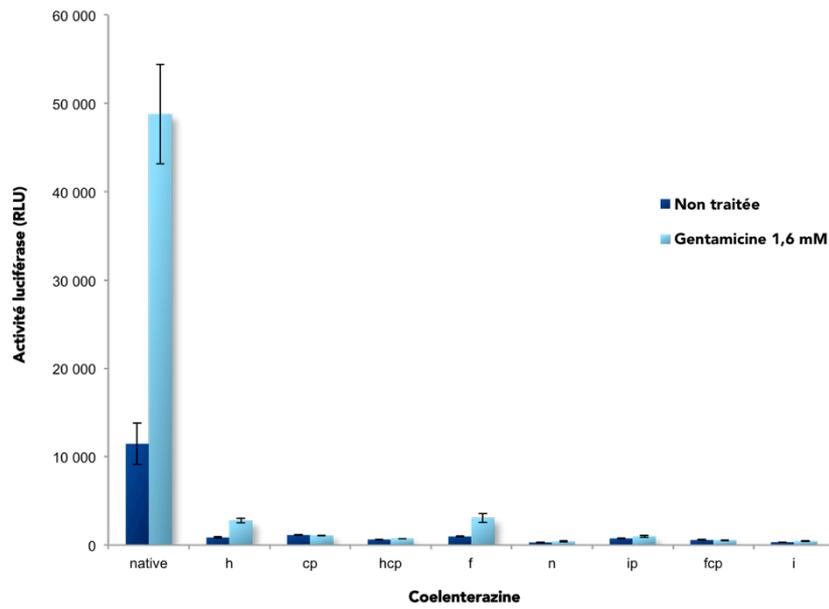
- pour  $SSMD \geq 4,7$  : le criblage est d'excellente qualité
- pour  $4,7 > SSMD \geq 3$  : le criblage est bon
- pour  $3 > SSMD \geq 2$  le criblage est de qualité inférieure
- pour  $SSMD < 2$  : le criblage est de pauvre qualité

La présence de valeurs extrêmes au sein des populations contrôles influencent leur moyenne alors que ces valeurs ne sont pas représentatives de la population globale et sont le plus souvent liées à des biais techniques ou biologiques ponctuels. Le remplacement de la moyenne par la médiane augmente la robustesse du test statistique et permet de minimiser l'effet des valeurs extrêmes. Néanmoins il requiert un minimum de 8 points pour chacun des contrôles. Cette condition étant remplie dans ce crible, le facteur Z et le SSMD ont donc été calculés à partir de la médiane des populations contrôles.

### C. Mise au point du dosage pour un crible à grande échelle

Le développement d'un protocole de dosage de l'activité luciférase de *Metridia* pour un crible à grande échelle et la validation du contrôle qualité ont nécessité un important de travail de mise au point que j'ai réalisé avec Goulven Merer de la plateforme de criblage haut débit de l'équipe de J-C Cintrat (Figure 9). En effet, deux contraintes principales ont été identifiées : le facteur d'augmentation du signal de luminescence entre les puits non traités et ceux traités par la gentamicine doit être suffisamment élevé pour passer les contrôles qualité et le substrat une fois repris dans son tampon doit être stable sur une longue période pour permettre le dosage de plusieurs plaques sans introduire de biais de mesure.

Dans un premier temps, j'ai recherché un tampon permettant de doser l'activité luciférase de *Metridia*. Un tel tampon a été décrit par Lupold et ses collaborateurs et j'ai confirmé son efficacité sur les cellules 3T3pML213 <sup>449</sup>.



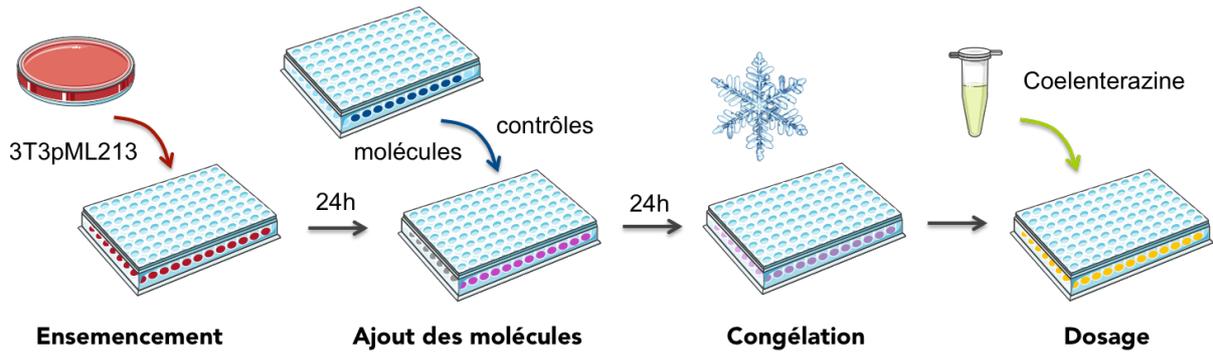
**Figure 10 : Analyse du signal de luminescence en présence des différents substrats (coelenterazine)**

Les cellules 3T3pML213 sont traitées ou non pendant 24h par la gentamicine 1,6 mM. Le dosage de l'activité luciférase est réalisé après ajout de coelenterazine native ou d'un des analogues à 24  $\mu$ M final (Biotium).

Nous nous sommes alors concentrés sur l'optimisation de l'intensité du signal de luminescence en faisant varier plusieurs paramètres. J'ai commencé par rechercher un substrat dont les propriétés de luminescence répondaient à trois critères un signal de luminescence élevé en présence de l'enzyme, un faible bruit de fond et une stabilité du signal de luminescence la plus importante possible. J'ai testé la coelenterazine native et 8 analogues (Biotium) repris à 5 mM en éthanol pur et dilués à 24  $\mu$ M finale dans le tampon Metridia, puis j'ai dosé l'activité luciférase en présence ou non de gentamicine à 1,6 mM (Figure 10). J'ai ainsi pu montrer que la coelenterazine native était très nettement supérieure par rapport à ses analogues structuraux. Plusieurs substrats provenant d'autres fournisseurs ont également été testés et l'un d'eux émet 4 fois plus de signal (Nanolight), tout en maintenant le même facteur d'augmentation entre les cellules contrôle et celles traitées à la gentamicine 1,6 mM. De plus, la bioluminescence étant optimale entre 10 et 12°C, la température de la pièce a été abaissée au maximum possible soit 16°C<sup>450</sup>. La concentration en gentamicine a également été augmentée à 2,5 mM pour accroître la luminescence des contrôles positifs. Les mesures de viabilité réalisées sur la lignée NIH3T3 en présence de gentamicine 2,5 mM n'ont pas mis en évidence d'effets toxiques importants (Figure 22). Enfin, le nombre de cellulesensemencées a été optimisé. L'ensemble de ces mises au point nous a permis d'augmenter le signal d'un facteur 16 par rapport aux premiers essais.

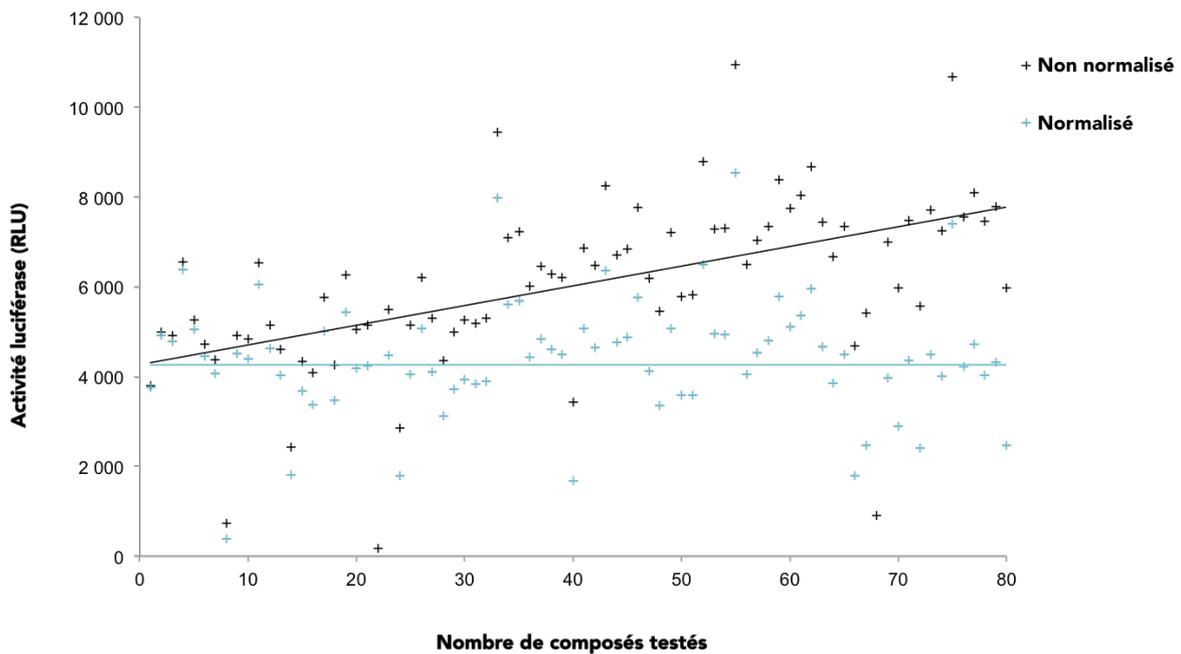
Dans un second temps, nous nous sommes concentrés sur la stabilisation du substrat en limitant son oxydation par l'oxygène de l'air. La coelenterazine est fournie dans des tubes sous argon, mais une fois repris dans de l'éthanol pur, la coelenterazine s'altère en s'oxydant. Le substrat a donc été remis en solution avec de l'éthanol dégazé à l'azote puis stocké à -20°C où il est plus stable<sup>451</sup>. Lors de l'étape de dosage, la coelenterazine est diluée dans du tampon Metridia dégazé à l'azote puis est immédiatement répartis en plaque 96 et recouvert d'huile minérale pour minimiser son contact avec l'oxygène. De plus, ce substrat est particulièrement sensible à la lumière, nous avons donc diminué le plus possible la luminosité de la pièce de dosage et protégé la coelenterazine tout au long du processus de mesure. La stabilité de ce substrat a ainsi été multipliée par 12, ce qui permet de doser consécutivement 20 plaques sans refaire de mix tampon-substrat.

Les mises au point précédemment décrites ont largement contribué à la validation du contrôle qualité des plaques. Toutefois, il fut tout de même nécessaire d'optimiser chacune des étapes du protocole de criblage décrit ci-dessous afin de limiter les biais instrumentaux et biologiques.



**Figure 11 : Protocole du criblage haut débit**

Schéma représentant le protocole simplifié du criblage des chimiothèques Prestwick et Chembridge dans l'objectif d'identifier des molécules inductrices de translecture. Les cellules 3T3pML213 sont tout d'abord ensemencées en plaque 96 puits. 24 heures plus tard, les contrôles et les molécules des chimiothèques sont ajoutés dans les puits de culture. Après 24 heures de traitement, les plaques sont congelées à  $-20^{\circ}\text{C}$  et l'activité luciférase est dosée ultérieurement en ajoutant la coelenterazine dans chacun des puits et en réalisant les mesures à l'aide d'un luminomètre à plaques.



**Figure 12 : Correction de la pente de la mesure de l'activité luciférase**

Exemple de plaque de molécules nécessitant une correction de la pente de l'activité luciférase mesurée en raison de l'augmentation progressive du signal entre le premier et dernier puits. Les valeurs d'activités luciférases sont rangées dans l'ordre dans lequel ils ont été dosés. La pente est calculée à partir des puits contrôles négatifs situés aux deux extrémités de la plaque et est ensuite appliquée à ces mêmes contrôles ainsi qu'à l'ensemble des puits traités par des molécules issues de la chimiothèque.

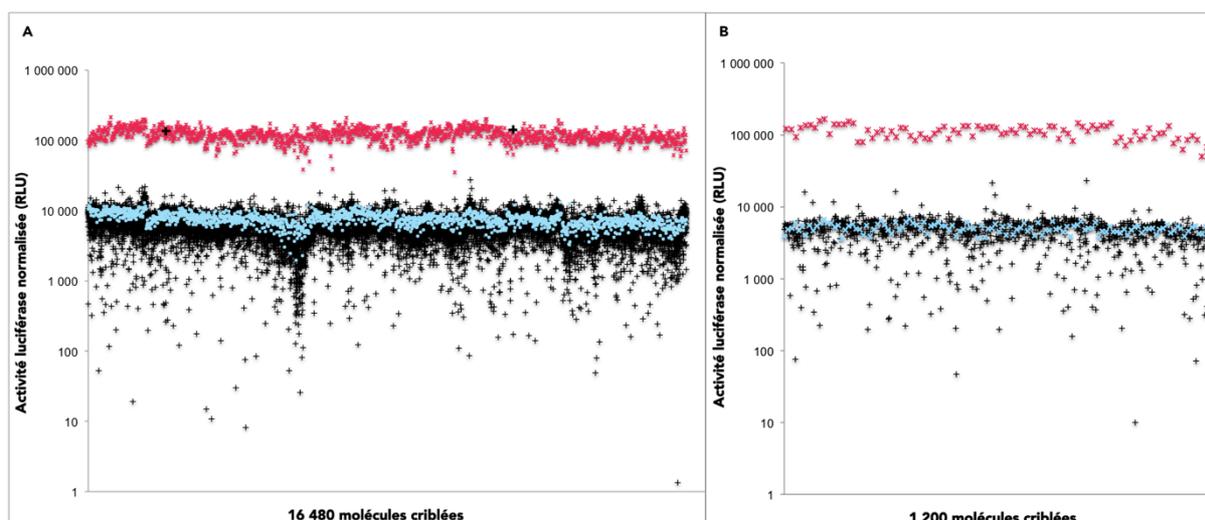
## D. Protocole de criblage

L'utilisation de la luciférase de *Metridia* qui est sécrétée directement dans le milieu de culture permet de réaliser les étapes de culture, traitement et dosage dans la même plaque. Nous avons ainsi pu développer un protocole en 4 étapes (Figure 11) : les cellules 3T3pML213 sontensemencées en plaque 96 puits. 24 heures plus tard, le milieu de culture est aspiré. Les molécules tests et contrôles dilués dans du milieu de culture sont déposées dans les plaques de culture. Après 24 heures de traitement, les plaques sont congelées à -20°C. Le dosage des plaques est réalisé en ajoutant le substrat dans chacun des puits puis en mesurant le signal de luminescence.

Ce protocole a été vérifié par trois essais en conditions réelles de criblage qui ont été réalisés sur des plaques de chimiothèque. Le facteur Z et le SSMD ont été validés dans les deux premiers essais de 4 et 5 plaques. Sur un jeu de 10 plaques, 3 plaques ont obtenu un facteur Z compris entre 0,45 et 0,5 suite à la présence d'une valeur aberrante dans un des puits des contrôles positifs. Le SSMD qui est plus robuste a validé la totalité des plaques.

## E. Nature des banques criblées

Ce crible a été réalisé sur deux banques de molécules. La Prestwick Chemical Library qui est une collection unique de 1 280 petites molécules, dont l'autorisation de mise sur le marché européen et américain a été délivrée. Ces molécules ont été sélectionnées pour leur diversité chimique et pharmacologique, pour leur biodisponibilité et leur absence de toxicité chez l'homme. La deuxième chimiothèque est un sous-groupe de 16 480 composés de la banque ChemBridge qui en comprend plus d'un million à ce jour. Ces molécules ont été sélectionnées par l'équipe de J-C. Cintrat sur des critères de diversité chimique et pharmacologique. Les 17 680 molécules sont stockées en plaques 96 puits à 10 mM et ont été criblées à 50 µM. Les résultats du crible primaire et secondaire, ainsi que l'analyse des résultats ont été synthétisée dans un arbre décisionnel (Annexe 2).



**Figure 13 : Crible primaire de 17 680 molécules**

Résultats du crible des chimiothèques Chembridge (A) et Prestwick (B) après correction des erreurs systématiques (n=1). Les mesures d'activité luciférase en présence des molécules testées sont représentées en noir, les contrôles positifs (gentamicine) sont représentés en rouge et les contrôles négatifs (sans molécule) en bleu.

$$\text{SSMD}^* = \frac{Y_i - \tilde{Y}_N}{\sqrt{\frac{2}{K}(n_N - 1)} \text{MAD}_N} \quad \text{MAD} = 1,4826 \text{ médiane} (|Y_i - \text{médiane}(Y)|)$$

$$K \approx n_N - 2,48$$

**Figure 14 : Formule de calcul du SSMD pour l'identification des hits**

Le calcul du SSMD évalue l'effet de chaque composé testé par rapport à l'ensemble des molécules criblées d'une plaque. SSMD\* correspond à la version robuste du SSMD remplaçant la déviation standard par le MAD. Le MAD représente la médiane des déviations absolues.  $Y_i$  est la valeur normalisée des  $i$  puits traités.  $\tilde{Y}_N$ ,  $n_N$ ,  $\text{MAD}_N$  et  $Y$  sont la médiane, le nombre de puits, le MAD et les valeurs normalisées des puits traités par une molécule sur une plaque.

## F. Correction des erreurs systématiques

Une fois les chimiothèques criblées, les données générées sont soumises à plusieurs tests permettant d'identifier la présence de variations aléatoires qui nécessiteraient une normalisation ponctuelle des résultats ou la présence d'erreurs systématiques provoquées par un biais lors de l'expérimentation et qui s'appliquent à une partie voir la totalité des plaques criblées. Ce type d'erreur doit être corrigé avant de vérifier le contrôle qualité des plaques et d'analyser les résultats en vue d'identifier les hits.

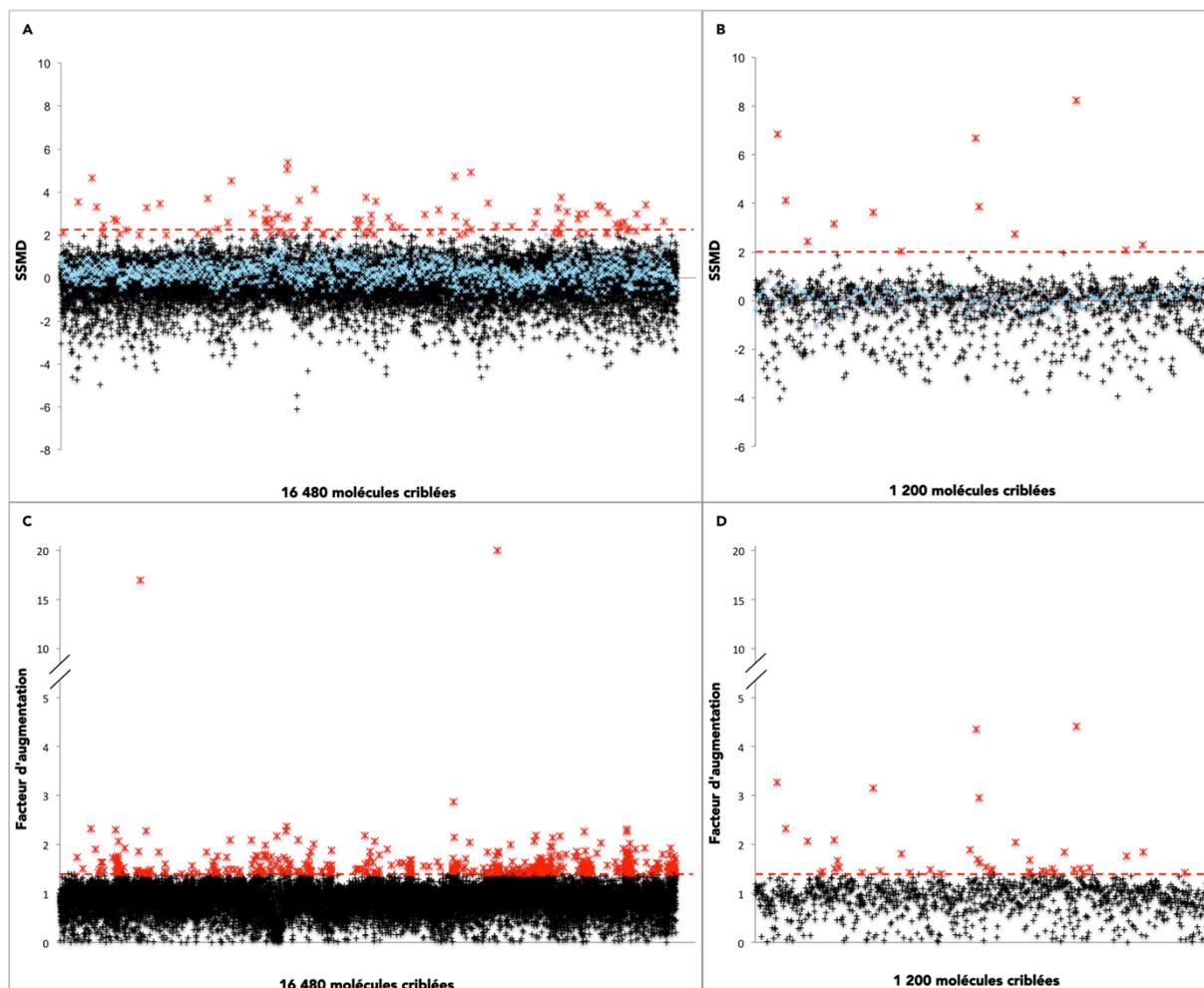
Dans notre crible, nous avons pu identifier la présence d'une erreur systématique liée au temps de dosage de l'ensemble des puits d'une plaque (Figure 12). En effet, après l'ajout du substrat, l'activité dosée est stable sur une période limitée puis décroît rapidement, provoquant ainsi une variation du signal entre le premier et le dernier puits. Pour limiter ce biais, les plaques sont dosées durant la première phase de stabilité du signal. De plus, la pente des contrôles négatifs est calculée en fonction de leur position sur la plaque et de l'ordre de lecture du luminomètre <sup>452</sup>. La valeur d'activité luciférase des puits de molécules tests et contrôles négatifs est ensuite corrigée en fonction de leur position et de la pente calculée. Ce même processus est appliqué aux contrôles positifs.

## G. Analyse des résultats

J'ai ensuite analysé les données issues de l'étape de correction des erreurs systématiques à l'aide d'une table que j'ai développée (Figure 13). J'ai ainsi pu vérifier le contrôle qualité de chaque plaque et 12% d'entre elles ont un facteur Z inférieur à 0,5 en raison d'une valeur aberrante. Le SSMD valide quant à lui l'ensemble des plaques criblées, avec toutefois un score compris entre 3 et 4,7 pour 9,5% d'entre elles.

L'identification des hits positifs s'est faite selon deux critères : le calcul du facteur d'augmentation entre la condition traitée et la médiane des contrôles négatifs, et la version robuste du SSMD appliquée à la sélection des hits, qui diffère du SSMD calculé pour le contrôle qualité des plaques.

Le SSMD est un indice statistique qui tient compte de la médiane des mesures et de la variabilité de l'effet des composés. Cependant, contrairement aux puits contrôles, les puits traités par les molécules de la chimiothèque n'ont pas de répliques. En partant du principe que la quasi-totalité des molécules criblées n'ont pas d'effet, la seule façon de calculer cette variabilité est d'admettre que l'effet des composés varie globalement comme celui des contrôles négatifs.



**Figure 15 : Sélection des hits primaires**

Les molécules hits ont été sélectionnées à l'issue du crible, après normalisation des résultats, selon leur score SSMD (**A**, **B**) et selon le facteur d'augmentation de l'activité luciférase entre la condition traitée et la médiane des contrôles négatifs (**C**, **D**). Les molécules provenant de la chimiothèque Chembridge sont représentées en **A** et **C**, et celles de Prestwick en **B** et **D**. Les croix rouges correspondent aux molécules dont le SSMD  $\geq 2$  ou le facteur d'augmentation est  $\geq 1,4$  soit un total de 465 molécules.

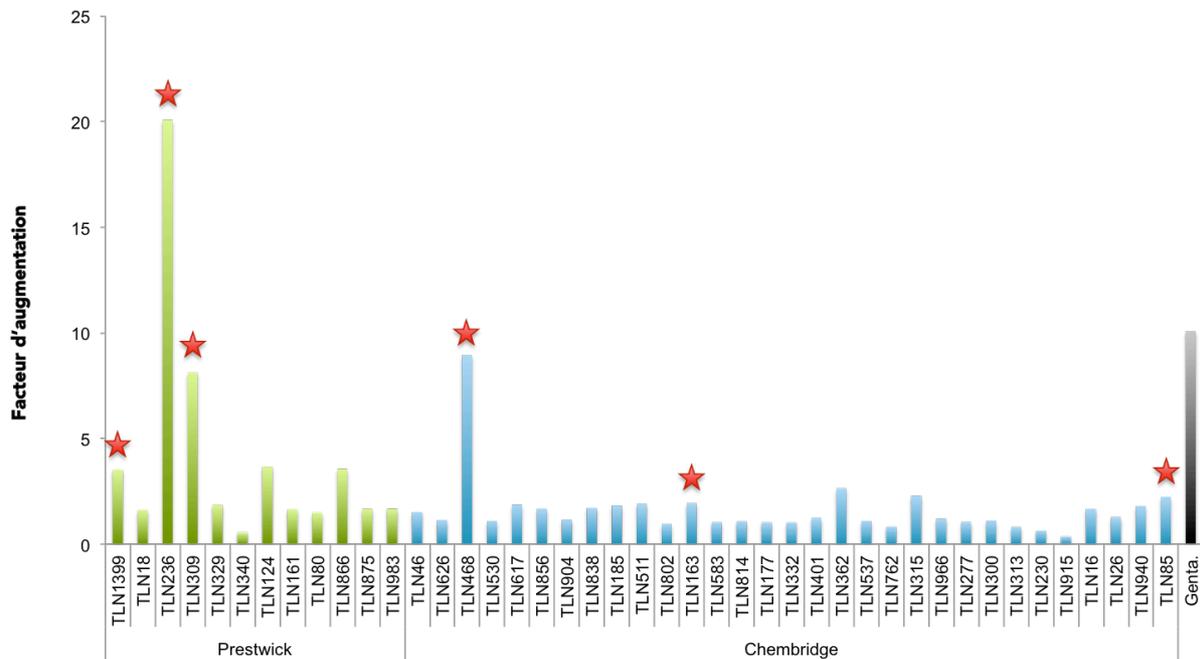
Le calcul du SSMD va donc permettre de donner un score correspondant à l'effet du composé testé par rapport à l'effet médian de l'ensemble des composés d'une plaque (Figure 14) <sup>453</sup>. En revanche, ceci n'est vrai que si la distribution des molécules est aléatoire et que les « hits » ne sont donc pas concentrés sur une même plaque. L'effet des composés est ainsi catégorisé en 11 classes, de l'absence d'effet (SSMD = 0) à un effet extrêmement fort (SSMD  $\geq$  5).

Nous avons alors choisi de ne sélectionner que les molécules dont le facteur d'augmentation entre la condition traitée et la médiane des contrôles négatifs est supérieur ou égal à 1,4 ou dont le SSMD est supérieur ou égal à 2. Nous avons ainsi sélectionné 465 molécules qui ont été recriblées (Figure 15).

## H. Crible secondaire

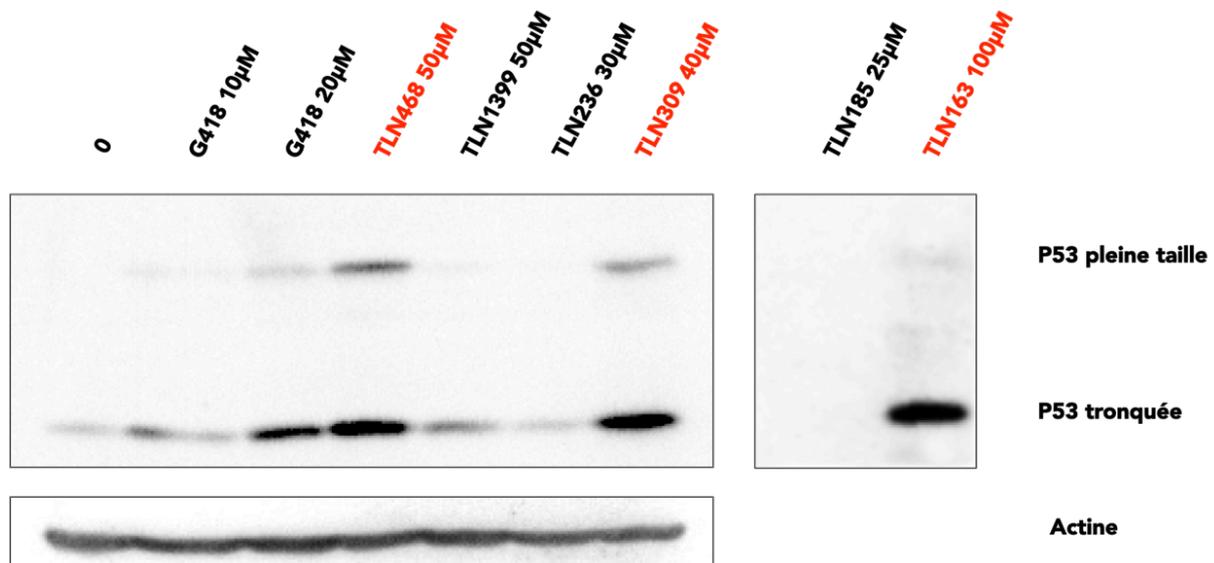
Lors du crible secondaire, nous avons cherché à valider les hits primaires en procédant à un nouveau crible en duplicas des 465 molécules sélectionnées. Nous avons tout d'abord vérifié que ces molécules n'étaient pas des faux positifs capables d'augmenter artificiellement l'émission de photon. Pour cela, les cellules NIH3T3 qui ne contiennent pas le système rapporteur ont été traitées avec ces mêmes molécules (n=1). Nous avons pu montrer qu'aucune des molécules sélectionnées n'induisait de signal indépendamment du gène rapporteur.

L'ensemble des plaques a passé les tests de contrôle qualité après normalisation des résultats. Le facteur d'augmentation a été calculé indépendamment pour les deux répliques. Le SSMD ici ne peut être utilisé, car la distribution n'est plus aléatoire et que les deux répliques ne suffisent pas à calculer la variabilité de l'effet des molécules. Les molécules ont ensuite été sélectionnées sur leur capacité à induire une augmentation de l'activité luciférase d'au moins un facteur 2 par rapport à la médiane des contrôles négatifs dans les deux répliques. Les 48 molécules qui ont été retenues ont été commandées chez leur fournisseur respectif, toutefois, seulement 43 molécules ont pu être obtenues. Parmi elles, 12 proviennent de la chimiothèque Prestwick et 31 de la Chembridge.



**Figure 16 : Quantification de l'effet des 43 molécules hits sur la translecture**

La capacité des molécules à induire une augmentation de l'activité luciférase par translecture du codon stop a été mesurée sur le plasmide double rapporteur pAC99-R213X. Les cellules ont été traitées durant 24h avec le contrôle gentamicine et les 43 molécules sélectionnées lors du crible à grande échelle (TLN) (n=3). Ces 43 composés ont été testés à 50 µM final et la gentamicine à 2,5 mM. Les mesures d'activités luciférase en présence et en absence de molécules permettent de déterminer le facteur d'augmentation. Les étoiles rouges représentent les molécules sélectionnées pour avoir une capacité d'induction du niveau de translecture supérieur à 2 et une toxicité modérée ou faible.



**Figure 17 : Capacité des molécules d'intérêts à induire la réexpression d'une protéine p53 complète**

Western blots réalisés à partir d'extraits protéiques de la lignée HDQ-P1 qui possède la mutation non-sens endogène P53-R213X. 32 µg (WB de gauche) ou 45 µg (WB de droite) d'extraits protéiques ont été déposés dans chaque puits. La protéine est détectée par l'anticorps p53-DO1 reconnaissant la région N-terminale. Le traitement par le G418 à 20 µM permet de réexprimer la forme complète de la protéine p53 et de stabiliser la forme tronquée. En présence de 3 des 6 molécules TLN testées (rouge), la forme pleine taille est détectée. Cette dernière est associée à une importante bande correspondant à la protéine p53 tronquée. Le contrôle actine n'a été obtenu que dans un seul des deux western blot.

## I. Caractérisation des molécules hits

### a. Capacité à induire de la translecture

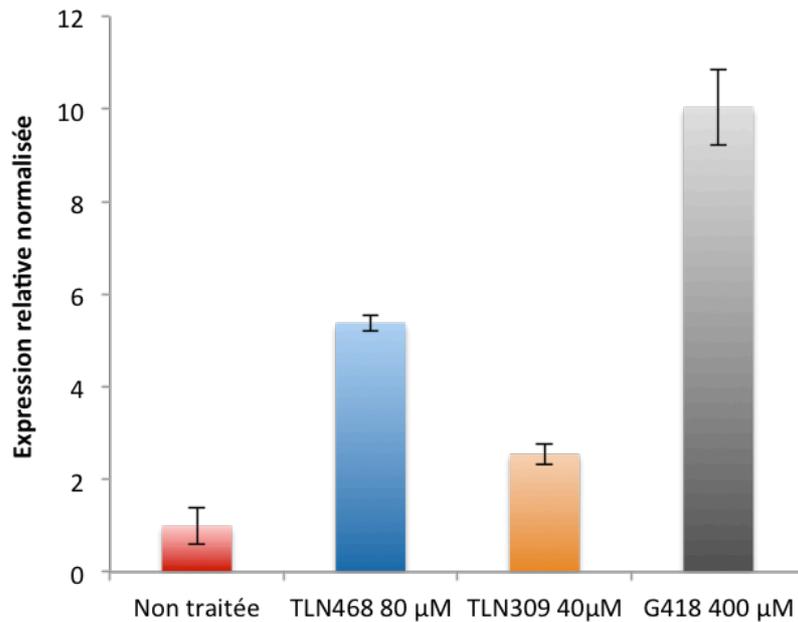
Les 43 composés ont été repris à 10 mM dans du DMSO et leur capacité à induire de la translecture a été de nouveau quantifiée grâce à un système rapporteur différent de celui du crible : le vecteur double rapporteur pAC99-R213X à 50  $\mu$ M finale pendant 24 heures (n=1) (Figure 16). Les résultats montrent que la majorité des molécules n'induisent pas d'augmentation de l'activité luciférase firefly et donc pas de translecture à la dose testée. En revanche, 10 molécules ont été identifiées pour augmenter l'activité luciférase d'un facteur supérieur à 2. L'analyse de leur niveau d'activité  $\beta$ -galactosidase montre que la moitié d'entre elles présentent une importante toxicité.

Pour la suite, nous nous sommes focalisés sur les molécules présentant une toxicité modérée ou faible et dont le facteur d'augmentation est supérieur à 2, soit TLN1399, 236, 309, 468 et 163 et sur la TLN185 qui a concentration plus faible présente une toxicité modérée et un facteur d'augmentation supérieur à 3. De plus, TLN185 et TLN468 sont les molécules qui induisaient le plus d'activité luciférase lors du crible primaire et secondaire. Chacune des 6 molécules a également été testée à plusieurs concentrations sur ce même système rapporteur pour définir la concentration associée au meilleur rapport activité/toxicité.

### b. Capacité à réexprimer une protéine de pleine taille

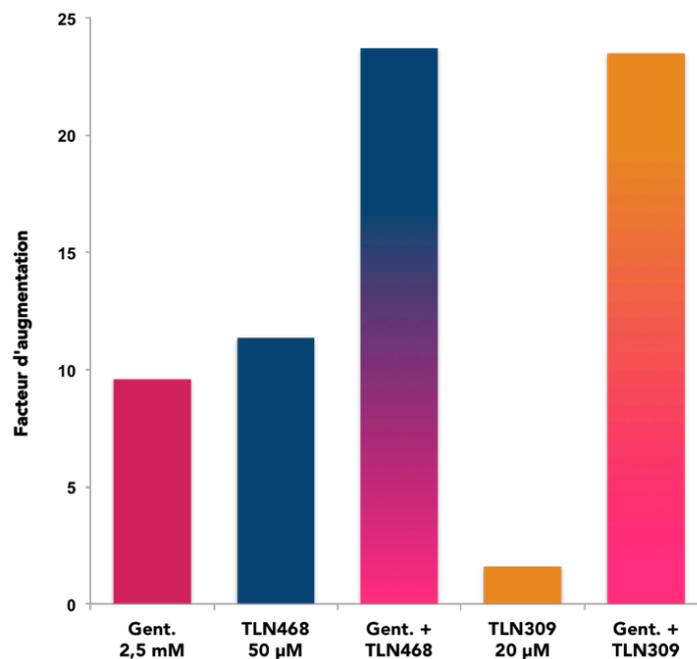
Nous avons cherché à déterminer si le taux de translecture induit par ces 6 molécules permettait de réexprimer une protéine p53 entière. Des expériences de western blot ont été réalisées à partir d'extraits protéiques de la lignée humaine HDQ-P1 qui a été établie à partir d'un carcinome primaire du sein (Figure 17)<sup>454</sup>. Les cellules HDQ-P1 sont homozygotes pour la mutation R213X du gène *P53*. Cette mutation est localisée 50 nucléotides en amont de la dernière jonction exon-exon, le transcrit généré est donc sensible à la dégradation par la voie NMD. Les cellules HDQ-P1 ont été traitées pendant 48 heures avec les 6 molécules à la concentration optimale définie précédemment.

En absence de traitement, une seule bande d'environ 30 kDa est détectée et correspond à la forme tronquée de la protéine. La faible intensité de cette bande est corrélée avec le faible niveau d'ARNm *P53* ciblé par la voie NMD. Après traitement avec le contrôle G418, une bande de 53 kDa associée à la forme pleine-taille de la protéine est détectée et son intensité varie en fonction de la dose testée.



**Figure 18 : Capacité des molécules à stabiliser le transcrit mutant P53**

Les cellules HDQ-P1 portant la mutation non-sens R213X endogène ont été traitées ou non avec le G418 (400 µM), TLN468 (80 µM) ou TLN309 (40 µM) pendant 48 heures. Les niveaux d'ARNm de p53 mutant ont été déterminés par PCR quantitative (n=3). Les résultats de chaque condition sont exprimés par rapport à la quantité relative normalisée d'ARNm en l'absence de traitement fixée à 1.



**Figure 19 : Effet sur la translecture de la synergie des molécules TLN avec la gentamicine**

La synergie des molécules TLN avec la gentamicine a été mesurée sur les cellules NIH3T3 transfectées par le vecteur pAC99-R213X (n=2). Les cellules ont ensuite été traitées durant 24h avec chacune des molécules TLN, soit seules, soit en synergie avec la gentamicine. La gentamicine a été utilisée à 2,5 mM finale, TLN468 à 50 µM et TLN309 à 20 µM. Le facteur d'augmentation a été calculé en faisant le rapport de l'activité luciférase en présence et en absence de molécules.

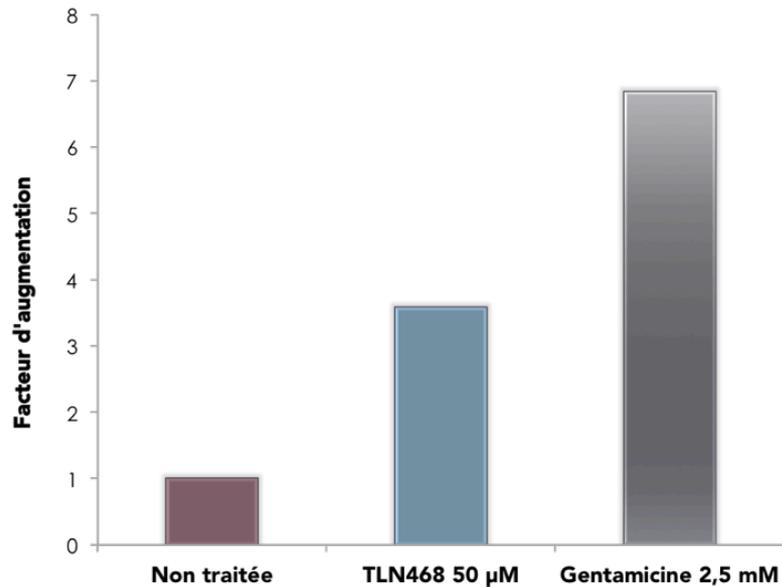
La forme tronquée de la protéine p53 est retrouvée, quelle que soit la molécule TLN utilisée, excepté pour TLN185 dont la toxicité est trop élevée sur cette lignée cellulaire. La forme pleine taille est retrouvée après traitement par TLN468, 309 et 163, ce qui indique que ces molécules sont capables de restaurer l'expression de la protéine p53 complète. De plus, le niveau de p53 tronquée est bien plus élevé lorsque les cellules sont traitées avec ces trois molécules, ce qui suggère qu'elles seraient capables de stabiliser l'ARNm ou la forme tronquée de la protéine. Toutefois, le niveau de protéine réexprimée en présence de TLN163 étant particulièrement faible malgré une concentration élevée, nous n'avons pas poursuivi la caractérisation de cette molécule (Figure 17).

#### *c. Capacité à stabiliser le transcrit P53*

Nous savons que les inducteurs de translecture déjà caractérisés comme le G418 ont un effet stabilisateur sur les ARNm mutants présentant un PTC et qu'il existe un effet antagoniste entre la translecture et la voie NMD<sup>381,455</sup>. Dans le but de tester l'effet des nouvelles molécules sur la stabilité des ARNm mutants, nous avons analysé par RT-PCR quantitative l'expression relative normalisée des transcrits P53 en présence et en de TLN468 (80 µM) et TLN309 (40 µM) sur les cellules HDQ-P1 en utilisant le G418 comme contrôle (n=3) (Figure 18). Comme attendu, le traitement des cellules par le G418 augmente les niveaux de transcrit P53 R213X d'un facteur 10. L'ajout de TLN309 sur les cellules HDQ-P1 augmente d'un facteur 2,5 le niveau de transcrit et la TLN468 permet de stabiliser d'un facteur 5,4. Ces deux molécules permettraient donc également de stabiliser les transcrits mutants probablement via leur effet sur la translecture du PTC.

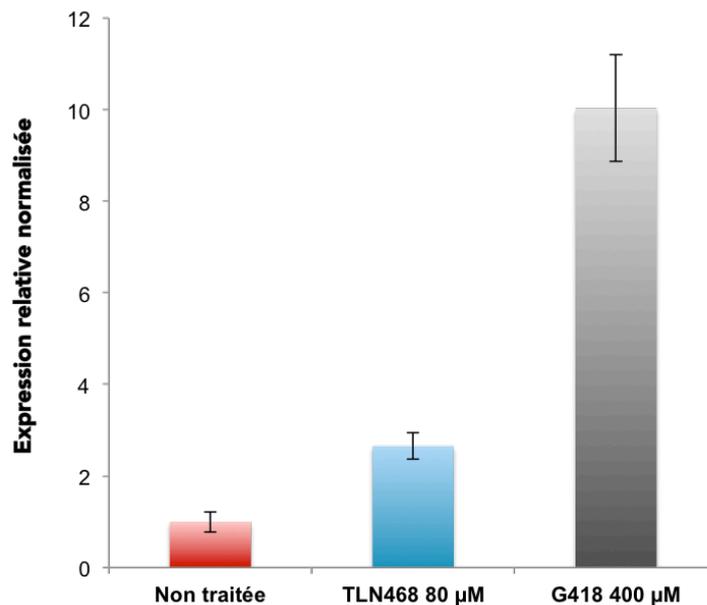
#### *d. Synergie des molécules TLN avec un aminoglycoside*

N'ayant aucun a priori concernant le mode d'action de ces nouvelles molécules, nous avons analysé leur synergie en présence d'un inducteur de translecture de la famille des aminoglycosides, la gentamicine (n=2) (Figure 19). Les cellules NIH3T3 transfectées par le pAC99-R213X ont été traitées par les molécules TLN468 (50 µM) et TLN309 (20 µM) couplées ou non à la gentamicine (2,5 mM) durant 24 heures. Les résultats montrent que le facteur d'augmentation induit par TLN468 s'additionne à celui induit par la gentamicine lorsque les cellules sont traitées par les deux molécules. En revanche, le facteur d'augmentation induit par TLN309 passe de 1,6 à 23 lorsque la gentamicine est ajoutée, mettant en évidence un effet synergique. Ces résultats indiquent que le mode d'action de TLN309 diffère de celui des aminoglycosides. Pour la suite, nous nous sommes concentrés sur la caractérisation de la molécule TLN468 dont le rapport efficacité de translecture sur toxicité est bien meilleur que TLN309.



**Figure 20 : Capacité de TLN468 à induire la réexpression d'une protéine p53 fonctionnelle**

Les cellules humaines H1299 (*P53*<sup>-/-</sup>) sont cotransfectées par les vecteurs pCMV-R213X, pCMV-LacZ et p53BS-luc. Le vecteur pCMV-R213X exprime le gène *P53* portant la mutation non-sens R213X. Le vecteur p53BS-luc exprime le gène de la luciférase sous contrôle d'un promoteur p53-dépendant et permet de détecter la synthèse d'une p53 transcriptionnellement active. Le vecteur pCMVLacZ permet de normaliser les résultats. Le composé TLN468 et la gentamicine ont été ajoutés 24 heures après transfection à des concentrations de 50 µM et 2,5 mM respectivement (n=1). Les dosages des activités luciférase en présence et en absence de traitement permettent de déterminer le facteur d'augmentation de l'activité luciférase.



**Figure 21 : Capacité de la protéine p53 réexprimée à réguler la niveau de transcrits du gène Bax.**

Les cellules HDQ-P1 portant la mutation non-sens R213X endogène ont été traitées ou non avec le G418 (400 µM), la TLN468 (80 µM) ou la TLN309 (40 µM) pendant 48 heures. Les niveaux d'ARNm du gène Bax, une des cibles majeures de la p53, ont été déterminés par PCR quantitative (n=3). Les résultats de chaque condition sont exprimés par rapport à la quantité relative normalisée d'ARNm en l'absence de traitement fixée à 1.

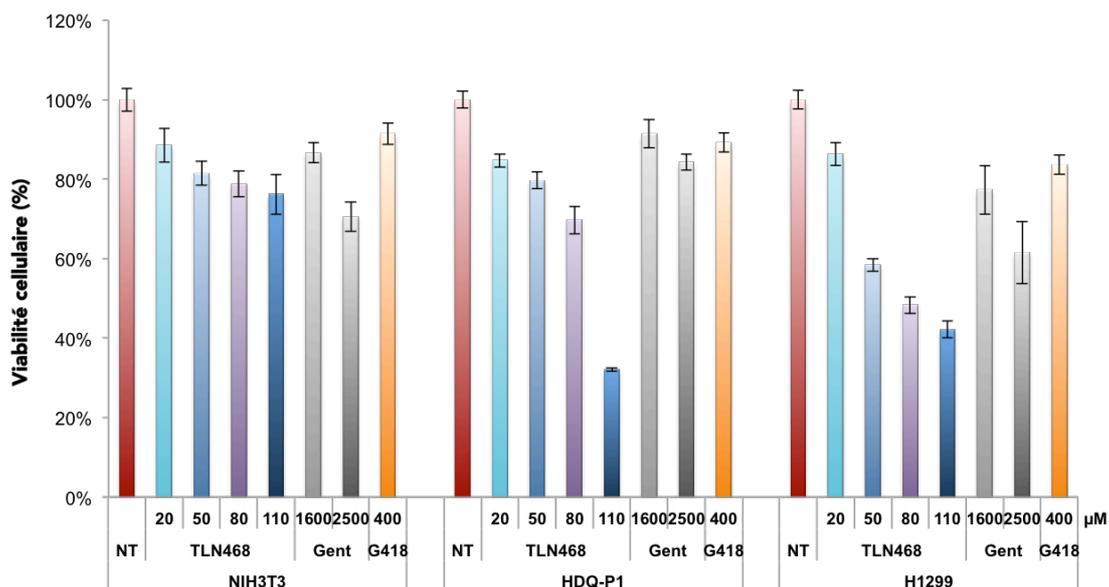
*e. Capacité de TLN468 à induire la réexpression d'une protéine p53 active*

La capacité de transactivation de la protéine p53 sauvage est à la base de sa fonction de gène suppresseur de tumeur. Nous avons analysé si la protéine p53 complète produite après traitement par la molécule TLN468 était un facteur de transcription actif (n=1) (Figure 20). Pour cela, nous avons utilisé la lignée cellulaire H1299 qui dérive d'un carcinome des poumons et qui porte une délétion homozygote du gène *P53*. Les cellules ont été cotransfectées avec un vecteur portant le gène *P53* muté à la position 213 (cDNA) sous contrôle d'un promoteur fort (pCMV-R213X), un vecteur exprimant constitutivement le gène codant la  $\beta$ -galactosidase (pCMV-LacZ) pour normaliser les résultats et un vecteur rapporteur contenant sept sites de liaisons pour la protéine p53 situés en amont du gène de la luciférase firefly (p53BS-luc). Puis les cellules ont été traitées pendant 24h avec la TLN468 en utilisant la gentamicine en contrôle. Ainsi, seule une protéine p53 active peut induire l'expression du gène rapporteur luciférase. Le traitement avec la gentamicine induit l'expression de la luciférase firefly d'un facteur 7 par rapport à la condition non traitée. En présence de TLN468, l'expression de la luciférase est 3,5 fois plus élevée que le niveau de base. La molécule TLN468 est donc capable d'induire la réexpression d'une protéine p53 active.

*f. Capacité de la protéine p53 réexprimée à activer ses gènes cibles*

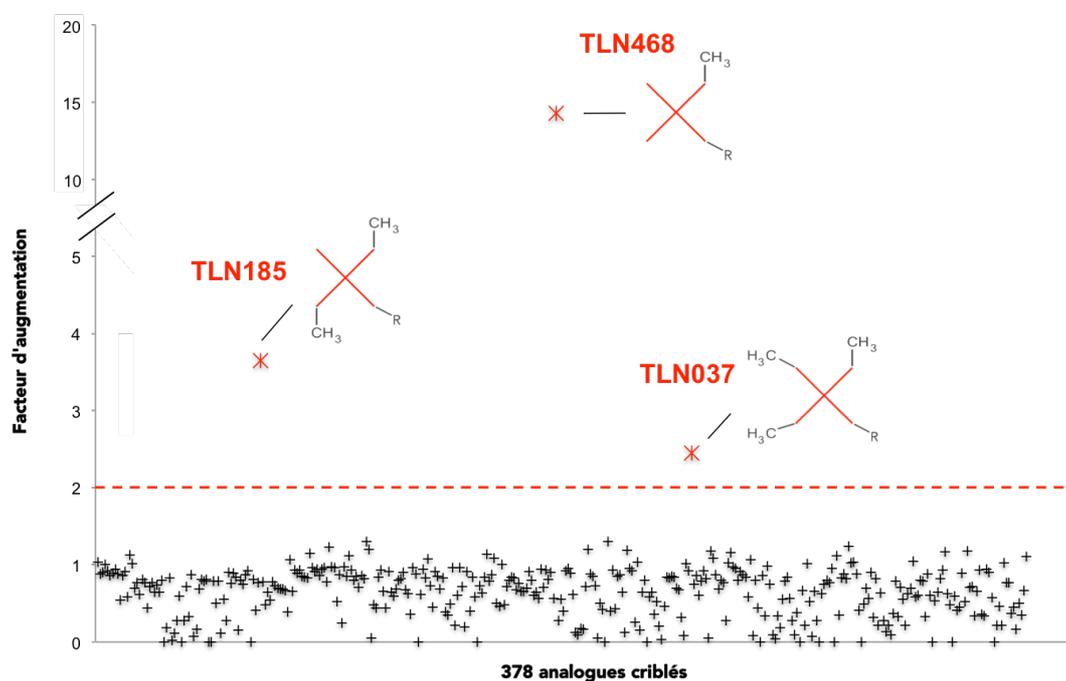
Nous avons ensuite étudié si la protéine p53 complète, réexprimée après traitement par la molécule TLN468, régule une de ses cibles transcriptionnelles majeures, le gène *Bax* qui un gène proapoptotique bien caractérisé. Des essais ont été réalisés en PCR quantitative sur les cellules HDQ-P1 en présence ou non de G418 (400  $\mu$ M) ou de la molécule TLN468 (80  $\mu$ M) durant 48 heures (n=3) (Figure 21). Dans cet essai, l'expression relative normalisée des transcrits est quantifiée par rapport aux cellules non traitées (normalisé à 1). Le traitement des cellules par le G418 augmente les niveaux de transcrit P

53 R213X d'un facteur 10. En présence de TLN468, la quantité de transcrits *Bax* est multipliée par 2,5. Cette augmentation est corrélée avec la quantité de protéine réexprimée active, ce qui suggère que la TLN468 permet de réexprimer suffisamment de protéine p53 active pour activer l'expression du gène *Bax*.



**Figure 22 : Viabilité cellulaire des lignées NIH3T3, HDQ-P1 et H1299 après traitement par la molécule TLN468**

Les composés G418, gentamicine et TLN468 sont ajoutés aux lignées cellulaires NIH3T3, HDQ-P1 et H1299 durant 24h (n=4). Le WST1 est alors ajouté sur les cellules et l'absorbance est quantifiée. Ce composé est réduit en formazan (détectable à 440nm), qui s'accumule de façon proportionnelle au nombre de cellules vivantes. La viabilité cellulaire est alors déterminée en faisant le rapport des mesures d'absorbance dans la condition traitée par rapport à la médiane des absorbances de la condition contrôle.



**Figure 23 : Crible des analogues de TLN468**

Résultats du crible de 378 analogues structuraux de la structure de base de la TLN468 représentée ici par une croix rouge pour des raisons de confidentialité. Ce crible a été réalisé sur les cellules 3T3pML213 selon le même protocole que le crible haut débit. Les résultats sont normalisés et les contrôles validés (n=1). Les molécules induisant une augmentation de l'activité luciférase d'un facteur supérieure à 2 sont indiqués par une petit croix rouge et leur structure chimique est schématisée sur le graphique.

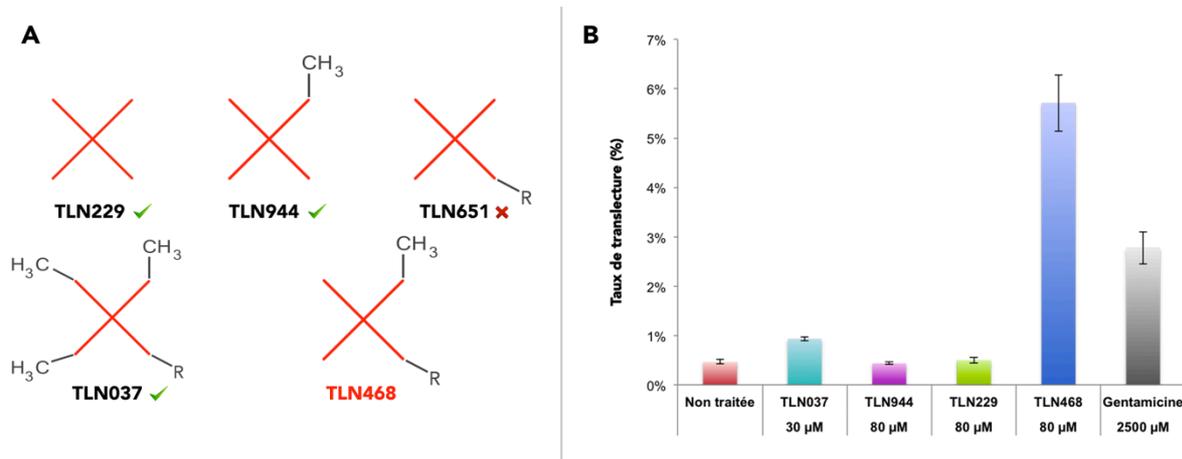
### *g. Toxicité cellulaire de TLN468*

Afin de déterminer avec précision la toxicité du composé TLN468, nous avons évalué son effet sur la viabilité des cellules NIH3T3, HDQ-P1 et H1299 en utilisant un kit de dosage de la prolifération et de la viabilité cellulaire (Roche, WST1). Les cellules ont été traitées pendant 24 heures avec les contrôles gentamicine et G418 aux doses auxquelles elles sont décrites pour induire de la translecture avec une toxicité modérée et avec la molécule TLN468 sur une gamme de concentration allant de 30  $\mu\text{M}$  à 110  $\mu\text{M}$  (Figure 22). Les résultats obtenus sur la lignée NIH3T3 montrent que la concentration de la TLN468 utilisée dans nos essais (80  $\mu\text{M}$ ) est associée à une faible toxicité. Les cellules HDQ-P1 semblent plus sensibles au traitement par la molécule TLN468, principalement à 110  $\mu\text{M}$  où la toxicité est supérieure à 60%. Ce phénomène est encore plus marqué sur la lignée H1299 où la toxicité apparaît clairement dès la dose 50  $\mu\text{M}$  de TLN468 et 2,5 mM de gentamicine. La molécule TLN468 présente une toxicité très variable selon la lignée cellulaire traitée. L'optimisation de cette molécule, par modification ou par recherche d'analogues déjà existants, pourrait permettre de diminuer sa cytotoxicité.

## **J. Les analogues de TLN468**

### *a. Recherche et criblage de 378 analogues*

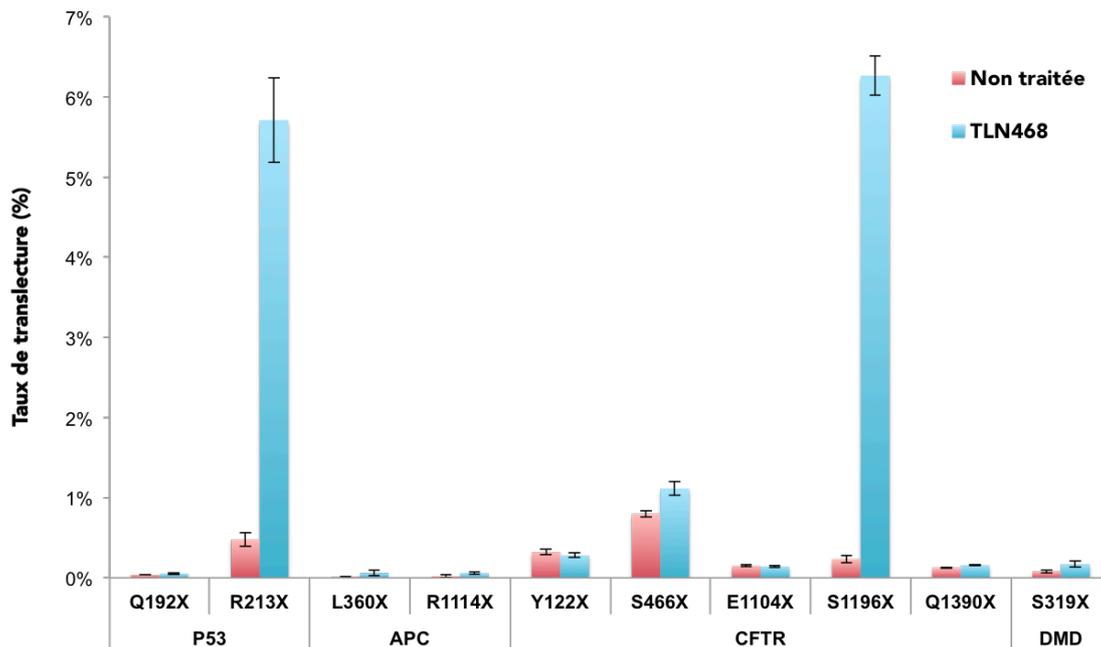
La toxicité de la molécule TLN468 est un frein important à son utilisation. Nous avons donc entrepris de rechercher et tester la capacité de translecture des analogues de la TLN468. Une première approche consiste à cribler informatiquement la base de données de plus de 80 000 molécules de l'équipe de J-C Cintrat à partir de la structure chimique de base de TLN468 et de tester les molécules analogues sur les cellules 313pML213. 378 composés ont ainsi été identifiés et préparés en plaque 96 pour être criblés selon le même protocole que précédemment. Une fois les contrôles validés, les résultats ont été analysés et 3 molécules induisant une augmentation de l'activité luciférase supérieure à 2 ont été identifiées (Figure 23). Parmi ces 3 molécules, nous avons retrouvé la TLN468 et TLN185 et une nouvelle molécule, TLN037 dont la structure chimique est très similaire à celle de TLN468 et TLN185. Dans une seconde approche, nous avons cherché à nous procurer les molécules constituées de la structure de base, associée ou non à chacun des deux groupements de TLN468. Nous avons identifié 4 molécules analogues, appelées TLN229, TLN944, TLN037 et TLN651 (Figure 24-A). TLN651 est toujours en cours de synthèse et n'a donc pas pu être testé ici.



**Figure 24 : Schéma structurel et transfecture des molécules analogues à TLN468**

**A** – Structure chimique partielle des analogues de TLN468 : TLN229, TLN944, TLN651 et TLN037. La structure de base est indiquée par une croix rouge.

**B** – La capacité des molécules analogues à induire une augmentation du taux de translecture a été mesurée sur le plasmide double rapporteur pAC99-R213X dans les cellules NIH3T3. Les cellules ont été traitées ou non durant 24h avec le contrôle gentamicine, TLN468 et ses 3 analogues TLN037, TLN944, TKN229 (n=3). Ces 3 composés ont été testés à la concentration la plus élevée à laquelle elles présentent une toxicité compatible avec l'expérience. La TLN651 étant toujours en cours de synthèse, son taux de translecture sera mesurée ultérieurement.



**Figure 25 : Efficacité de la translecture induite par la TLN468 sur différentes mutations non-sens**

Le taux de translecture induit par la molécule TLN468 a été mesurée sur plusieurs mutations non-sens impliquées dans des maladies humaines tels que le cancer (*P53* et *APC*), la mucoviscidose (*CFTR*) et la myopathie de duchenne<sup>22</sup>. Les cellules NIH3T3 ont été transfectées par les vecteurs doubles rapporteurs portant ces mutations ou par le vecteur contrôle pour lequel le codon stop est remplacé par un codon sens. Les cellules ont été traitées ou non par la molécule TLN468 à 50 µM (S466X, E1104X et Q1390X) ou 80 µM pendant 24h (n=6). Les activités luciférase et β-galactosidase sont dosées et le taux de translecture est calculé en faisant le rapport de l'activité luciférase normalisée du vecteur test sur le vecteur contrôle.

### *b. Efficacité de translecture des analogues de TLN468*

Pour des raisons de confidentialité, la structure de base de la TLN468 est représentée par un X sur lequel sont liés un groupement méthyle et un groupement R. La capacité de translecture des molécules TLN037, TLN229 et TLN944 a été testée sur le système double rapporteur pAC99-R213X dans les cellules NIH3T3 (Figure 24-B). Le taux de translecture de TLN944 et TLN229 n'est pas significativement différent des cellules non traitées. A l'inverse, TLN037 induit une augmentation du taux de translecture d'un facteur 2 par rapport au taux de base. Toutefois, à concentration identique à TLN468, l'induction de translecture de TLN037 n'a pas pu être quantifiée, car sa toxicité était supérieure à 80%. Par conséquent, si TLN037 semble être capable d'induire de la translecture, sa toxicité, probablement liée à la présence des deux groupements méthyles absents de TLN468, est trop importante.

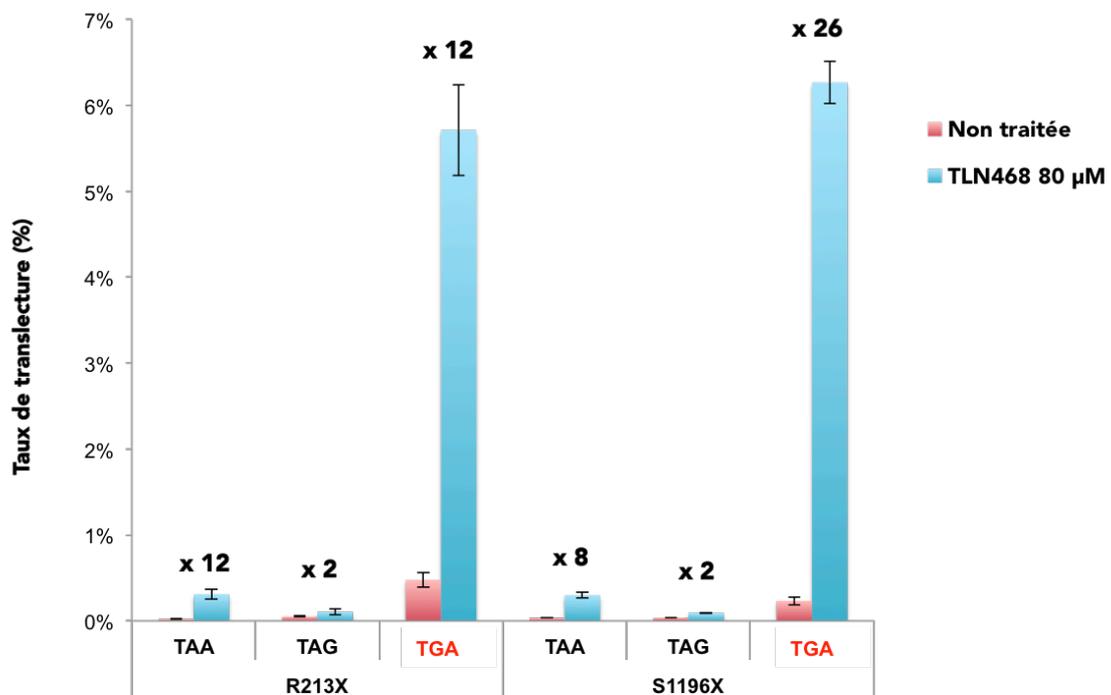
### **K. Spécificité de la molécule TLN468**

La molécule TLN468 a été identifiée comme inducteur de translecture sur la mutation R213X. Nous avons cherché à évaluer l'effet de cette nouvelle molécule sur d'autres mutations non-sens impliquées dans diverses maladies (Tableau 1). Nous avons sélectionné des mutations présentant une forte suppression traductionnelle en présence de gentamicine, mutation dite « bonne répondeuse » comme la S319X, ainsi que des mutations peu sensibles à l'effet de la gentamicine comme la Q1390. Les séquences cibles de la translecture, constituée du codon stop et du contexte nucléotidique naturel, ont été insérées dans le vecteur double rapporteur pAC99 avant mon arrivée. Nous avons ainsi pu quantifier le taux de translecture en présence et en absence de TLN468 (50 ou 80  $\mu$ M) dans les cellules NIH3T3 (n=6) (Figure 25). Les résultats montrent que seules les mutations R213X et S1196X répondent au traitement par la molécule TLN468 avec une efficacité de translecture équivalente. Ces résultats suggèrent que le spectre d'action de la molécule TLN468 est sensiblement différent de celui des aminoglycosides inducteurs de translecture.

Maladie	Gène	Mutation	Séquence 5' -> 3'
Cancer	<i>P53</i>	Q192X R213X	GCC CCT CCT <b>TAG</b> CAT CTT ATC AAC ACT TTT <b>TGA</b> CAT AGT GTG
	<i>APC</i>	L360X R1114X	ATC CAG CTT <b>TGA</b> CAT GGC AAT GAA ACA AAT <b>TGA</b> GTG GGT TCT
Mucoviscidose	<i>CFTR</i>	Y122X	TCT ATC GCG <b>TAA</b> CTA GGC ATA
S466X		GGC AAG ACT <b>TGA</b> CTT CTA ATG	
S1196X		ATT GAG AAT <b>TGA</b> CAC GTG AAG	
E1104X		ATG AGA ATA <b>TAA</b> ATG ATT TTT	
Q1390X		ACT CTA AAA <b>TAA</b> GCA TTT GCT	
Myopathie de Duchenne	<i>DMD</i>	S319X	CCA TTT CCT <b>TGA</b> CAG CAT TTG

**Tableau 1 : Séquences dont la translecture a été testée en présence de la TLN468**

La séquence de la cible de translecture (codon stop dans son contexte nucléotidique naturel) a été insérée dans le vecteur double rapporteur pAC99 afin de quantifier le taux de translecture dans la lignée cellulaire NIH3T3. Les mutations sont appelées par leur position et par la nature de l'acide aminé de la séquence sauvage. Le codon stop est indiqué en rouge.



**Figure 26 : Effet de la nature du codon stop sur la translecture induite par TLN468**

Le taux de translecture induit par la molécule TLN468 a été mesurée sur les mutations non-sens R213X et S1196X. Pour ces deux mutations, le codon stop d'origine, indiqué en rouge, a été remplacé par les deux autres stop en conservant le contexte nucléotidique environnant. Ces séquences ont été clonées dans le vecteur double-rapporteur pAC99. Les cellules NIH3T3 ont été transfectées par les vecteurs pAC99 portant ces mutations ou par le vecteur en phase (pAC-TQ) et ont été traitées ou non par la molécule TLN468 à 80 µM pendant 24h (n=6). Les activités luciférase et β-galactosidase sont dosées et le taux de translecture est calculé en faisant le rapport de l'activité luciférase normalisée du vecteur portant la mutation non-sens sur le vecteur en phase. Le facteur d'augmentation entre la condition traitée et la condition contrôle est indiqué au dessus de chaque histogramme.

## L. Effet de la nature du codon stop

Nous avons vu que la nature du codon stop influence fortement l'efficacité de translecture induite par les aminoglycosides. Le codon stop TGA permettant le plus de translecture, suivi de UAG et de UAA. Bien que le mode d'action de TLN468 semble différent de la gentamicine, nous avons vérifié si cet ordre était maintenu avec la molécule TLN468 sur les mutations non-sens R213X et S1196X. Pour chacune de ces mutations, le codon stop d'origine a été remplacé par les deux autres stop en conservant le contexte nucléotidique environnant (Figure 27) Ces séquences ont été clonées dans le système double rapporteur pAC99 puis utilisées pour transfecter les cellules NIH3T3. Ces cellules ont ensuite été traitées ou non pendant 24 heures avec la molécule TLN468 à 80  $\mu$ M. Le taux de translecture et le facteur d'augmentation entre la condition traitée et non traitée sont systématiquement plus élevés avec le codon stop TGA. En revanche, la hiérarchie connue n'est plus respectée pour les deux autres stop puisque le taux de translecture induit par la TLN468 est plus important sur le stop TAA que sur le stop TAG. Néanmoins, ce taux ne dépasse pas 0,3%, ce qui reste marginal comparativement aux 5,5 et 6% de translecture obtenus avec le stop TGA sur les séquences R213X et S1196X.

# ***Discussion***

***Tout problème simple a une solution complexe  
qui ne fonctionne pas  
(Olivier Lockert)***

## Discussion

La recherche de molécules capables d'induire la suppression des codons stop prématurés est une stratégie ayant fait l'objet d'un nombre grandissant d'études depuis la publication princeps en 1996 sur le traitement de la CFTR par un aminoglycoside. En effet, la présence de nombreuses maladies à codons stop et la toxicité des aminoglycosides ont conduit la communauté scientifique à développer de nouvelles approches permettant d'identifier d'autres inducteurs de translecture.

### 1. Les approches développées permettant l'identification de nouvelles molécules inductrices de translecture

#### ★ Le design rationnel

La première approche est basée sur la modification de molécules inductrices de translecture déjà existantes et a été appliquée aux antibiotiques de la famille des aminoglycosides par le laboratoire de Timor Baasov. La généticine est décrite pour être l'aminoglycoside le plus efficace sur la translecture et la paromomycine, pour présenter le moins de toxicité. La combinaison de ces deux molécules a conduit à l'obtention d'un premier dérivé, le NB30. Plusieurs groupements ont été greffés sur cette molécule, améliorant ainsi sa capacité à supprimer les codons stop prématurés et son affinité pour le ribosome cytoplasmique eucaryote. J'ai été amené à tester plusieurs de ces dérivés et à caractériser la molécule NB124 qui présente le meilleur rapport efficacité sur toxicité.

#### ★ Le criblage à grande échelle

La seconde approche, qui est aussi la plus répandue, consiste à cribler des chimiothèques de molécules sur un système biologique permettant d'identifier les molécules actives. En effet, la majorité des molécules inductrices de translecture découvertes ces dernières années sont issues de cribles à grande échelle (HTS). Toutefois, l'efficacité de ces composés est encore à démontrer. De plus, le défaut majeur des molécules déjà caractérisées est leur forte toxicité.

Pour ces raisons, nous avons réalisé un crible à grande échelle sans a priori sur le mode d'action des molécules recherchées (collaboration avec le laboratoire de J-C. Cintrat).

Le principe de ce crible repose sur le gène rapporteur de luciférase de *Metridia* interrompu par un PTC afin de conditionner son expression à un événement de translecture. Les molécules capables de supprimer ce PTC permettent la synthèse d'une luciférase complète. Contrairement aux luciférases Firefly ou *Renilla*, la luciférase de *Metridia* a la particularité d'être sécrétée directement dans le milieu de culture, ce qui nous a permis de grandement simplifier le protocole du crible en évitant l'étape de lyse des cellules. Ainsi, l'ensemencement, l'ajout des molécules et le dosage de l'activité luciférase sont réalisés dans le même puits.

Le succès du crible à grande échelle repose avant tout sur le développement d'un test biologique pertinent, limitant ainsi la proportion de faux positifs. Ce développement nécessite entre autres de valider les contrôles qualité. Ce fut particulièrement complexe dans notre cas, car ces tests requièrent à la fois une faible dispersion des contrôles positifs et négatifs et un écart entre leur médiane très important. Il est donc préconisé d'utiliser un contrôle positif dont l'effet est le plus marqué possible. Dans notre cas le G418 est la molécule dont l'effet sur la translecture est le plus fort, néanmoins son utilisation n'a pas été possible, car le plasmide contenant le système rapporteur porte le gène de résistance à la généticine. Nous avons donc utilisé la gentamicine même si son effet sur la translecture est un peu inférieur. De plus, la luciférase de *Metridia* est une enzyme dont la forte intensité de signal est associée à une cinétique de décroissance très rapide, dite « flash ». Ces caractéristiques facilitent sa détection, mais son instabilité dans le temps est responsable de biais de mesures au cours des dosages qui doivent être maîtrisés et corrigés. Le substrat de la luciférase de *Metridia*, la coelenterazine, est sensible à la lumière et à l'oxygène, ce qui nous a contraints à optimiser sa stabilisation pour améliorer la reproductibilité des résultats. Plus globalement, les cribles utilisant un gène rapporteur codant une enzyme sont sujets à la présence de faux positifs liés à la stabilisation de l'enzyme. La luciférase est une des enzymes les plus utilisées de par ses caractéristiques de bioluminescence et n'est pas exempte de ce défaut. En effet, de nombreux papiers ont rapporté la présence de faux positifs liés à son utilisation. C'est pourquoi nous avons réalisé, dès le crible secondaire, un essai sur des cellules ne contenant pas le gène rapporteur *metluc* muté par un codon stop prématuré. Ce test permet ainsi de discriminer les molécules capables de stabiliser l'enzyme. L'ensemble de ces optimisations nous a permis de valider notre test biologique, le crible à grande échelle et de sélectionner 48 molécules hits.

Je vais maintenant m'intéresser plus précisément aux deux molécules identifiées durant ma thèse dont j'ai caractérisé l'activité de translecture.

## 2. Le NB124 et la TLN468 induisent la translecture des codons stop

### ★ Le NB124 est le dérivé le plus efficace sur la translecture

Parmi les cinq dérivés d'aminoglycosides qui nous ont été transmis, NB74, NB84, NB122, NB124 et NB128 il apparaît clairement que la molécule NB124 présente la meilleure efficacité de translecture (Figure 1 – Article). D'un point de vue de sa structure chimique, le NB124 ne diffère du NB128 que par l'absence du groupement AHB. Cette différence est associée à un taux de translecture du NB128 1,4 fois plus faible que le NB124 à la dose la plus élevée. Le groupement AHB a pourtant été ajouté dans la première génération de dérivés pour améliorer l'efficacité de translecture et diminuer la toxicité, ce que confirment nos résultats obtenus entre le NB74 et le NB84. Il semble donc que le groupement AHB antagonise l'effet du 5''-(R)-Me ajouté sur l'amine du cycle III. En revanche, aucune amélioration du taux de translecture n'est observée avec le NB122 qui possède ce même groupement méthyle en configuration S.

Il est connu que l'environnement nucléotidique du codon stop influence fortement l'efficacité de terminaison. Nous avons donc voulu tester son impact sur la translecture induite par le NB124 en analysant les 5 autres mutations non-sens les plus fréquemment retrouvées dans les gènes suppresseurs de tumeurs *P53* et *APC*. Le NB124 supprime efficacement l'ensemble de ces mutations, ce qui indique que son spectre d'action ne se limite pas au contexte décrit dans la littérature comme le plus favorable (U en -1 et C en +4) (Figure 2 – Article). Plusieurs études ont également montré que le codon stop UGA est le codon stop associé au plus fort taux de translecture, suivi de UAG et UAA. Or, les mutations R213X et L360X qui répondent le mieux au traitement par le NB124 sont des codons stop UGA, et la mutation la moins bonne répondeuse (Q1131X), correspond au codon stop UAA. Ces résultats confirment que, sur les mutations testées, les règles de décodage qui régissent le taux de translecture en présence de NB124 ne sont pas différentes de celles des autres aminoglycosides inducteurs de translecture.

Comparativement aux aminoglycosides de référence, le NB124 induit, quelque soit la mutation testée, davantage de translecture que la gentamicine à concentration identique et autant que le G418 à la concentration la plus élevée compatible avec la viabilité cellulaire.

### ★ La molécule TLN468 induit de la translecture sur les mutations R213X et S1196X

A l'issue du crible à grande échelle de 16 480 molécules, nous avons testé 48 molécules pour leur capacité à induire la suppression de la mutation non-sens R213X et nous avons sélectionné la TLN468 qui semble être la plus prometteuse (Figure 16). Le spectre d'action de la TLN468 a été précisé en réalisant des essais de translecture sur 10 mutations impliquées dans différentes maladies telles que le cancer, la mucoviscidose et la myopathie de Duchenne. Une induction de translecture est observée pour seulement deux mutations, la R213X sur laquelle elle a été criblée et la S1196X (Figure 25). Ces deux mutations ont un codon stop UGA, un U en position -1 et un C en +1. Toutefois, cette même séquence est retrouvée dans la mutation L360X et S319X qui ne répondent pourtant pas au traitement par la TLN468. En revanche, il a été montré au laboratoire que la gentamicine induit de la translecture sur toutes ces cibles exceptées sur la E1104X et Q1390X dont le taux induit est trop faible pour être statistiquement différent du taux de base. Nous avons également remplacé le codon stop UGA des mutations R213X et S1196X par les deux autres codons stop et une induction de translecture significative est observée uniquement sur le codon UAA (Figure 26). Or, ce codon stop est globalement moins bien translu par les aminoglycosides que le codon UAG. Ces résultats démontrent donc que la molécule TLN468 induit la suppression de codons stop prématurés puisqu'une activité de translecture est retrouvée sur une mutation différente de celle utilisée dans le crible.

Nos résultats montrent clairement que cette molécule a un spectre d'action différent des aminoglycosides. Une étude approfondie d'un plus grand nombre de mutations permettrait de rationaliser son utilisation et éventuellement d'identifier d'autres séquences répondeuses.

### 3. Le NB124 et la TLN468 stabilisent l'ARNm et restaurent la synthèse d'une protéine de pleine taille

#### ★ Le NB124 et la TLN468 stabilisent l'ARNm soumis à la voie NMD

Nous avons mis en évidence que le traitement par le NB124 et par la TLN468 stabilise l'ARNm P53 portant la mutation R213X dans la lignée tumorale humaine HDQ-P1 (Figure 3A – Article ; Figure 18). Des résultats similaires avaient également été obtenus avec la gentamicine et le G418 sur les cellules HDQ-P1, mais l'efficacité du

NB124 est encore une fois nettement supérieure. Quelle que soit la molécule testée, la stabilisation des transcrits est directement corrélée au taux de translecture induit et de façon proportionnelle pour le NB124. Les modèles décrits dans la littérature expliquent cette corrélation par la présence des complexes EJC et/ou des protéines UPF1 sur l'ARNm. Lors du premier cycle de traduction, l'entrée d'un PTC au site A du ribosome interrompt la traduction et active la voie de dégradation NMD via l'intermédiaire des EJC et/ou d'UPF1. Dans le cas où le codon stop prématuré est translu, la traduction continue jusqu'au codon stop naturel, ce qui permet au ribosome de décrocher les EJC / UPF1 présents sur l'ARNm via son activité hélicase et d'empêcher l'activation ultérieure de la voie NMD. Ainsi, plus le taux de translecture induit est élevé, plus l'ARNm est stabilisé. Les résultats obtenus avec l'apramycine, un aminoglycoside non inducteur de translecture qui se fixe également au site A du ribosome, ne montrent aucune stabilisation des transcrits P53 et confirment le modèle exposé précédemment.

D'un point de vue thérapeutique, si l'ARNm porteur d'un PTC est la cible de la NMD, l'utilisation de ces molécules permet à la fois d'induire la translecture des codons stop et de stabiliser les transcrits et donc de produire davantage de protéines, ce qui est particulièrement bénéfique pour le patient. Il serait également possible d'utiliser des inhibiteurs de la NMD comme le NMDI14 qui potentialiseraient l'effet des inducteurs de translecture et augmenteraient la synthèse de protéines complètes.

#### **★ Le traitement par le NB124 ou la TLN468 permet la synthèse d'une protéine complète**

La quantité de protéine pleine taille réexprimée dépend à la fois de la quantité d'ARNm et de l'acide aminé incorporé au niveau du codon stop qui doit être compatible avec la stabilité de la protéine. Nous avons vu que le traitement des cellules HDQ-P1 par le NB124 et la TLN468 restaure l'expression de la protéine p53 complète (Figure 3B – Article ; Figure 17). De plus, la quantité de protéine réexprimée est directement proportionnelle aux taux de translecture et à la quantité d'ARNm P53 stabilisée.

Le traitement des HDQ-P1 par ces deux inducteurs de translecture a également conduit à la synthèse de protéines p53 tronquées. La quantité de cette forme tronquée dépend de l'efficacité de la voie NMD et du taux de translecture induit. Nous nous serions donc attendus à une augmentation de la forme tronquée de la protéine p53 de façon proportionnelle au taux de translecture. Or, les résultats de western-blot ne vont pas dans ce sens, la quantité de forme tronquée étant approximativement identique, quelle que soit la concentration de NB124 utilisée. L'augmentation de la synthèse, dose-dépendante, de la forme pleine taille pourrait expliquer partiellement ces

résultats. Néanmoins, il est plus probable que les protéines tronquées soient dégradées par le protéasome, ce qui expliquerait également pourquoi à la dose la plus élevée de NB124, la quantité de p53 tronquée est inférieure à celle de la protéine complète. En effet, des mécanismes co-traductionnels et post-traductionnels d'ubiquitylation des peptides naissants ont été mis en évidence et permettraient entre autres de dégrader les protéines mal repliées, dont les protéines tronquées, via la présence de régions hautement hydrophobes<sup>456,457</sup>. Bien que l'accumulation de ces protéines tronquées puisse être délétère pour la cellule, dans le cas de p53, le domaine de tétramérisation indispensable à la fonctionnalité de la protéine et le domaine d'adressage au noyau sont situés en C-terminal. Par conséquent, les protéines p53 tronquées ne devraient pas interférer directement avec les protéines p53 sauvages. Dans le cas des cancers, empêcher la dégradation de la protéine p53 pourrait ainsi conduire à une amélioration clinique de certains patients.

#### 4. Le traitement par le NB124 et la TLN468 conduit à la synthèse d'une protéine fonctionnelle

##### ★ La capacité de transactivation de la protéine p53 est restaurée en présence de NB124 et de TLN468

La réexpression de la protéine est une étape essentielle qui n'est pas toujours associée à une protéine active. En effet, la fonctionnalité de la protéine dépend de la nature de l'acide aminé incorporé à la place du codon stop prématuré. Nous avons donc vérifié la capacité de transactivation de la protéine p53 pleine taille réexprimée après traitement par le NB124 et la TLN468 sur la mutation endogène P53 R213X. Les résultats montrent que ces deux molécules restaurent l'expression de protéines p53 actives capables de reconnaître leurs sites de liaisons (Figure 4A – Article ; Figure 20). En comparaison avec les aminoglycosides de référence, le NB124 est nettement plus efficace que le G418, mais la capacité de transactivation des protéines réexprimées après traitement par la TLN468 est inférieure à celle de la gentamicine. Nous avons pourtant vu précédemment que le taux de translecture induit était aussi élevé en présence de TLN468 qu'avec la gentamicine. Deux hypothèses sont envisageables : soit l'acide aminé majoritairement incorporé au niveau du PTC conduit à une restauration partielle de la fonctionnalité de la protéine, soit les différents acides aminés incorporés le sont dans des proportions égales et seule une fraction restaure l'activité de la protéine p53.

Les données issues de la littérature tendent à montrer que plusieurs ARNt proche-cognat sont capables de reconnaître chacun des codons stop et que le traitement par la gentamicine ne modifierait que leur proportion d'incorporation. Le codon TGA serait décodé chez la levure très majoritairement par le tryptophane puis par la cystéine et très faiblement par l'arginine <sup>245</sup>. Des mutants de la protéine p53 ont été construits au laboratoire afin d'analyser la fonctionnalité de la protéine selon la nature de l'acide aminé présent à la position 213. L'arginine est compatible avec l'activité de p53 puisqu'il s'agit de l'acide aminé présent à cette position dans la protéine sauvage. En revanche, en présence d'un tryptophane ou d'une cystéine, la protéine p53 n'est plus active. L'acide aminé préférentiellement incorporé à cette position ne serait donc pas compatible avec une fonctionnalité de la p53.

Ces résultats ne sont pas cohérents avec l'activité de la p53 observée en présence de la TLN468 et du NB124, car seule l'arginine est compatible avec l'activité de la protéine et cet acide-aminé serait incorporé dans moins de 5% des cas. Ces données pourraient s'expliquer par une différence de la nature de l'acide aminé incorporé chez la levure et chez les mammifères. De plus, il est possible que la structure chimique, le mode d'action de la molécule et même le contexte de la mutation non-sens modifient la nature de l'ARNt proche-cognat suppresseur. Ceci pourrait être vérifié en analysant l'acide aminé incorporé à la position du PTC dans les cellules de mammifères avec différents aminoglycosides et avec la TLN468 sur plusieurs mutations non-sens. Toutefois, l'identification d'un nombre restreint d'acides-aminés limiterait fortement la probabilité de restaurer l'activité des protéines contenant une mutation non-sens. L'efficacité de cette stratégie dépendrait alors à la fois de la nature des acides aminés pouvant être incorporés sur un PTC spécifique, de leur proportion d'incorporation et de la position de la mutation qui peut être plus moins délétères pour l'activité de la protéine.

**★ La protéine p53 réexprimée par le NB124 et la TLN468 est capable de réguler la transcription de ses gènes cibles**

Nous avons ensuite mis en évidence que la protéine p53 réexprimée après traitement par le NB124 dans la lignée HDQ-P1 est capable de réguler positivement l'expression du gène *p21*, impliqué dans l'arrêt du cycle cellulaire et du gène *Bax* impliqué dans l'induction de l'apoptose (Figure 4B – Article). Nous avons également observé une régulation positive de *Bax* après traitement avec la TLN468 à 80  $\mu$ M (Figure 21). L'ensemble de nos résultats indique que même si la translecture conduit à la synthèse d'une fraction de protéine ne possédant pas l'acide aminé sauvage à la position 213, l'ensemble des protéines générées permet de réguler les gènes cibles

dans le même sens qu'une p53 sauvage. Le traitement par l'une ou l'autre de ces deux molécules permet donc de réexprimer suffisamment de protéine p53 fonctionnelle pour réguler des gènes cibles impliqués dans l'arrêt du cycle cellulaire ou dans l'apoptose.

## 5. La TLN468 est une molécule qui nécessite d'être optimisée

### ★ TLN468 est toxique sur les cellules HDQ-P1 et H1299

La TLN468 n'affecte que faiblement la croissance cellulaire, mais l'augmentation de sa concentration est associée à une toxicité sévère. Nous avons donc procédé à des tests de viabilité cellulaire avec la TLN468 afin de mieux définir sa toxicité sur trois lignées cellulaires : des fibroblastes de souris NIH3T3 et deux lignées tumorales humaines HDQ-P1 et H1299 (Figure 22). TLN468 présente une faible toxicité sur les NIH3T3 quelle que soit la dose testée. En revanche, la toxicité est plus marquée sur les H1299, principalement aux doses les plus élevées qui semblent particulièrement sensibles à la présence de TLN468. Quant aux HDQ-P1 la toxicité devient particulièrement marquée à la plus forte concentration. Les données obtenues après 48h de traitement montrent que la TLN468 présente une toxicité toujours modérée sur les NIH3T3 et plus marquée sur les HDQ-P1 et les H1299 (Annexe 3). La TLN468 nécessite donc d'être optimisée pour pouvoir être utilisée à des concentrations plus élevées et sur des temps de traitement aussi ou plus important que 48h.

### ★ Aucun analogue de la TLN468 n'induit de translecture sans toxicité cellulaire

Nous avons alors cherché à identifier si certains groupements chimiques de la TLN468 pouvaient être responsables spécifiquement de la toxicité ou de la capacité d'induction de la translecture. Nous avons dans un premier temps criblé informatiquement l'ensemble des chimiothèques de l'équipe de J.C Cintrat, soit environ 80 000 molécules et sélectionné les molécules constituées de la structure chimique de base de la TLN468, soit 378 molécules. Nous avons ensuite criblé ces analogues sur notre système rapporteur et identifié une molécule, TLN037, qui est structurellement très proche de TLN468 (Figure 23). TLN037 ne possède que deux groupements méthyles supplémentaires sur la structure de base par rapport à TLN468. Les mesures d'activité de translecture montrent que TLN229 et TLN944 ne sont pas capables de supprimer les codons stop alors que TLN037 induit une augmentation du taux de translecture d'un facteur 2 (Figure 24). Or, TLN229 et TLN944 ne diffèrent de TLN468

que par l'absence du groupement « R ». Ce dernier pourrait donc être responsable de l'activité de translecture retrouvée chez TLN468, TLN037 et TLN185. Les résultats de translecture sur TLN651, qui ne possède que le groupement « R », devraient confirmer cette hypothèse. Nous n'avons cependant pas pu identifier de groupements responsables de la toxicité. Les molécules TLN185 et TLN037 sont particulièrement toxiques sur les NIH3T3 et bien qu'elles possèdent toutes deux des groupements méthyles, ces derniers ne sont pas situés sur les mêmes positions. La présence de méthyle pourrait être corrélée à la toxicité cellulaire de cette famille de molécules, mais TLN944 a la même structure que TLN468 sans le groupement « R » et aucune toxicité n'est observée. La toxicité est donc probablement multifactorielle et proviendrait à la fois du groupement « R » et des groupements méthyles. La réalisation d'une étude de relation structure activité (SAR) permettrait d'identifier plus précisément le rôle des différents groupements de TLN468 et ainsi d'améliorer son efficacité et diminuer sa toxicité.

## 6. La protéine p53 pleine taille réexprimée par le NB124 est capable de limiter la prolifération cellulaire en induisant l'apoptose

### ★ La réexpression d'une protéine p53 complète active la voie apoptotique

Nous avons montré que le NB124 induit la réexpression des protéines p53 capables d'activer l'expression de deux de ses principaux gènes cibles. Cependant, cette induction doit être suffisante pour induire l'apoptose des cellules cancéreuses et ainsi limiter la prolifération tumorale. Une des cascades d'activation de l'apoptose la mieux décrite implique l'activation de caspases (cysteinyl-aspartate protéases) qui vont initier la mort cellulaire en clivant plusieurs protéines clés telles que PARP-1. Nous avons mis en évidence que 91% de PARP-1 était clivée après traitement par le NB124 dans les cellules tumorales humaines H1299, dans lesquelles nous avons stablement intégré un ADNc P53 R213X (Figure 5A – Article). En revanche, la gentamicine n'est pas suffisamment efficace pour déclencher l'activation des caspases et le clivage de PARP-1. Il semble donc essentiel d'atteindre un certain seuil d'efficacité de translecture pour réexprimer suffisamment de protéines p53 afin de déclencher l'activation de la voie apoptotique. Nous avons confirmé ces résultats en analysant la quantité de cellules dans lesquelles l'apoptose était induite. Dans ces expériences, la protéine p53 pleine taille réexprimée après traitement par le NB124 est capable d'augmenter

significativement le nombre de cellules en apoptose précoce et tardive (Figure 5B – Article). Ces résultats sont très prometteurs et devront être confirmés *in vivo* en stoppant la progression tumorale d'une xénogreffe sur le modèle murin.

## 7. La molécule unique, un outil d'analyse dynamique de la traduction eucaryote

### ★ Stratégie de marquage du système biologique

La molécule unique est une technique permettant l'étude de processus dynamiques et asynchrones sans en moyenniser le comportement individuel. Cette technique est donc particulièrement adaptée à l'étude de la traduction. Néanmoins, une des difficultés majeures est de marquer le système biologique afin de suivre sa cinétique. Dans le cas de la traduction, le marquage est le plus souvent réalisé sur le ribosome procaryote, sur les ARNt et sur la 40S de *S. cerevisiae*<sup>458</sup>. Le risque de cibler la machinerie traductionnelle est d'altérer le mécanisme de la traduction à cause de l'encombrement stérique des fluorophores ou de leur interaction avec le ribosome ou l'un des facteurs impliqués dans ce processus. C'est pourquoi nous avons choisi de marquer l'ARNm à l'aide de deux oligonucléotides qui ont été désignés pour ne pas altérer la traduction<sup>211</sup>. Ces nucléotides sont associés à deux fluorophores différents (ATTO647 et ATTO565) et sont détachés au cours de l'élongation par l'activité hélicase du ribosome. La fluorescence émise par ces sondes est détectée et localisée par un système de microscopie à fluorescence par réflexion totale interne (TIRF) développée par le laboratoire de Karen Perronet et Nathalie Westbrook. Ce système assure à la fois l'excitation grâce à la formation d'une onde évanescente à la surface de la lamelle et la collection du signal renvoyé sur un capteur.

L'inconvénient principal de ce type de marquage provient du fait qu'un événement de traduction est associé à une disparition des fluorophores et non à une apparition. Or plusieurs biais existent dans la disparition prématurée d'un signal et de nombreuses optimisations ont été nécessaires pour s'en affranchir :

- Cette disparition peut être liée au photoblanchiment des fluorophores qui correspond à la perte irréversible des propriétés de fluorescence. Il est connu que l'oxygène est le facteur principal de photoblanchiment des fluorophores organiques. Nous avons pu observer qu'en présence de RRL, le photoblanchiment des fluorophores est accéléré. En effet, le RRL est particulièrement riche en hémoglobine et il est probable que l'abondance d'oxygène soit responsable du photoblanchiment observé.

Des essais utilisant la méthode ROXS (Reducing and OXydizing System), qui appauvrit la solution en oxygène, ont été réalisés sur différents extraits cellulaires et ont permis de stabiliser les fluorophores d'un facteur 2 après quelques dizaines de minutes<sup>459</sup>. Cependant, le photoblanchiment n'est plus constant au cours du temps. Nous avons donc diminué le photoblanchiment en utilisant un objectif à plus grande ouverture numérique qui collecte davantage de photons et permet de mieux confiner l'onde évanescente. Le rapport signal à bruit a ainsi été largement amélioré, ce qui nous a permis de diminuer la puissance d'excitation des fluorophores sans compromis sur la détection. La part de photoblanchiment évaluée en réalisant des expériences où seul le photoblanchiment est mesuré est nettement inférieure aux départs liés à la traduction.

- Nous observons également des départs spontanés, c'est-à-dire des disparitions de signal qui ne sont pas liées au photoblanchiment. En effet, les oligonucléotides peuvent se déshybrider de l'ARN et sortir de l'onde évanescente, ce qui provoque aussi une disparition du signal. Ces départs sont liés à une déstabilisation de l'hybridation qui peut avoir deux origines: i) les propriétés physico-chimiques de l'environnement ont été modifiées lors de l'injection du lysat de réticulocytes de lapin (RRL) pour initier la traduction. Dans ce cas, les départs ont lieu très rapidement après l'injection et sont facilement éliminés lors de l'analyse; ii) par la présence de fluctuations de température et/ou suite à une hybridation partielle. Une étape d'incubation à 37°C permet de détacher la plupart des oligonucléotides mal appariés.

- Nous savons que d'autres facteurs peuvent conduire à la disparition du signal de fluorescence, tels que les RNases, c'est pourquoi nous avons également réalisé des contrôles sans site d'initiation de la traduction et avec un inhibiteur la cycloheximide. Il y a systématiquement moins de départs dans ces contrôles que dans la traduction (Figure 3A - Article), ce qui montre que le différentiel entre ces deux conditions correspond à des départs liés à un événement de traduction.

### **★ La traduction est majoritairement réalisée par les ribosomes purifiés**

Une des difficultés majeures pour mettre au point un tel système de traduction in vitro en molécule unique est de reconstituer l'étape d'initiation qui nécessite plus de 40 facteurs chez les eucaryotes. Nous avons contourné cette difficulté en utilisant un site d'entrée interne pour le ribosome<sup>441</sup> provenant du Cricket Paralysis Virus (CrPV) qui a la particularité de recruter le ribosome 80S indépendamment de tous les facteurs d'initiation et de l'ARNt initiateur. En effet, il est décrit dans la littérature que le domaine PKI de l'IRES mime la présence d'un ARNt interagissant avec le codon au site A<sup>460</sup>. Nous avons montré que les ribosomes 80S purifiés sont capables d'initier une traduction sur l'IRES CrPV malgré la présence d'une sonde marquée à proximité du site

d'initiation. Le ribosome traduit ensuite l'ARNm et décroche successivement les deux sondes (Figure 1 - Article).

Toutefois, le RRL étant également constitué de ribosomes, nous avons cherché à savoir si la traduction était réalisée par les ribosomes purifiés et préincubés ou par ceux apportés par le RRL. Les résultats des essais *in vitro* avec de la méthionine S<sup>35</sup> ont démontré que les ribosomes du RRL traduisent l'ARNm avec une efficacité légèrement supérieure que les ribosomes purifiés (Figure 7). En revanche, les essais de traduction en molécule unique ont mis en évidence une plus faible quantité d'ARNm traduits en présence uniquement de ribosomes du RRL. Ces différences s'expliquent par le fait que seul le premier cycle de traduction est observé en molécule unique, contrairement aux essais en tube où la méthionine radiomarquée est incorporée à chaque cycle. De plus, il a été montré que l'initiation des ribosomes 80S est plus lente que lorsque seule la 40S se fixe à l'IRES<sup>461</sup>. Il est donc possible que les ribosomes provenant du RRL nécessitent davantage de temps pour initier sur l'IRES. Des analyses statistiques ont été conduites par nos collaboratrices à partir des données de cinétiques de traduction afin de déterminer la probabilité que les ribosomes préincubés traduisent un ARNm. Les résultats indiquent que 80% des ARNm traduits le sont par des ribosomes préincubés et 20% par les ribosomes du RRL (Annexe 4). Les ribosomes du RRL ne sont donc pas responsables de la traduction observée.

#### ★ Le ribosome eucaryote traduit l'ARN à une vitesse de 1,4s/codon

Les récentes analyses de la dynamique de la traduction utilisent des systèmes où les constituants de la machinerie traductionnelle, provenant d'organismes variés, sont purifiés et reconstitués pour former un système de traduction *in vitro*<sup>461,462</sup>. La concentration de chacun des constituants peut alors être modifiée, mais les vitesses de traduction observées s'éloignent fortement des données obtenues à partir de lysats cellulaires tels que le RRL et de ceux obtenus *in vivo*. Ces systèmes permettent d'étudier le rôle de chacun des facteurs, mais de façon non physiologique. Dans le système de traduction que nous avons mis au point, seuls les ribosomes sont purifiés et ils proviennent du RRL. La présence des deux sondes marquées sur l'ARNm permet de calculer la vitesse d'élongation de la traduction. Cette vitesse a été déterminée en analysant le différentiel de temps entre la perte de signal du fluorophore le plus éloigné de l'IRES (T<sub>down</sub>) et le plus rapproché (T<sub>up</sub>)<sup>463</sup>. Les ribosomes purifiés et préincubés traduisent à une vitesse approximative de 1,4 secondes par codon, ce qui confirme les résultats d'autres travaux utilisant le RRL<sup>464</sup>. Comparativement, la traduction *in vivo* est environ 6 fois plus rapide avec une vitesse comprise entre 6 et 10 codons par seconde<sup>465,466</sup>. Il a été montré dans plusieurs lignées cellulaires eucaryotes que la machinerie

traductionnelle est hautement organisée permettant à la cellule d'optimiser le taux de traduction. En effet, la coordination des activités synthétases faciliterait l'acheminement des ARNt aminoacylés aux ribosomes en cours d'élongation <sup>467</sup>. Ainsi, la synthèse protéique d'une cellule de mammifère perméabilisée serait 40 fois plus rapide que dans un système exempt de cellules préparé en réalisant un lysat de ces mêmes cellules <sup>468</sup>. Ces données pourraient expliquer pourquoi la traduction in vitro en tube, mais également en molécule unique est plus lente que la traduction cellulaire. En ce qui concerne notre système, nous ne pouvons pas non plus exclure que cette différence ne soit pas due à l'effet de la température ou à une gêne stérique du fait que le système soit accroché à la surface de la lamelle. Enfin, il est également envisageable que les facteurs intervenant au cours de la traduction soient ralentis par l'augmentation de la viscosité à proximité d'une surface plane.

**\* L'initiation IRES dépendante ralentit la ou les première(s) étape(s) d'élongation de la traduction**

Les sites d'entrées internes pour le ribosome permettent d'initier la traduction de façon indépendante de la coiffe. Bien que de nombreuses études biochimiques et structurales ont analysé l'étape d'initiation de la traduction, les conséquences d'une initiation via l'IRES sur l'élongation n'ont jamais été abordées. Les analyses de nos données de cinétiques de traduction montrent la présence de deux temps caractéristiques, un long de 40s et un court de 1,4s (fois le nombre de cycles d'élongation). Ces temps caractéristiques sont définis dans un modèle mathématique décrivant les événements de départs des sondes marquées observés au cours des cinétiques de traduction. Le temps long correspondrait au premier ou aux deux premiers cycles d'élongation et le temps court à chaque cycle suivant (Figure 4 – Article). En effet, la première translocation nécessite de déplacer l'imposante structure secondaire qu'est l'IRES dont son pseudonoeud <sup>59</sup> qui mime l'ARNt au site A du ribosome. Le facteur d'élongation eEF2 est recruté sur le ribosome et l'hydrolyse de son GTP lui permet de transloquer l'IRES du site A au site P <sup>460,461</sup>. Un ARNt aminoacylé peut alors rentrer au site A du ribosome. Il est très probable que les nombreuses interactions de l'IRES dans le ribosome soient responsables du ralentissement de la traduction. A l'issue du second cycle d'élongation, le PKI est au site E. Il est alors envisageable que l'IRES n'interfère plus que faiblement avec le ribosome dont la vitesse de traduction augmente et tend vers 1,4s par codon. Ces résultats démontrent clairement que l'initiation de la traduction par une IRES ralentit la vitesse de traduction du ribosome au cours du ou des premiers cycles d'élongation. Très récemment, le laboratoire de B.S. Cooperman a publié des données de cinétique de traduction via une initiation

dépendante de l'IRES<sup>462</sup>. Dans leur modèle, les quatre premiers cycles d'élongations sont ralentis. Toutefois, ils utilisent un système de traduction *in vitro* reconstitué dont le cycle d'élongation le plus rapide est de 20s par codon, soit 15 fois plus lent que les ribosomes purifiés du RRL. De plus, le « temps long » est également 15 fois plus élevé que ce que nous avons mesuré ( $\approx 600s$ ). Ces données, bien que peu physiologiques, corroborent nos résultats et confirment que l'initiation de la traduction sur l'IRES CrPV est fortement ralentie lors des premiers cycles d'élongation.

### ★ Les ribosomes du RRL sont plus actifs que les ribosomes purifiés

Nous avons vu que seulement 20% des ribosomes du RRL participent à la traduction des ARNm. Pour cette raison, nous avons dû purifier les ribosomes 80S du RRL et les préincuber avec l'ARNm pour faciliter leur initiation sur l'IRES. De plus, les analyses de cinétiques de traduction ont mis en évidence que les ribosomes provenant du RRL sont deux à trois fois plus rapides que ceux purifiés. Lors de l'étape de purification, la perte d'une protéine ribosomale et/ou l'endommagement de l'ARNr peuvent être responsables de la diminution de l'activité du ribosome. Nous avons alors cherché à optimiser notre système de traduction et à nous affranchir de cette étape de purification des ribosomes. Ainsi, excepté l'ARNm, la totalité des intervenants de la traduction seraient alors apportés par le RRL. Une hypothèse qui pourrait expliquer la faible proportion d'ARNm traduits par les ribosomes du RRL serait liée à la proximité du site d'initiation de la surface de la lamelle. En effet, d'importantes contraintes spatiales pourraient empêcher la majorité des ribosomes d'initier sur les IRES. Les ribosomes préincubés ne sont pas soumis à ces contraintes, car ils sont déjà associés aux IRES. La distance séparant le site d'accroche de l'ARNm à l'IRES a alors été allongée de 14 codons et les résultats préliminaires obtenus montrent que la proportion de ribosomes du RRL qui traduisent les ARNm augmente d'un facteur 4 (Figure 9). De plus, le temps long n'est pas modifié par rapport au design précédent, ce qui signifie que la vitesse d'initiation des ribosomes sur l'IRES et la vitesse des premiers cycles d'élongation restent identiques. Par conséquent, la cinétique de liaison des ribosomes ne dépendrait pas des contraintes spatiales. Ces résultats corroborent notre hypothèse et confirment que l'accroche des ribosomes était limitée par la faible accessibilité de l'IRES. Une seconde hypothèse peut être envisagée où la majorité des ribosomes seraient en réalité capables d'initier la traduction sur l'IRES proche de la lamelle, mais pas de quitter cette structure secondaire et donc de traduire l'ARNm. Dans cette seconde hypothèse, l'étape limitante ne serait plus l'accroche du ribosome sur l'IRES mais le départ du ribosome de son site d'initiation. L'éloignement de l'IRES permettrait alors au ribosome

de traduire plus facilement et donc de se séparer de l'IRES qui ralentit les premiers cycles d'élongation. Il est toutefois difficile d'expliquer pourquoi les ribosomes purifiés et préincubés n'étaient pas gênés de la même manière. Quoi qu'il en soit, ces nouveaux résultats sont très prometteurs et l'amélioration de notre système va nous permettre de l'utiliser pour étudier les événements de recodage d'une manière originale.

# ***Conclusions et Perspectives***

***Le futur est déjà là,  
il n'est simplement pas encore distribué de façon équitable  
(William Gibson)***

## Conclusion - Perspectives

A ce jour, seuls le PTC124 et la gentamicine qui ont été utilisés pour le traitement des maladies à codons stop ont été à même de démontrer un bénéfice thérapeutique chez les patients traités. Toutefois, l'effet du PTC124 n'a pas été confirmé sur les essais cliniques de phase 3 et la toxicité de la gentamicine ne lui permet pas d'être utilisée dans le cadre d'un traitement à long terme. Il y a donc un réel déficit de molécules capables d'induire la translecture des PTC et ainsi réexprimer une protéine fonctionnelle. D'importants efforts ont été faits en ce sens, comme le démontre le nombre croissant de molécules découvertes ces dernières années telles que le PTC124, les RTC, l'escine et très récemment, le CDX5-1<sup>469</sup>. Toutefois, l'efficacité de ces molécules est restreinte à un petit nombre de mutations dont le contexte et la nature du codon stop sont favorables à leur suppression. De plus, leur mode d'action n'étant pas identifié, il est particulièrement complexe de les optimiser. Enfin, la toxicité de ces composés peut également être un frein à leur utilisation. L'identification de leur cible moléculaire ainsi que la compréhension des mécanismes de la traduction impliqués lors du traitement de cellules par ces composés permettraient pourtant d'accélérer leur développement.

Au cours de ma thèse, je me suis donc intéressé à la suppression traductionnelle des codons stop chez les mammifères sous deux aspects. J'ai d'une part cherché à identifier de nouvelles molécules capables d'induire de la translecture puis j'ai analysé leur efficacité dans le cadre d'une thérapie anti-cancer à travers la réexpression du gène *P53* contenant une mutation non-sens. J'ai également développé un système de traduction *in vitro* eucaryote qui permet l'étude dynamique de l'élongation de la traduction et des mécanismes impliqués au cours de cette étape.

Plusieurs stratégies ont été développées dans le but d'identifier de nouvelles molécules inductrices de translecture. La plus connue est le crible à grande échelle et permet de discriminer rapidement un très grand nombre d'agents actifs des inactifs. Nous avons développé un système de criblage innovant basé sur la synthèse d'une luciférase sécrétée dans le milieu de culture sous condition d'un événement de translecture. Nous avons ainsi criblé avec l'équipe de J.C Cintrat plus de 17 000 molécules qui ont conduit à l'identification d'un hit, la TLN468. Cette molécule s'est révélée aussi efficace que la gentamicine sur certaines mutations, mais présente une toxicité encore trop importante

pour être utilisée *in vivo*. Nous avons entamé la phase d'optimisation de la TLN468 en analysant l'efficacité de translecture et la toxicité de plusieurs analogues dont les résultats suggèrent que le groupement R serait responsable de l'activité de translecture de la molécule. Ces résultats ont encore besoin d'être confirmés sur l'analogue TLN651 avant de procéder à une étude de relation structure-activité.

Cette étude permettra à la fois d'identifier plus précisément le rôle de chaque groupement et de marquer la molécule dans le but de déterminer sa cible moléculaire. Dans le cas où il ne serait pas possible de marquer la TLN468 sans altérer son activité, de nombreuses méthodes ont été développées pour identifier la cible d'une molécule sans modifications chimiques<sup>470</sup>.

En revanche, le spectre d'action de la TLN468, qui n'est pas encore complètement caractérisé, diffère de celui des aminoglycosides. Ces résultats suggèrent que les règles du contexte nucléotidique qui régissent le taux de translecture sont différentes en présence de gentamicine et de TLN468. Il est donc très probable que la cible cellulaire de la TLN468 diffère de celle de la gentamicine et de la généticine. L'obtention d'un plus grand nombre de données sur d'autres mutations non-sens permettra de mieux caractériser l'effet du contexte sur la translecture induite par la TLN468.

Plus récemment, une autre stratégie a vu le jour : la conception de médicaments basés sur la connaissance de la cible biologique, également appelée « drug design » ou design rationnel. Cette stratégie a été utilisée par T. Baasov et son équipe pour concevoir de nouveaux composés optimisés à partir de la paromomycine et de la généticine. Plusieurs dérivés ont été synthétisés et nous avons montrés que non seulement le NB124 est la molécule dont le taux de translecture est le plus élevé, mais qu'elle est aussi deux fois plus efficace que la gentamicine. Cependant, les aminoglycosides sont connus pour leur ototoxicité qui rend leur utilisation clinique incompatible avec un traitement à long terme. Ce point a également été amélioré sur le NB124 qui présente une ototoxicité quatre fois plus faible que la gentamicine<sup>413</sup>. Le NB124 est donc le premier dérivé d'aminoglycosides à avoir un réel potentiel thérapeutique, qui a déjà été démontré sur la CFTR et que nous avons analysé sur le cancer.

Le design rationnel est donc une stratégie parfaitement adaptée à l'optimisation de molécules dont l'efficacité de translecture a déjà été démontrée. Ce type de stratégie est cependant encore difficile à mettre en œuvre, car il est nécessaire d'avoir établi une corrélation entre chacun des groupements chimiques et l'activité biologique (où l'affinité). Il est également préférable d'avoir résolu la structure de la molécule liée à sa cible. Ainsi, l'obtention de la structure tridimensionnelle de la généticine liée au ribosome 80S de levure a déjà permis à l'équipe de T. Baasov d'optimiser l'affinité de

ces dérivés pour le ribosome eucaryote (NB156-NB157). Ce type de modification a deux avantages, il permet à la fois d'augmenter la proportion de ribosomes cytoplasmiques capables de translire le stop efficacement et de diminuer l'affinité de ces dérivés pour le ribosome mitochondrial. En effet, cette affinité est partiellement responsable de la toxicité des aminoglycosides. La compréhension des mécanismes traductionnels associés aux avancées de la biologie structurale joue donc un rôle prépondérant dans le développement et la réussite de cette stratégie pouvant être appliquée à un large panel de molécules.

La compréhension des mécanismes de la traduction permet non seulement d'améliorer les molécules qui ciblent l'appareil traductionnel, mais aussi de développer une approche raisonnée de la recherche de nouvelles molécules thérapeutique. C'est pourquoi nous avons développé pour la première fois, avec l'équipe de K. Perronet, un outil d'analyse dynamique de la traduction eucaryote en molécule unique sans modifier ni le ribosome, ni les ARNt, ni aucun des facteurs protéiques intervenant au cours de ce processus. Nous avons pu montrer que ce système permet de détecter et d'analyser la vitesse de la traduction d'un ARNm marqué par des ribosomes purifiés du RRL et par des ribosomes directement apportés par le RRL. Les résultats ont mis en évidence que les ribosomes traduisent efficacement cet ARNm à une vitesse proche de celle estimée *in vivo* (0,5 seconde par codon). De plus, les données dynamiques de la traduction que nous avons obtenues ont montré que l'initiation de la traduction par une IRES ralentissait le ribosome lors des premiers cycles d'élongation de la traduction. Lorsque les données de traduction obtenues à partir des ribosomes non purifiés seront confirmées, il sera alors possible d'étudier la fidélité de la traduction en analysant la suppression traductionnelle des codons stop. Dans un premier temps, la cinétique de la traduction sera mesurée en absence de traitements sur la mutation R213X, ce qui permettra de vérifier que les ribosomes effectuent la translecture du codon stop. Puis, les molécules inductrices de translecture telles que le NB124 et la TLN468 seront ajoutées et la cinétique de franchissement du codon stop sera déterminée pour chaque ARNm ayant été traduit par un ribosome. Enfin, davantage d'essais seront réalisés en faisant varier la nature du codon stop et le contexte nucléotidique. Ces données devraient permettre de déterminer les paramètres cinétiques qui gouvernent le passage des codons stop. Il sera par la suite possible d'étudier d'autres événements de recodage tels que le décalage de phase de lecture (frameshift) en faisant varier la séquence glissante et la structure secondaire. L'obtention de données dynamiques sera une avancée majeure dans la compréhension de la traduction eucaryote.

Le traitement des maladies à codon stop par des inducteurs de translecture ne conduit pas systématiquement à un bénéfice thérapeutique. Un certain nombre de paramètres autres que l'efficacité de translecture et que la toxicité doivent être prises en compte. Tout d'abord, il est nécessaire d'avoir un ARNm en quantité suffisante afin d'atteindre un niveau suffisant de protéine complète. Ces protéines doivent être fonctionnelles, ce qui implique que l'acide aminé incorporé doit être compatible avec l'activité de la protéine et que la protéine complète réexprimée soit correctement repliée.

Nous avons pu montrer que le NB124 est la première molécule inductrice de translecture à induire l'apoptose de cellules tumorales via la réexpression de la protéine p53 tout en ayant une faible ototoxicité. Ces résultats confirment le potentiel thérapeutique du NB124 et ouvrent la voie au traitement de nombreuses maladies génétiques liées à la présence de mutations stop. La prochaine étape consiste à analyser le potentiel métastatique des cellules tumorales traitées ou non au NB124 dans des chambres d'invasion. Puis, des essais seront réalisés sur la souris afin de confirmer la capacité du NB124 à empêcher la progression tumorale sur un modèle de xénogreffe. La possibilité d'utiliser le NB124 en synergie avec d'autres molécules sera également évaluée, et notamment avec la molécule CDX5-1 dont les premiers résultats ont démontré la capacité à potentialiser l'effet de la généticine. L'inhibition de la voie NMD permettrait encore d'accroître l'effet de ce type de traitements et plutôt que d'utiliser un inhibiteur général de cette voie de dégradation telle que le NMDI14, il serait moins toxique de cibler spécifiquement les sites d'EJC à proximité du PTC en utilisant des oligonucléotides anti-sens (ASO). Ce type d'approche permettrait d'augmenter la quantité de protéine fonctionnelle et d'obtenir un bénéfice thérapeutique sur des mutations qui répondent mal aux aminoglycosides inducteurs de translecture. A plus long terme, l'ADME-Tox (Absorption, Distribution, Métabolisation, Elimination – Toxicité) du NB124 sera analysée avant d'envisager une utilisation clinique.

Nous avons vu que la molécule TLN468 permet aussi d'induire la réexpression de protéines p53, mais celles-ci semblent peu fonctionnelles. Nous avons émis l'hypothèse que l'acide aminé incorporé à la position du PTC ne soit pas ou partiellement compatible avec l'activité de la protéine. Toutefois, il n'existe à ce jour que très peu de données sur l'identité des acides aminés incorporés au niveau des codons stop par translecture. Notre équipe a développé *in vivo* une méthode d'analyse systématique des acides aminés incorporés au niveau des trois codons stop chez la levure. Dans un premier temps, il serait intéressant d'identifier les acides aminés incorporés sur la mutation R213X en présence de TLN468 chez la levure et de comparer ces résultats avec ceux obtenus avec le NB124. Puis, il sera important de réitérer ces expériences sur

les cellules de mammifères pour lesquelles ce système sera mis au point très prochainement. Ces données permettront à la fois de mieux comprendre les règles de décodage du codon stop et à plus long terme, de pouvoir prédire l'identité de l'acide aminé incorporé, ce qui constituerait un sérieux avantage dans la mise en place de cette stratégie thérapeutique.

Au cours de ma thèse, j'ai donc caractérisé un nouvel aminoglycoside. Mes travaux de recherche ont aussi permis d'identifier une nouvelle molécule inductrice de translecture, la TLN468. Cette molécule semble agir d'une manière totalement originale. L'identification de la cible moléculaire permettra non seulement de mieux comprendre le mécanisme de translecture, mais aussi devrait permettre la mise au point de nouvelles molécules inductrices de translecture.

# ***Références bibliographiques***

***Nous ne connaissons pas le millionième  
d'un pourcent de quoi que ce soit  
(Thomas Edison)***

## Références bibliographiques

- 1 Schwanhausser, B. et al. Global quantification of mammalian gene expression control. *Nature* 473, 337-342, doi:10.1038/nature10098 (2011).
- 2 Crick, F. H., Barnett, L., Brenner, S. & Watts-Tobin, R. J. General nature of the genetic code for proteins. *Nature* 192, 1227-1232 (1961).
- 3 Nirenberg, M. et al. RNA codewords and protein synthesis, VII. On the general nature of the RNA code. *Proceedings of the National Academy of Sciences of the United States of America* 53, 1161-1168 (1965).
- 4 Knight, R. D., Freeland, S. J. & Landweber, L. F. Rewiring the keyboard: evolvability of the genetic code. *Nature reviews. Genetics* 2, 49-58, doi:10.1038/35047500 (2001).
- 5 Barrell, B. G., Bankier, A. T. & Drouin, J. A different genetic code in human mitochondria. *Nature* 282, 189-194 (1979).
- 6 Uzman, A. *Molecular Cell Biology* (4th edition) - Harvey Lodish, Arnold Berk, S. Lawrence Zipursky, Paul Matsudaira, David Baltimore and James Darnell; Freeman & Co., New York, NY, 2000, 1084 pp., list price \$102.25, ISBN 0-7167-3136-3. *Biochemistry and Molecular Biology Education*, 126-128, doi:doi: 10.1016/s1470-8175(01)00023-6 (2001).
- 7 Bentley, D. L. Coupling mRNA processing with transcription in time and space. *Nature reviews. Genetics* 15, 163-175, doi:10.1038/nrg3662 (2014).
- 8 Shatkin, A. J. Capping of eucaryotic mRNAs. *Cell* 9, 645-653 (1976).
- 9 Gonatopoulos-Pournatzis, T. & Cowling, V. H. Cap-binding complex (CBC). *The Biochemical journal* 457, 231-242, doi:10.1042/BJ20131214 (2014).
- 10 Clark, D. P. & Pazdernik, N. J. in *Biotechnology (Second Edition)* 33-61 (Academic Cell, 2016).
- 11 Goldman, E. in *eLS* (John Wiley & Sons, Ltd, 2001).
- 12 Elkon, R., Ugalde, A. P. & Agami, R. Alternative cleavage and polyadenylation: extent, regulation and function. *Nature reviews. Genetics* 14, 496-506, doi:10.1038/nrg3482 (2013).
- 13 Goss, D. J. & Kleiman, F. E. Poly(A) binding proteins: are they all created equal? *Wiley interdisciplinary reviews. RNA* 4, 167-179, doi:10.1002/wrna.1151 (2013).
- 14 Hou, Y. M. CCA addition to tRNA: implications for tRNA quality control. *IUBMB life* 62, 251-260, doi:10.1002/iub.301 (2010).
- 15 Anger, A. M. et al. Structures of the human and Drosophila 80S ribosome. *Nature* 497, 80-85, doi:10.1038/nature12104 (2013).
- 16 Borek, E. Modification of nucleic acids in relation to differentiation. *Trends in biochemical sciences* 2, 3-6, doi: 10.1016/0968-0004(77)90314-0 (1977).
- 17 Barciszewska, M. Z., Perrigue, P. M. & Barciszewski, J. tRNA - the golden standard in molecular biology. *Molecular bioSystems* 12, 12-17, doi:10.1039/C5MB00557D (2016).

- 18 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).
- 19 Cochella, L. & Green, R. Fidelity in protein synthesis. *Current biology : CB* 15, R536-540, doi:10.1016/j.cub.2005.07.018 (2005).
- 20 Klinge, S., Voigts-Hoffmann, F., Leibundgut, M. & Ban, N. Atomic structures of the eukaryotic ribosome. *Trends in biochemical sciences* 37, 189-198, doi:10.1016/j.tibs.2012.02.007 (2012).
- 21 Sonenberg, N., Wilchek, M. & Zamir, A. Mapping of 23-S rRNA at the ribosomal peptidyl-transferase center by photo-affinity labeling. *European journal of biochemistry / FEBS* 77, 217-222 (1977).
- 22 Bushby, K. et al. Ataluren treatment of patients with nonsense mutation dystrophinopathy. *Muscle Nerve* 50, 477-487, doi:10.1002/mus.24332 (2014).
- 23 Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905-920 (2000).
- 24 Korobeinikova, A. V., Garber, M. B. & Gongadze, G. M. Ribosomal proteins: structure, function, and evolution. *Biochemistry. Biokhimiia* 77, 562-574, doi:10.1134/S0006297912060028 (2012).
- 25 Ban, N. et al. A new system for naming ribosomal proteins. *Current opinion in structural biology* 24, 165-169, doi:10.1016/j.sbi.2014.01.002 (2014).
- 26 Mailliot, J. et al. Crystal Structures of the uL3 Mutant Ribosome: Illustration of the Importance of Ribosomal Proteins for Translation Efficiency. *Journal of molecular biology*, doi:10.1016/j.jmb.2016.02.013 (2016).
- 27 Wilson, D. N. & Nierhaus, K. H. Ribosomal proteins in the spotlight. *Crit Rev Biochem Mol Biol* 40, 243-267, doi:10.1080/10409230500256523 (2005).
- 28 Wawiorka, L. et al. Functional analysis of the uL11 protein impact on translational machinery. *Cell cycle* 15, 1060-1072, doi:10.1080/15384101.2016.1154245 (2016).
- 29 Panopoulos, P., Dresios, J. & Synetos, D. Biochemical evidence of translational infidelity and decreased peptidyltransferase activity by a sarcin/ricin domain mutation of yeast 25S rRNA. *Nucleic acids research* 32, 5398-5408, doi:10.1093/nar/gkh860 (2004).
- 30 Loenarz, C. et al. Hydroxylation of the eukaryotic ribosomal decoding center affects translational accuracy. *Proceedings of the National Academy of Sciences of the United States of America* 111, 4019-4024, doi:10.1073/pnas.1311750111 (2014).
- 31 Aitken, C. E. & Lorsch, J. R. A mechanistic overview of translation initiation in eukaryotes. *Nature structural & molecular biology* 19, 568-576, doi:10.1038/nsmb.2303 (2012).
- 32 Voigts-Hoffmann, F., Klinge, S. & Ban, N. Structural insights into eukaryotic ribosomes and the initiation of translation. *Current opinion in structural biology* 22, 768-777, doi:10.1016/j.sbi.2012.07.010 (2012).
- 33 Jackson, R. J., Hellen, C. U. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature reviews. Molecular cell biology* 11, 113-127, doi:10.1038/nrm2838 (2010).

- 34 Hinnebusch, A. G. The scanning mechanism of eukaryotic translation initiation. *Annual review of biochemistry* 83, 779-812, doi:10.1146/annurev-biochem-060713-035802 (2014).
- 35 Brina, D., Miluzio, A., Ricciardi, S. & Biffo, S. eIF6 anti-association activity is required for ribosome biogenesis, translational control and tumor progression. *Biochimica et biophysica acta* 1849, 830-835, doi:10.1016/j.bbagr.2014.09.010 (2015).
- 36 Weis, F. et al. Mechanism of eIF6 release from the nascent 60S ribosomal subunit. *Nature structural & molecular biology* 22, 914-919, doi:10.1038/nsmb.3112 (2015).
- 37 Haimov, O., Sinvani, H. & Dikstein, R. Cap-dependent, scanning-free translation initiation mechanisms. *Biochimica et biophysica acta* 1849, 1313-1318, doi:10.1016/j.bbagr.2015.09.006 (2015).
- 38 Sinvani, H. et al. Translational tolerance of mitochondrial genes to metabolic energy stress involves TISU and eIF1-eIF4G1 cooperation in start codon selection. *Cell Metab* 21, 479-492, doi:10.1016/j.cmet.2015.02.010 (2015).
- 39 Jang, S. K. et al. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *Journal of virology* 62, 2636-2643 (1988).
- 40 Macejak, D. G. & Sarnow, P. Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature* 353, 90-94, doi:10.1038/353090a0 (1991).
- 41 Guerrero, S. et al. HIV-1 replication and the cellular eukaryotic translation apparatus. *Viruses* 7, 199-218, doi:10.3390/v7010199 (2015).
- 42 Komar, A. A. & Hatzoglou, M. Exploring Internal Ribosome Entry Sites as Therapeutic Targets. *Front Oncol* 5, 233, doi:10.3389/fonc.2015.00233 (2015).
- 43 Gilbert, W. V., Zhou, K., Butler, T. K. & Doudna, J. A. Cap-independent translation is required for starvation-induced differentiation in yeast. *Science* 317, 1224-1227, doi:10.1126/science.1144467 (2007).
- 44 Noller, H. F., Yusupov, M. M., Yusupova, G. Z., Baucom, A. & Cate, J. H. Translocation of tRNA during protein synthesis. *FEBS letters* 514, 11-16 (2002).
- 45 Dever, T. E. & Green, R. The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harbor perspectives in biology* 4, a013706, doi:10.1101/cshperspect.a013706 (2012).
- 46 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).
- 47 Rozov, A., Demeshkina, N., Westhof, E., Yusupov, M. & Yusupova, G. Structural insights into the translational infidelity mechanism. *Nature communications* 6, 7251, doi:10.1038/ncomms8251 (2015).
- 48 Lewin, B. in *Genes IX* (ed Inc Jones and Bartlett Publishers) Ch. 8, (2007).
- 49 Agirrezabala, X. et al. Visualization of the hybrid state of tRNA binding promoted by spontaneous ratcheting of the ribosome. *Molecular cell* 32, 190-197, doi:10.1016/j.molcel.2008.10.001 (2008).
- 50 Sengupta, J. et al. Visualization of the eEF2-80S ribosome transition-state complex by cryo-electron microscopy. *Journal of molecular biology* 382, 179-187, doi:10.1016/j.jmb.2008.07.004 (2008).

- 51 Kisselev, L., Ehrenberg, M. & Frolova, L. Termination of translation: interplay of mRNA, rRNAs and release factors? *The EMBO journal* 22, 175-182, doi:10.1093/emboj/cdg017 (2003).
- 52 Nguyen, V. T., Morange, M. & Bensaude, O. Firefly luciferase luminescence assays using scintillation counters for quantitation in transfected mammalian cells. *Analytical Biochemistry* 171, 404-408, doi:10.1016/0003-2697(88)90505-2 (1988).
- 53 Matheisl, S., Berninghausen, O., Becker, T. & Beckmann, R. Structure of a human translation termination complex. *Nucleic acids research* 43, 8615-8626, doi:10.1093/nar/gkv909 (2015).
- 54 Kryuchkova, P. et al. Two-step model of stop codon recognition by eukaryotic release factor eRF1. *Nucleic acids research* 41, 4573-4586, doi:10.1093/nar/gkt113 (2013).
- 55 Frolova, L. et al. A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor. *Nature* 372, 701-703, doi:10.1038/372701a0 (1994).
- 56 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).
- 57 Frolova, L. Y., Merkulova, T. I. & Kisselev, L. L. Translation termination in eukaryotes: polypeptide release factor eRF1 is composed of functionally and structurally distinct domains. *Rna* 6, 381-390 (2000).
- 58 Kononenko, A. V. et al. Role of the individual domains of translation termination factor eRF1 in GTP binding to eRF3. *Proteins* 70, 388-393, doi:10.1002/prot.21544 (2008).
- 59 Feng, T. et al. Optimal translational termination requires C4 lysyl hydroxylation of eRF1. *Molecular cell* 53, 645-654, doi:10.1016/j.molcel.2013.12.028 (2014).
- 60 Liger, D. et al. Mechanism of activation of methyltransferases involved in translation by the Trm112 'hub' protein. *Nucleic acids research* 39, 6249-6259, doi:10.1093/nar/gkr176 (2011).
- 61 Kallmeyer, A. K., Keeling, K. M. & Bedwell, D. M. Eukaryotic release factor 1 phosphorylation by CK2 protein kinase is dynamic but has little effect on the efficiency of translation termination in *Saccharomyces cerevisiae*. *Eukaryotic cell* 5, 1378-1387, doi:10.1128/EC.00073-06 (2006).
- 62 Saito, K. et al. Omnipotent role of archaeal elongation factor 1 alpha (EF1alpha) in translational elongation and termination, and quality control of protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 107, 19242-19247, doi:10.1073/pnas.1009599107 (2010).
- 63 Kong, C. et al. Crystal structure and functional analysis of the eukaryotic class II release factor eRF3 from *S. pombe*. *Molecular cell* 14, 233-245 (2004).
- 64 Jerbi, S., Jolles, B., Bouceba, T. & Jean-Jean, O. Studies on human eRF3-PABP interaction reveal the influence of eRF3a N-terminal glycine repeat on eRF3-PABP binding affinity and the lower affinity of eRF3a 12-GGC allele involved in cancer susceptibility. *RNA biology* 13, 306-315, doi:10.1080/15476286.2015.1137421 (2016).

- 65 Chauvin, C. & Jean-Jean, O. Proteasomal degradation of human release factor eRF3a regulates translation termination complex formation. *Rna* 14, 240-245, doi:10.1261/rna.728608 (2008).
- 66 des Georges, A. et al. Structure of the mammalian ribosomal pre-termination complex associated with eRF1\*eRF3\*GDPNP. *Nucleic acids research*, doi:10.1093/nar/gkt1279 (2013).
- 67 Kushnirov, V. V. et al. Nucleotide sequence of the SUP2 (SUP35) gene of *Saccharomyces cerevisiae*. *Gene* 66, 45-54 (1988).
- 68 Cosson, B. et al. Poly(A)-Binding Protein Acts in Translation Termination via Eukaryotic Release Factor 3 Interaction and Does Not Influence [PSI<sup>+</sup>] Propagation. *Molecular and cellular biology* 22, 3301-3315, doi:10.1128/mcb.22.10.3301-3315.2002 (2002).
- 69 Liu, J. J., Sondheimer, N. & Lindquist, S. L. Changes in the middle region of Sup35 profoundly alter the nature of epigenetic inheritance for the yeast prion [PSI<sup>+</sup>]. *Proceedings of the National Academy of Sciences of the United States of America* 99 Suppl 4, 16446-16453, doi:10.1073/pnas.252652099 (2002).
- 70 Cheng, Z. et al. Structural insights into eRF3 and stop codon recognition by eRF1. *Genes & development* 23, 1106-1118, doi:10.1101/gad.1770109 (2009).
- 71 Jackson, R. J., Hellen, C. U. & Pestova, T. V. Termination and post-termination events in eukaryotic translation. *Advances in protein chemistry and structural biology* 86, 45-93, doi:10.1016/B978-0-12-386497-0.00002-5 (2012).
- 72 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).
- 73 Hatin, I., Fabret, C., Rousset, J. P. & Namy, O. Molecular dissection of translation termination mechanism identifies two new critical regions in eRF1. *Nucleic acids research* 37, 1789-1798, doi:10.1093/nar/gkp012 (2009).
- 74 Cridge, A. G. et al. Comparison of characteristics and function of translation termination signals between and within prokaryotic and eukaryotic organisms. *Nucleic acids research* 34, 1959-1973, doi:10.1093/nar/gkl074 (2006).
- 75 Eliseev, B., Kryuchkova, P., Alkalaeva, E. & Frolova, L. A single amino acid change of translation termination factor eRF1 switches between bipotent and omnipotent stop-codon specificity. *Nucleic acids research* 39, 599-608, doi:10.1093/nar/gkq759 (2011).
- 76 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).
- 77 Blanchet, S. et al. New insights into stop codon recognition by eRF1. *Nucleic acids research* 43, 3298-3308, doi:10.1093/nar/gkv154 (2015).
- 78 Preis, A. et al. Cryoelectron Microscopic Structures of Eukaryotic Translation Termination Complexes Containing eRF1-eRF3 or eRF1-ABCE1. *Cell reports* 8, 59-65, doi:10.1016/j.celrep.2014.04.058 (2014).
- 79 Klaholz, B. P. Molecular recognition and catalysis in translation termination complexes. *Trends in biochemical sciences* 36, 282-292, doi:10.1016/j.tibs.2011.02.001 (2011).

- 80 Rodnina, M. V. The ribosome as a versatile catalyst: reactions at the peptidyl transferase center. *Current opinion in structural biology* 23, 595-602, doi:10.1016/j.sbi.2013.04.012 (2013).
- 81 Weixlbaumer, A. et al. Insights into translational termination from the structure of RF2 bound to the ribosome. *Science* 322, 953-956, doi:10.1126/science.1164840 (2008).
- 82 Shoemaker, C. J. & Green, R. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proceedings of the National Academy of Sciences of the United States of America* 108, E1392-1398, doi:10.1073/pnas.1113956108 (2011).
- 83 Becker, T. et al. Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature* 482, 501-506, doi:10.1038/nature10829 (2012).
- 84 Brumm, H. et al. Rescue of melanocortin 4 receptor (MC4R) nonsense mutations by aminoglycoside-mediated read-through. *Obesity (Silver Spring)* 20, 1074-1081, doi:10.1038/oby.2011.202 (2012).
- 85 Franckenberg, S., Becker, T. & Beckmann, R. Structural view on recycling of archaeal and eukaryotic ribosomes after canonical termination and ribosome rescue. *Current opinion in structural biology* 22, 786-796, doi:10.1016/j.sbi.2012.08.002 (2012).
- 86 Barthelme, D. et al. Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1. *Proceedings of the National Academy of Sciences of the United States of America* 108, 3228-3233, doi:10.1073/pnas.1015953108 (2011).
- 87 Pisarev, A. V. et al. The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Molecular cell* 37, 196-210, doi:10.1016/j.molcel.2009.12.034 (2010).
- 88 Pisarev, A. V., Hellen, C. U. & Pestova, T. V. Recycling of eukaryotic posttermination ribosomal complexes. *Cell* 131, 286-299, doi:10.1016/j.cell.2007.08.041 (2007).
- 89 Skabkin, M. A. et al. Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes & development* 24, 1787-1801, doi:10.1101/gad.1957510 (2010).
- 90 Calvo, S. E., Pagliarini, D. J. & Mootha, V. K. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proceedings of the National Academy of Sciences of the United States of America* 106, 7507-7512, doi:10.1073/pnas.0810916106 (2009).
- 91 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).
- 92 Poyry, T. A., Kaminski, A. & Jackson, R. J. What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes & development* 18, 62-75, doi:10.1101/gad.276504 (2004).
- 93 Gray, N. K., Hrabalkova, L., Scanlon, J. P. & Smith, R. W. Poly(A)-binding proteins and mRNA localization: who rules the roost? *Biochemical Society transactions* 43, 1277-1284, doi:10.1042/BST20150171 (2015).
- 94 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. *The Journal of biological chemistry* 274, 16677-16680 (1999).

- 95 Keeling, K. M. et al. Leaky termination at premature stop codons antagonizes nonsense-mediated mRNA decay in *S. cerevisiae*. *Rna* 10, 691-703 (2004).
- 96 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. *The EMBO journal* 27, 736-747, doi:10.1038/emboj.2008.17 (2008).
- 97 Czaplinski, K., Majlesi, N., Banerjee, T. & Peltz, S. W. Mtt1 is a Upf1-like helicase that interacts with the translation termination factors and whose overexpression can modulate termination efficiency. *Rna* 6, 730-743 (2000).
- 98 Gross, T. et al. The DEAD-box RNA helicase Dbp5 functions in translation termination. *Science* 315, 646-649, doi:10.1126/science.1134641 (2007).
- 99 Baierlein, C. & Krebber, H. Translation termination: new factors and insights. *RNA biology* 7, 548-550 (2010).
- 100 Chen, C. Y. & Shyu, A. B. Mechanisms of deadenylation-dependent decay. *Wiley interdisciplinary reviews. RNA* 2, 167-183, doi:10.1002/wrna.40 (2011).
- 101 Meyer, S., Temme, C. & Wahle, E. Messenger RNA turnover in eukaryotes: pathways and enzymes. *Crit Rev Biochem Mol Biol* 39, 197-216, doi:10.1080/10409230490513991 (2004).
- 102 Camier, S. & Seraphin, B. [Destroy this (RNA) message after reading it!]. *Medecine sciences : M/S* 23, 850-856, doi:10.1051/medsci/20072310850 (2007).
- 103 Weill, L., Belloc, E., Bava, F. A. & Mendez, R. Translational control by changes in poly(A) tail length: recycling mRNAs. *Nature structural & molecular biology* 19, 577-585, doi:10.1038/nsmb.2311 (2012).
- 104 Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. & Tollervey, D. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases. *Cell* 91, 457-466 (1997).
- 105 Lam, J. K., Chow, M. Y., Zhang, Y. & Leung, S. W. siRNA Versus miRNA as Therapeutics for Gene Silencing. *Mol Ther Nucleic Acids* 4, e252, doi:10.1038/mtna.2015.23 (2015).
- 106 Cheadle, C. et al. Stability regulation of mRNA and the control of gene expression. *Annals of the New York Academy of Sciences* 1058, 196-204, doi:10.1196/annals.1359.026 (2005).
- 107 Wickens, M., Bernstein, D. S., Kimble, J. & Parker, R. A PUF family portrait: 3'UTR regulation as a way of life. *Trends in genetics : TIG* 18, 150-157 (2002).
- 108 Barreau, C., Paillard, L. & Osborne, H. B. AU-rich elements and associated factors: are there unifying principles? *Nucleic acids research* 33, 7138-7150, doi:10.1093/nar/gki1012 (2005).
- 109 Valencia-Sanchez, M. A., Liu, J., Hannon, G. J. & Parker, R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & development* 20, 515-524, doi:10.1101/gad.1399806 (2006).
- 110 Zamore, P. D. & Haley, B. Ribo-gnome: the big world of small RNAs. *Science* 309, 1519-1524, doi:10.1126/science.1111444 (2005).
- 111 Parker, J. S. & Barford, D. Argonaute: A scaffold for the function of short regulatory RNAs. *Trends in biochemical sciences* 31, 622-630, doi:10.1016/j.tibs.2006.09.010 (2006).

- 112 Tang, G. siRNA and miRNA: an insight into RISCs. *Trends in biochemical sciences* 30, 106-114, doi:10.1016/j.tibs.2004.12.007 (2005).
- 113 Fatscher, T., Boehm, V. & Gehring, N. H. Mechanism, factors, and physiological role of nonsense-mediated mRNA decay. *Cell Mol Life Sci* 72, 4523-4544, doi:10.1007/s00018-015-2017-9 (2015).
- 114 Doma, M. K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* 440, 561-564, doi:10.1038/nature04530 (2006).
- 115 Inada, T. Quality control systems for aberrant mRNAs induced by aberrant translation elongation and termination. *Biochimica et biophysica acta* 1829, 634-642, doi:10.1016/j.bbagr.2013.02.004 (2013).
- 116 Graille, M. & Seraphin, B. Surveillance pathways rescuing eukaryotic ribosomes lost in translation. *Nature reviews. Molecular cell biology* 13, 727-735, doi:10.1038/nrm3457 (2012).
- 117 Hudson, G. A., Bloomingdale, R. J. & Znosko, B. M. Thermodynamic contribution and nearest-neighbor parameters of pseudouridine-adenosine base pairs in oligoribonucleotides. *Rna* 19, 1474-1482, doi:10.1261/rna.039610.113 (2013).
- 118 Fourati, Z. & Graille, M. in *Fungal RNA Biology* (eds Ane Sesma & Tobias von der Haar) 195-216 (Springer International Publishing, 2014).
- 119 Becker, T. et al. Structure of the no-go mRNA decay complex Dom34-Hbs1 bound to a stalled 80S ribosome. *Nature structural & molecular biology* 18, 715-720, doi:10.1038/nsmb.2057 (2011).
- 120 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).
- 121 Pisareva, V. P., Skabkin, M. A., Hellen, C. U., Pestova, T. V. & Pisarev, A. V. Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *The EMBO journal* 30, 1804-1817, doi:10.1038/emboj.2011.93 (2011).
- 122 Shao, S., Brown, A., Santhanam, B. & Hegde, R. S. Structure and assembly pathway of the ribosome quality control complex. *Molecular cell* 57, 433-444, doi:10.1016/j.molcel.2014.12.015 (2015).
- 123 van den Elzen, A. M. et al. Dissection of Dom34-Hbs1 reveals independent functions in two RNA quality control pathways. *Nature structural & molecular biology* 17, 1446-1452, doi:10.1038/nsmb.1963 (2010).
- 124 Defenouillere, Q. et al. Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. *Proceedings of the National Academy of Sciences of the United States of America* 110, 5046-5051, doi:10.1073/pnas.1221724110 (2013).
- 125 Shen, P. S. et al. Protein synthesis. Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. *Science* 347, 75-78, doi:10.1126/science.1259724 (2015).
- 126 Yonashiro, R. et al. The Rqc2/Tae2 subunit of the ribosome-associated quality control (RQC) complex marks ribosome-stalled nascent polypeptide chains for aggregation. *eLife* 5, doi:10.7554/eLife.11794 (2016).

- 127 Lykke-Andersen, J. & Bennett, E. J. Protecting the proteome: Eukaryotic cotranslational quality control pathways. *The Journal of cell biology* 204, 467-476, doi:10.1083/jcb.201311103 (2014).
- 128 van Hoof, A. & Parker, R. Messenger RNA degradation: beginning at the end. *Current biology : CB* 12, R285-287 (2002).
- 129 Frischmeyer, P. A. et al. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science* 295, 2258-2261, doi:10.1126/science.1067338 (2002).
- 130 Kong, J. & Liebhaber, S. A. A cell type-restricted mRNA surveillance pathway triggered by ribosome extension into the 3' untranslated region. *Nature structural & molecular biology* 14, 670-676, doi:10.1038/nsmb1256 (2007).
- 131 Akimitsu, N., Tanaka, J. & Pelletier, J. Translation of nonSTOP mRNA is repressed post-initiation in mammalian cells. *The EMBO journal* 26, 2327-2338, doi:10.1038/sj.emboj.7601679 (2007).
- 132 Torres-Torronteras, J. et al. A novel nonstop mutation in TYMP does not induce nonstop mRNA decay in a MNGIE patient with severe neuropathy. *Human mutation* 32, E2061-2068, doi:10.1002/humu.21447 (2011).
- 133 Klauer, A. A. & van Hoof, A. Degradation of mRNAs that lack a stop codon: a decade of nonstop progress. *Wiley interdisciplinary reviews. RNA* 3, 649-660, doi:10.1002/wrna.1124 (2012).
- 134 Saito, S., Hosoda, N. & Hoshino, S. The Hbs1-Dom34 protein complex functions in non-stop mRNA decay in mammalian cells. *The Journal of biological chemistry* 288, 17832-17843, doi:10.1074/jbc.M112.448977 (2013).
- 135 Lu, J. & Deutsch, C. Electrostatics in the ribosomal tunnel modulate chain elongation rates. *Journal of molecular biology* 384, 73-86, doi:10.1016/j.jmb.2008.08.089 (2008).
- 136 Ito-Harashima, S., Kuroha, K., Tatematsu, T. & Inada, T. Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. *Genes & development* 21, 519-524, doi:10.1101/gad.1490207 (2007).
- 137 Matsuda, R., Ikeuchi, K., Nomura, S. & Inada, T. Protein quality control systems associated with no-go and nonstop mRNA surveillance in yeast. *Genes to cells : devoted to molecular & cellular mechanisms* 19, 1-12, doi:10.1111/gtc.12106 (2014).
- 138 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. *Molecular cell* 46, 518-529, doi:10.1016/j.molcel.2012.03.013 (2012).
- 139 Karousis, E. D., Nasif, S. & Muhlemann, O. Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact. *Wiley interdisciplinary reviews. RNA*, doi:10.1002/wrna.1357 (2016).
- 140 Losson, R. & Lacroute, F. Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proceedings of the National Academy of Sciences of the United States of America* 76, 5134-5137 (1979).
- 141 Maquat, L. E., Kinniburgh, A. J., Rachmilewitz, E. A. & Ross, J. Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia. *Cell* 27, 543-553 (1981).

- 142 Peltz, S. W., Brown, A. H. & Jacobson, A. mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. *Genes & development* 7, 1737-1754 (1993).
- 143 Medghalchi, S. M. et al. Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Human molecular genetics* 10, 99-105 (2001).
- 144 Jolly, L. A., Homan, C. C., Jacob, R., Barry, S. & Gecz, J. The UPF3B gene, implicated in intellectual disability, autism, ADHD and childhood onset schizophrenia regulates neural progenitor cell behaviour and neuronal outgrowth. *Human molecular genetics* 22, 4673-4687, doi:10.1093/hmg/ddt315 (2013).
- 145 Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F. & Dietz, H. C. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nature genetics* 36, 1073-1078, doi:10.1038/ng1429 (2004).
- 146 Yepiskoposyan, H., Aeschmann, F., Nilsson, D., Okoniewski, M. & Muhlemann, O. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *Rna* 17, 2108-2118, doi:10.1261/rna.030247.111 (2011).
- 147 Huang, L. et al. RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. *Molecular cell* 43, 950-961, doi:10.1016/j.molcel.2011.06.031 (2011).
- 148 Leeds, P., Peltz, S. W., Jacobson, A. & Culbertson, M. R. The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. *Genes & development* 5, 2303-2314 (1991).
- 149 Pulak, R. & Anderson, P. mRNA surveillance by the *Caenorhabditis elegans* smg genes. *Genes & development* 7, 1885-1897 (1993).
- 150 Hwang, J., Sato, H., Tang, Y., Matsuda, D. & Maquat, L. E. UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps. *Molecular cell* 39, 396-409, doi:10.1016/j.molcel.2010.07.004 (2010).
- 151 Hentze, M. W. & Izaurralde, E. Making sense of nonsense. *Nature structural & molecular biology* 20, 651-653, doi:10.1038/nsmb.2601 (2013).
- 152 Maquat, L. E. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nature reviews. Molecular cell biology* 5, 89-99, doi:10.1038/nrm1310 (2004).
- 153 Gehring, N. H., Lamprinaki, S., Kulozik, A. E. & Hentze, M. W. Disassembly of exon junction complexes by PYM. *Cell* 137, 536-548, doi:10.1016/j.cell.2009.02.042 (2009).
- 154 Durand, S. & Lykke-Andersen, J. Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. *Nature structural & molecular biology* 20, 702-709, doi:10.1038/nsmb.2575 (2013).
- 155 Ishigaki, Y., Li, X., Serin, G. & Maquat, L. E. Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell* 106, 607-617 (2001).
- 156 Hosoda, N., Kim, Y. K., Lejeune, F. & Maquat, L. E. CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells. *Nature structural & molecular biology* 12, 893-901, doi:10.1038/nsmb995 (2005).

- 157 Rufener, S. C. & Muhlemann, O. eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nature structural & molecular biology* 20, 710-717, doi:10.1038/nsmb.2576 (2013).
- 158 Metzger, S., Herzog, V. A., Ruepp, M. D. & Muhlemann, O. Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. *Rna*, doi:10.1261/rna.038893.113 (2013).
- 159 Schweingruber, C., Rufener, S. C., Zund, D., Yamashita, A. & Muhlemann, O. Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochimica et biophysica acta* 1829, 612-623, doi:10.1016/j.bbagr.2013.02.005 (2013).
- 160 Stalder, L. & Muhlemann, O. The meaning of nonsense. *Trends in cell biology* 18, 315-321, doi:10.1016/j.tcb.2008.04.005 (2008).
- 161 Peixeiro, I. et al. Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. *Nucleic acids research* 40, 1160-1173, doi:10.1093/nar/gkr820 (2012).
- 162 Silva, A. L., Ribeiro, P., Inacio, A., Liebhaber, S. A. & Romao, L. Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. *Rna* 14, 563-576, doi:10.1261/rna.815108 (2008).
- 163 Kononenko, A. V. et al. GTP-dependent structural rearrangement of the eRF1:eRF3 complex and eRF3 sequence motifs essential for PABP binding. *Nucleic acids research* 38, 548-558, doi:10.1093/nar/gkp908 (2010).
- 164 Singh, G., Rebbapragada, I. & Lykke-Andersen, J. A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. *PLoS biology* 6, e111, doi:10.1371/journal.pbio.0060111 (2008).
- 165 Hurt, J. A., Robertson, A. D. & Burge, C. B. Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. *Genome research* 23, 1636-1650, doi:10.1101/gr.157354.113 (2013).
- 166 Zund, D., Gruber, A. R., Zvolan, M. & Muhlemann, O. Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. *Nature structural & molecular biology* 20, 936-943, doi:10.1038/nsmb.2635 (2013).
- 167 Melero, R. et al. The RNA helicase DHX34 functions as a scaffold for SMG1-mediated UPF1 phosphorylation. *Nature communications* 7, 10585, doi:10.1038/ncomms10585 (2016).
- 168 Chamieh, H., Ballut, L., Bonneau, F. & Le Hir, H. NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nature structural & molecular biology* 15, 85-93, doi:10.1038/nsmb1330 (2008).
- 169 Chakrabarti, S. et al. Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Molecular cell* 41, 693-703, doi:10.1016/j.molcel.2011.02.010 (2011).
- 170 Melero, R. et al. Structures of SMG1-UPFs complexes: SMG1 contributes to regulate UPF2-dependent activation of UPF1 in NMD. *Structure* 22, 1105-1119, doi:10.1016/j.str.2014.05.015 (2014).
- 171 Lykke-Andersen, S. & Jensen, T. H. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nature reviews. Molecular cell biology* 16, 665-677, doi:10.1038/nrm4063 (2015).

- 172 Nicholson, P., Josi, C., Kurosawa, H., Yamashita, A. & Muhlemann, O. A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. *Nucleic acids research* 42, 9217-9235, doi:10.1093/nar/gku645 (2014).
- 173 Eberle, A. B., Lykke-Andersen, S., Muhlemann, O. & Jensen, T. H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nature structural & molecular biology* 16, 49-55, doi:10.1038/nsmb.1530 (2009).
- 174 Boehm, V., Haberman, N., Ottens, F., Ule, J. & Gehring, N. H. 3' UTR length and messenger ribonucleoprotein composition determine endocleavage efficiencies at termination codons. *Cell reports* 9, 555-568, doi:10.1016/j.celrep.2014.09.012 (2014).
- 175 Jonas, S., Weichenrieder, O. & Izaurralde, E. An unusual arrangement of two 14-3-3-like domains in the SMG5-SMG7 heterodimer is required for efficient nonsense-mediated mRNA decay. *Genes & development* 27, 211-225, doi:10.1101/gad.206672.112 (2013).
- 176 Cho, H., Kim, K. M. & Kim, Y. K. Human proline-rich nuclear receptor coregulatory protein 2 mediates an interaction between mRNA surveillance machinery and decapping complex. *Molecular cell* 33, 75-86, doi:10.1016/j.molcel.2008.11.022 (2009).
- 177 Lareau, L. F., Inada, M., Green, R. E., Wengrod, J. C. & Brenner, S. E. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* 446, 926-929, doi:10.1038/nature05676 (2007).
- 178 Kunkel, T. A. & Bebenek, K. DNA replication fidelity. *Annual review of biochemistry* 69, 497-529, doi:10.1146/annurev.biochem.69.1.497 (2000).
- 179 Sydow, J. F. & Cramer, P. RNA polymerase fidelity and transcriptional proofreading. *Current opinion in structural biology* 19, 732-739, doi:10.1016/j.sbi.2009.10.009 (2009).
- 180 Kurland, C. & Gallant, J. Errors of heterologous protein expression. *Current opinion in biotechnology* 7, 489-493 (1996).
- 181 Drummond, D. A. & Wilke, C. O. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* 134, 341-352, doi:10.1016/j.cell.2008.05.042 (2008).
- 182 Drummond, D. A. & Wilke, C. O. The evolutionary consequences of erroneous protein synthesis. *Nature reviews. Genetics* 10, 715-724, doi:10.1038/nrg2662 (2009).
- 183 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).
- 184 Schmeing, T. M. & Ramakrishnan, V. What recent ribosome structures have revealed about the mechanism of translation. *Nature* 461, 1234-1242, doi:10.1038/nature08403 (2009).
- 185 Triman, K. L. Mutational Analysis of the Ribosome. 58, 89-119, doi:10.1016/s0065-2660(06)58004-6 (2007).
- 186 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).

- 187 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).
- 188 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).
- 189 Rodnina, M. V. & Wintermeyer, W. Protein Elongation, Co-translational Folding and Targeting. *Journal of molecular biology*, doi:10.1016/j.jmb.2016.03.022 (2016).
- 190 Sweeney, B. A., Roy, P. & Leontis, N. B. An introduction to recurrent nucleotide interactions in RNA. *Wiley interdisciplinary reviews. RNA* 6, 17-45 (2015).
- 191 Yusupov, M. M. et al. Crystal structure of the ribosome at 5.5 Å resolution. *Science* 292, 883-896, doi:10.1126/science.1060089 (2001).
- 192 Voorhees, R. M. & Ramakrishnan, V. Structural basis of the translational elongation cycle. *Annual review of biochemistry* 82, 203-236, doi:10.1146/annurev-biochem-113009-092313 (2013).
- 193 Velichutina, I. V., Hong, J. Y., Mesecar, A. D., Chernoff, Y. O. & Liebman, S. W. Genetic interaction between yeast *Saccharomyces cerevisiae* release factors and the decoding region of 18 S rRNA. *Journal of molecular biology* 305, 715-727, doi:10.1006/jmbi.2000.4329 (2001).
- 194 O'Connor, M., Thomas, C. L., Zimmermann, R. A. & Dahlberg, A. E. Decoding fidelity at the ribosomal A and P sites: influence of mutations in three different regions of the decoding domain in 16S rRNA. *Nucleic acids research* 25, 1185-1193 (1997).
- 195 Chernoff, Y. O., Newnam, G. P. & Liebman, S. W. The translational function of nucleotide C1054 in the small subunit rRNA is conserved throughout evolution: genetic evidence in yeast. *Proceedings of the National Academy of Sciences of the United States of America* 93, 2517-2522 (1996).
- 196 Baudin-Baillieu, A. et al. Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic acids research* 37, 7665-7677, doi:10.1093/nar/gkp816 (2009).
- 197 Sharma, D., Cukras, A. R., Rogers, E. J., Southworth, D. R. & Green, R. Mutational analysis of S12 protein and implications for the accuracy of decoding by the ribosome. *Journal of molecular biology* 374, 1065-1076, doi:10.1016/j.jmb.2007.10.003 (2007).
- 198 Atkins, J. F., Elseviers, D. & Gorini, L. Low activity of -galactosidase in frameshift mutants of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 69, 1192-1195 (1972).
- 199 Ogle, J. M. et al. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* 292, 897-902, doi:10.1126/science.1060612 (2001).
- 200 Ogle, J. M., Murphy, F. V., Tarry, M. J. & Ramakrishnan, V. Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* 111, 721-732 (2002).
- 201 Jenner, L. B., Demeshkina, N., Yusupova, G. & Yusupov, M. Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nature structural & molecular biology* 17, 555-560, doi:10.1038/nsmb.1790 (2010).

- 202 Rozov, A. et al. Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nature communications* 7, 10457, doi:10.1038/ncomms10457 (2016).
- 203 Lu, Q. L. et al. Systemic delivery of antisense oligoribonucleotide restores dystrophin expression in body-wide skeletal muscles. *Proceedings of the National Academy of Sciences of the United States of America* 102, 198-203, doi:10.1073/pnas.0406700102 (2005).
- 204 Rozov, A., Westhof, E., Yusupov, M. & Yusupova, G. The ribosome prohibits the G\*U wobble geometry at the first position of the codon-anticodon helix. *Nucleic acids research*, doi:10.1093/nar/gkw431 (2016).
- 205 Baranov, P. V., Atkins, J. F. & Yordanova, M. M. Augmented genetic decoding: global, local and temporal alterations of decoding processes and codon meaning. *Nature reviews. Genetics* 16, 517-529, doi:10.1038/nrg3963 (2015).
- 206 Namy, O., Rousset, J. P., Naphthine, S. & Brierley, I. Reprogrammed genetic decoding in cellular gene expression. *Molecular cell* 13, 157-168 (2004).
- 207 Caliskan, N., Katunin, V. I., Belardinelli, R., Peske, F. & Rodnina, M. V. Programmed -1 Frameshifting by Kinetic Partitioning during Impeded Translocation. *Cell* 157, 1619-1631, doi:10.1016/j.cell.2014.04.041 (2014).
- 208 Atkins, J. F. & Gesteland, R. F. (eds John F. Atkins & Raymond F. Gesteland) (Springer, New York ;, 2010).
- 209 Yordanova, M. M., Wu, C., Andreev, D. E., Sachs, M. S. & Atkins, J. F. A Nascent Peptide Signal Responsive to Endogenous Levels of Polyamines Acts to Stimulate Regulatory Frameshifting on Antizyme mRNA. *The Journal of biological chemistry* 290, 17863-17878, doi:10.1074/jbc.M115.647065 (2015).
- 210 Kontos, H., Naphthine, S. & Brierley, I. Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. *Molecular and cellular biology* 21, 8657-8670, doi:10.1128/MCB.21.24.8657-8670.2001 (2001).
- 211 Takyar, S., Hickerson, R. P. & Noller, H. F. mRNA helicase activity of the ribosome. *Cell* 120, 49-58, doi:S0092867404011468 10.1016/j.cell.2004.11.042 (2005).
- 212 Namy, O., Moran, S. J., Stuart, D. I., Gilbert, R. J. & Brierley, I. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* 441, 244-247, doi:10.1038/nature04735 (2006).
- 213 Caliskan, N., Peske, F. & Rodnina, M. V. Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting. *Trends in biochemical sciences*, doi:10.1016/j.tibs.2015.03.006 (2015).
- 214 Belardinelli, R. et al. Choreography of molecular movements during ribosome progression along mRNA. *Nature structural & molecular biology* 23, 342-348, doi:10.1038/nsmb.3193 (2016).
- 215 Manktelow, E., Shigemoto, K. & Brierley, I. Characterization of the frameshift signal of Edr, a mammalian example of programmed -1 ribosomal frameshifting. *Nucleic acids research* 33, 1553-1563, doi:10.1093/nar/gki299 (2005).
- 216 Wills, N. M., Moore, B., Hammer, A., Gesteland, R. F. & Atkins, J. F. A functional -1 ribosomal frameshift signal in the human paraneoplastic Ma3 gene. *The Journal of biological chemistry* 281, 7082-7088, doi:10.1074/jbc.M511629200 (2006).

- 217 Sundararajan, A., Michaud, W. A., Qian, Q., Stahl, G. & Farabaugh, P. J. Near-cognate peptidyl-tRNAs promote +1 programmed translational frameshifting in yeast. *Molecular cell* 4, 1005-1015 (1999).
- 218 Vimaladithan, A. & Farabaugh, P. J. Special peptidyl-tRNA molecules can promote translational frameshifting without slippage. *Molecular and cellular biology* 14, 8107-8116 (1994).
- 219 Taliaferro, D. & Farabaugh, P. J. An mRNA sequence derived from the yeast EST3 gene stimulates programmed +1 translational frameshifting. *Rna* 13, 606-613, doi:10.1261/rna.412707 (2007).
- 220 Adamski, F. M., Donly, B. C. & Tate, W. P. Competition between frameshifting, termination and suppression at the frameshift site in the Escherichia coli release factor-2 mRNA. *Nucleic acids research* 21, 5074-5078 (1993).
- 221 Namy, O. et al. Epigenetic control of polyamines by the prion [PSI<sup>+</sup>]. *Nature cell biology* 10, 1069-1075, doi:10.1038/ncb1766 (2008).
- 222 Bidou, L., Allamand, V., Rousset, J. P. & Namy, O. Sense from nonsense: therapies for premature stop codon diseases. *Trends in molecular medicine* 18, 679-688, doi:10.1016/j.molmed.2012.09.008 (2012).
- 223 Belcourt, M. F. & Farabaugh, P. J. Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell* 62, 339-352 (1990).
- 224 Klobutcher, L. A. & Farabaugh, P. J. Shifty ciliates: frequent programmed translational frameshifting in euplotids. *Cell* 111, 763-766 (2002).
- 225 Namy, O., Hatin, I. & Rousset, J. P. Impact of the six nucleotides downstream of the stop codon on translation termination. *EMBO reports* 2, 787-793, doi:10.1093/embo-reports/kve176 (2001).
- 226 Skuzeski, J. M., Nichols, L. M., Gesteland, R. F. & Atkins, J. F. The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. *Journal of molecular biology* 218, 365-373 (1991).
- 227 Cassan, M. & Rousset, J. P. UAG readthrough in mammalian cells: effect of upstream and downstream stop codon contexts reveal different signals. *BMC molecular biology* 2, 3 (2001).
- 228 Bonetti, B., Fu, L., Moon, J. & Bedwell, D. M. The efficiency of translation termination is determined by a synergistic interplay between upstream and downstream sequences in *Saccharomyces cerevisiae*. *Journal of molecular biology* 251, 334-345, doi:10.1006/jmbi.1995.0438 (1995).
- 229 Namy, O., Duchateau-Nguyen, G. & Rousset, J. P. Translational readthrough of the PDE2 stop codon modulates cAMP levels in *Saccharomyces cerevisiae*. *Molecular microbiology* 43, 641-652 (2002).
- 230 Castets, P., Lescure, A., Guicheney, P. & Allamand, V. Selenoprotein N in skeletal muscle: from diseases to function. *Journal of molecular medicine* 90, 1095-1107, doi:10.1007/s00109-012-0896-x (2012).
- 231 Beznoskova, P., Wagner, S., Jansen, M. E., von der Haar, T. & Valasek, L. S. Translation initiation factor eIF3 promotes programmed stop codon readthrough. *Nucleic acids research* 43, 5099-5111, doi:10.1093/nar/gkv421 (2015).

- 232 Dinman, J. D. Mechanisms and implications of programmed translational frameshifting. *Wiley interdisciplinary reviews. RNA* 3, 661-673, doi:10.1002/wrna.1126 (2012).
- 233 Bulteau, A. L. & Chavatte, L. Update on selenoprotein biosynthesis. *Antioxid Redox Signal* 23, 775-794, doi:10.1089/ars.2015.6391 (2015).
- 234 Gupta, M. & Copeland, P. R. Functional analysis of the interplay between translation termination, selenocysteine codon context, and selenocysteine insertion sequence-binding protein 2. *The Journal of biological chemistry* 282, 36797-36807, doi:10.1074/jbc.M707061200 (2007).
- 235 Wills, N. M. et al. Translational bypassing without peptidyl-tRNA anticodon scanning of coding gap mRNA. *The EMBO journal* 27, 2533-2544, doi:10.1038/emboj.2008.170 (2008).
- 236 Gaston, M. A., Jiang, R. & Krzycki, J. A. Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Current opinion in microbiology* 14, 342-349, doi:10.1016/j.mib.2011.04.001 (2011).
- 237 Herr, A. J., Atkins, J. F. & Gesteland, R. F. Coupling of open reading frames by translational bypassing. *Annual review of biochemistry* 69, 343-372, doi:10.1146/annurev.biochem.69.1.343 (2000).
- 238 Bidou, L., Sandra, B. & Olivier, N. in *eLS Translational readthrough of premature termination codons in a therapeutic context* (John Wiley & Sons, Ltd, 2016).
- 239 Manuvakhova, M., Keeling, K. & Bedwell, D. M. Aminoglycoside antibiotics mediate context-dependent suppression of termination codons in a mammalian translation system. *Rna* 6, 1044-1055 (2000).
- 240 Floquet, C., Hatin, I., Rousset, J. P. & Bidou, L. Statistical analysis of readthrough levels for nonsense mutations in mammalian cells reveals a major determinant of response to gentamicin. *PLoS genetics* 8, e1002608, doi:10.1371/journal.pgen.1002608 (2012).
- 241 Fearon, K., McClendon, V., Bonetti, B. & Bedwell, D. M. Premature translation termination mutations are efficiently suppressed in a highly conserved region of yeast Ste6p, a member of the ATP-binding cassette (ABC) transporter family. *The Journal of biological chemistry* 269, 17802-17808 (1994).
- 242 Feng, Y. X., Copeland, T. D., Oroszlan, S., Rein, A. & Levin, J. G. Identification of amino acids inserted during suppression of UAA and UGA termination codons at the gag-pol junction of Moloney murine leukemia virus. *Proceedings of the National Academy of Sciences of the United States of America* 87, 8860-8863 (1990).
- 243 Urban, C. & Beier, H. Cysteine tRNAs of plant origin as novel UGA suppressors. *Nucleic acids research* 23, 4591-4597 (1995).
- 244 Nilsson, M. & Ryden-Aulin, M. Glutamine is incorporated at the nonsense codons UAG and UAA in a suppressor-free Escherichia coli strain. *Biochimica et biophysica acta* 1627, 1-6 (2003).
- 245 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 research* 42, 10061-10072, doi:10.1093/nar/gku663 (2014).
- 246 Dabrowski, M., Bukowy-Bieryllo, Z. & Zietkiewicz, E. Translational readthrough potential of natural termination codons in eucaryotes--The impact of RNA sequence. *RNA biology* 12, 950-958, doi:10.1080/15476286.2015.1068497 (2015).

- 247 Roy, B., Leszyk, J. D., Mangus, D. A. & Jacobson, A. Nonsense suppression by near-cognate tRNAs employs alternative base pairing at codon positions 1 and 3. *Proceedings of the National Academy of Sciences of the United States of America* 112, 3038-3043, doi:10.1073/pnas.1424127112 (2015).
- 248 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 research* 18, 6339-6345 (1990).
- 249 Loughran, G. et al. Evidence of efficient stop codon readthrough in four mammalian genes. *Nucleic acids research* 42, 8928-8938, doi:10.1093/nar/gku608 (2014).
- 250 Arkov, A. L., Korolev, S. V. & Kisselev, L. L. 5' contexts of Escherichia coli and human termination codons are similar. *Nucleic acids research* 23, 4712-4716 (1995).
- 251 Tork, S., Hatin, I., Rousset, J. P. & Fabret, C. The major 5' determinant in stop codon read-through involves two adjacent adenines. *Nucleic acids research* 32, 415-421, doi:10.1093/nar/gkh201 (2004).
- 252 Mottagui-Tabar, S., Tuite, M. F. & Isaksson, L. A. The influence of 5' codon context on translation termination in Saccharomyces cerevisiae. *European journal of biochemistry / FEBS* 257, 249-254 (1998).
- 253 Beier, H. & Grimm, M. Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. *Nucleic acids research* 29, 4767-4782 (2001).
- 254 Bertram, G., Innes, S., Minella, O., Richardson, J. & Stansfield, I. Endless possibilities: translation termination and stop codon recognition. *Microbiology* 147, 255-269 (2001).
- 255 Capecchi, M. R., Hughes, S. H. & Wahl, G. M. Yeast super-suppressors are altered tRNAs capable of translating a nonsense codon in vitro. *Cell* 6, 269-277 (1975).
- 256 Rijal, K., Maraia, R. J. & Arimbasseri, A. G. A methods review on use of nonsense suppression to study 3' end formation and other aspects of tRNA biogenesis. *Gene* 556, 35-50, doi:10.1016/j.gene.2014.11.034 (2015).
- 257 Schull, C. & Beier, H. Three Tetrahymena tRNA(Gln) isoacceptors as tools for studying unorthodox codon recognition and codon context effects during protein synthesis in vitro. *Nucleic acids research* 22, 1974-1980 (1994).
- 258 Sprinzl, M., Horn, C., Brown, M., loudovitch, A. & Steinberg, S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic acids research* 26, 148-153 (1998).
- 259 Kuchino, Y., Beier, H., Akita, N. & Nishimura, S. Natural UAG suppressor glutamine tRNA is elevated in mouse cells infected with Moloney murine leukemia virus. *Proceedings of the National Academy of Sciences of the United States of America* 84, 2668-2672 (1987).
- 260 Spenkuch, F., Motorin, Y. & Helm, M. Pseudouridine: still mysterious, but never a fake (uridine)! *RNA biology* 11, 1540-1554, doi:10.4161/15476286.2014.992278 (2014).
- 261 Valle, R. P., Morch, M. D. & Haenni, A. L. Novel amber suppressor tRNAs of mammalian origin. *The EMBO journal* 6, 3049-3055 (1987).
- 262 Triana-Alonso, F. J., Chakraburttty, K. & Nierhaus, K. H. The elongation factor 3 unique in higher fungi and essential for protein biosynthesis is an E site factor. *The Journal of biological chemistry* 270, 20473-20478 (1995).

- 263 Zerfass, K. & Beier, H. The leaky UGA termination codon of tobacco rattle virus RNA is suppressed by tobacco chloroplast and cytoplasmic tRNAs(Trp) with CmCA anticodon. *The EMBO journal* 11, 4167-4173 (1992).
- 264 Grosjean, H. & Chantrenne, H. On codon- anticodon interactions. *Mol Biol Biochem Biophys* 32, 347-367 (1980).
- 265 Ericson, J. U. & Bjork, G. R. tRNA anticodons with the modified nucleoside 2-methylthio-N6-(4-hydroxyisopentenyl)adenosine distinguish between bases 3' of the codon. *Journal of molecular biology* 218, 509-516 (1991).
- 266 Li, G. P. & Rice, C. M. Mutagenesis of the in-frame opal termination codon preceding nsP4 of Sindbis virus: studies of translational readthrough and its effect on virus replication. *Journal of virology* 63, 1326-1337 (1989).
- 267 Ge, J. & Yu, Y. T. RNA pseudouridylation: new insights into an old modification. *Trends in biochemical sciences* 38, 210-218, doi:10.1016/j.tibs.2013.01.002 (2013).
- 268 Kierzek, E. et al. The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic acids research* 42, 3492-3501, doi:10.1093/nar/gkt1330 (2014).
- 269 Hamma, T. & Ferre-D'Amare, A. R. Pseudouridine synthases. *Chem Biol* 13, 1125-1135, doi:10.1016/j.chembiol.2006.09.009 (2006).
- 270 Watanabe, Y. & Gray, M. W. Evolutionary appearance of genes encoding proteins associated with box H/ACA snoRNAs: cbf5p in *Euglena gracilis*, an early diverging eukaryote, and candidate Gar1p and Nop10p homologs in archaeobacteria. *Nucleic acids research* 28, 2342-2352 (2000).
- 271 Huang, S., Mahanta, N., Begley, T. P. & Ealick, S. E. Pseudouridine monophosphate glycosidase: a new glycosidase mechanism. *Biochemistry* 51, 9245-9255, doi:10.1021/bi3006829 (2012).
- 272 Preumont, A., Rzem, R., Vertommen, D. & Van Schaftingen, E. HDHD1, which is often deleted in X-linked ichthyosis, encodes a pseudouridine-5'-phosphatase. *The Biochemical journal* 431, 237-244, doi:10.1042/BJ20100174 (2010).
- 273 Yarian, C. S. et al. Structural and functional roles of the N1- and N3-protons of psi at tRNA's position 39. *Nucleic acids research* 27, 3543-3549 (1999).
- 274 Durant, P. C. & Davis, D. R. Stabilization of the anticodon stem-loop of tRNA<sup>Lys,3</sup> by an A+C base-pair and by pseudouridine. *Journal of molecular biology* 285, 115-131, doi:10.1006/jmbi.1998.2297 (1999).
- 275 Lecointe, F. et al. Lack of pseudouridine 38/39 in the anticodon arm of yeast cytoplasmic tRNA decreases in vivo recoding efficiency. *The Journal of biological chemistry* 277, 30445-30453, doi:10.1074/jbc.M203456200 (2002).
- 276 Hu, L. F. et al. Hydrogen sulfide regulates Na<sup>+</sup>/H<sup>+</sup> exchanger activity via stimulation of phosphoinositide 3-kinase/Akt and protein kinase G pathways. *J Pharmacol Exp Ther* 339, 726-735, doi:10.1124/jpet.111.184754 (2011).
- 277 Wilson, D. N. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat Rev Microbiol* 12, 35-48, doi:10.1038/nrmicro3155 (2014).
- 278 Jack, K. et al. rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Molecular cell* 44, 660-666, doi:10.1016/j.molcel.2011.09.017 (2011).
- 279 Karijolich, J. & Yu, Y. T. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* 474, 395-398, doi:10.1038/nature10165 (2011).

- 280 Parisien, M., Yi, C. & Pan, T. Rationalization and prediction of selective decoding of pseudouridine-modified nonsense and sense codons. *Rna* 18, 355-367, doi:10.1261/rna.031351.111 (2012).
- 281 Kariko, K. et al. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Molecular therapy : the journal of the American Society of Gene Therapy* 16, 1833-1840, doi:10.1038/mt.2008.200 (2008).
- 282 Hainrichson, M., Nudelman, I. & Baasov, T. Designer aminoglycosides: the race to develop improved antibiotics and compounds for the treatment of human genetic diseases. *Organic & biomolecular chemistry* 6, 227-239, doi:10.1039/b712690p (2008).
- 283 Schatz, A., Bugie, E. & Waksman, S. A. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. 1944. *Clin Orthop Relat Res*, 3-6 (1944).
- 284 Moazed, D. & Noller, H. F. Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature* 327, 389-394, doi:10.1038/327389a0 (1987).
- 285 Borovinskaya, M. A. et al. Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nature structural & molecular biology* 14, 727-732, doi:10.1038/nsmb1271 (2007).
- 286 Carter, A. P. et al. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 407, 340-348, doi:10.1038/35030019 (2000).
- 287 Hermann, T. Aminoglycoside antibiotics: old drugs and new therapeutic approaches. *Cell Mol Life Sci* 64, 1841-1852, doi:10.1007/s00018-007-7034-x (2007).
- 288 Karijolich, J. & Yu, Y. T. Therapeutic suppression of premature termination codons: mechanisms and clinical considerations (review). *International journal of molecular medicine* 34, 355-362, doi:10.3892/ijmm.2014.1809 (2014).
- 289 Pape, T., Wintermeyer, W. & Rodnina, M. V. Conformational switch in the decoding region of 16S rRNA during aminoacyl-tRNA selection on the ribosome. *Nature structural biology* 7, 104-107, doi:10.1038/72364 (2000).
- 290 Francois, B. et al. Crystal structures of complexes between aminoglycosides and decoding A site oligonucleotides: role of the number of rings and positive charges in the specific binding leading to miscoding. *Nucleic acids research* 33, 5677-5690, doi:10.1093/nar/gki862 (2005).
- 291 Garreau de Loubresse, N. et al. Structural basis for the inhibition of the eukaryotic ribosome. *Nature* 513, 517-522, doi:10.1038/nature13737 (2014).
- 292 Fan-Minogue, H. & Bedwell, D. M. Eukaryotic ribosomal RNA determinants of aminoglycoside resistance and their role in translational fidelity. *Rna* 14, 148-157, doi:10.1261/rna.805208 (2008).
- 293 Pierdet, A., Nédélec, L., Delaroff, V. & Allais, A. Synthèse totale de la ( $\pm$ ) negamycine. *Tetrahedron* 36, 1763-1772, doi:http://dx.doi.org/10.1016/0040-4020(80)80072-X (1980).
- 294 Polikanov, Y. S. et al. Negamycin interferes with decoding and translocation by simultaneous interaction with rRNA and tRNA. *Molecular cell* 56, 541-550, doi:10.1016/j.molcel.2014.09.021 (2014).

- 295 Arakawa, M. et al. Negamycin restores dystrophin expression in skeletal and cardiac muscles of mdx mice. *Journal of biochemistry* 134, 751-758 (2003).
- 296 Olivier, N. B. et al. Negamycin induces translational stalling and miscoding by binding to the small subunit head domain of the Escherichia coli ribosome. *Proceedings of the National Academy of Sciences of the United States of America* 111, 16274-16279, doi:10.1073/pnas.1414401111 (2014).
- 297 Karimi, R. & Ehrenberg, M. Dissociation rate of cognate peptidyl-tRNA from the A-site of hyper-accurate and error-prone ribosomes. *European journal of biochemistry / FEBS* 226, 355-360 (1994).
- 298 Peske, F., Savelsbergh, A., Katunin, V. I., Rodnina, M. V. & Wintermeyer, W. Conformational changes of the small ribosomal subunit during elongation factor G-dependent tRNA-mRNA translocation. *Journal of molecular biology* 343, 1183-1194, doi:10.1016/j.jmb.2004.08.097 (2004).
- 299 Mort, M., Ivanov, D., Cooper, D. N. & Chuzhanova, N. A. A meta-analysis of nonsense mutations causing human genetic disease. *Human mutation* 29, 1037-1047, doi:10.1002/humu.20763 (2008).
- 300 Gorgoni, B., Marshall, E., McFarland, M. R., Romano, M. C. & Stansfield, I. Controlling translation elongation efficiency: tRNA regulation of ribosome flux on the mRNA. *Biochemical Society transactions* 42, 160-165, doi:10.1042/BST20130132 (2014).
- 301 Howard, M., Frizzell, R. A. & Bedwell, D. M. Aminoglycoside antibiotics restore CFTR function by overcoming premature stop mutations. *Nature medicine* 2, 467-469 (1996).
- 302 Bordeira-Carrico, R., Pego, A. P., Santos, M. & Oliveira, C. Cancer syndromes and therapy by stop-codon readthrough. *Trends in molecular medicine* 18, 667-678, doi:10.1016/j.molmed.2012.09.004S1471-4914(12)00178-5 [pii] (2012).
- 303 Kellermayer, R. Translational readthrough induction of pathogenic nonsense mutations. *Eur J Med Genet* 49, 445-450, doi:10.1016/j.ejmg.2006.04.003 (2006).
- 304 Lee, H. L. & Dougherty, J. P. Pharmaceutical therapies to recode nonsense mutations in inherited diseases. *Pharmacology & therapeutics* 136, 227-266, doi:10.1016/j.pharmthera.2012.07.007 (2012).
- 305 Kole, R. & Krieg, A. M. Exon skipping therapy for Duchenne muscular dystrophy. *Advanced drug delivery reviews* 87, 104-107, doi:10.1016/j.addr.2015.05.008 (2015).
- 306 Falzarano, M. S., Scotton, C., Passarelli, C. & Ferlini, A. Duchenne Muscular Dystrophy: From Diagnosis to Therapy. *Molecules* 20, 18168-18184, doi:10.3390/molecules201018168 (2015).
- 307 Muntoni, F., Torelli, S. & Ferlini, A. Dystrophin and mutations: one gene, several proteins, multiple phenotypes. *Lancet Neurol* 2, 731-740 (2003).
- 308 Wein, N., Alfano, L. & Flanigan, K. M. Genetics and emerging treatments for Duchenne and Becker muscular dystrophy. *Pediatr Clin North Am* 62, 723-742, doi:10.1016/j.pcl.2015.03.008 (2015).
- 309 Ferlini, A., Neri, M. & Gualandi, F. The medical genetics of dystrophinopathies: molecular genetic diagnosis and its impact on clinical practice. *Neuromuscul Disord* 23, 4-14, doi:10.1016/j.nmd.2012.09.002 (2013).

- 310 Malik, V., Rodino-Klapac, L. R., Viollet, L. & Mendell, J. R. Aminoglycoside-induced mutation suppression (stop codon readthrough) as a therapeutic strategy for Duchenne muscular dystrophy. *Ther Adv Neurol Disord* 3, 379-389, doi:10.1177/1756285610388693 (2010).
- 311 Sicinski, P. et al. The molecular basis of muscular dystrophy in the mdx mouse: a point mutation. *Science* 244, 1578-1580 (1989).
- 312 Quon, B. S. & Rowe, S. M. New and emerging targeted therapies for cystic fibrosis. *BMJ* 352, i859, doi:10.1136/bmj.i859 (2016).
- 313 Grounds, M. D., Radley, H. G., Lynch, G. S., Nagaraju, K. & De Luca, A. Towards developing standard operating procedures for pre-clinical testing in the mdx mouse model of Duchenne muscular dystrophy. *Neurobiol Dis* 31, 1-19, doi:10.1016/j.nbd.2008.03.008 (2008).
- 314 Falzarano, M. S., Passarelli, C. & Ferlini, A. Nanoparticle delivery of antisense oligonucleotides and their application in the exon skipping strategy for Duchenne muscular dystrophy. *Nucleic Acid Ther* 24, 87-100, doi:10.1089/nat.2013.0450 (2014).
- 315 Brewington, J. & Clancy, J. P. Diagnostic Testing in Cystic Fibrosis. *Clin Chest Med* 37, 31-46, doi:10.1016/j.ccm.2015.10.005 (2016).
- 316 McKone, E. F., Emerson, S. S., Edwards, K. L. & Aitken, M. L. Effect of genotype on phenotype and mortality in cystic fibrosis: a retrospective cohort study. *Lancet* 361, 1671-1676, doi:10.1016/S0140-6736(03)13368-5 (2003).
- 317 Tasselli, L. & Chua, K. F. Cancer: Metabolism in 'the driver's seat. *Nature* 492, 362-363, doi:10.1038/492362a (2012).
- 318 Elborn, J. S. Cystic fibrosis. *Lancet*, doi:10.1016/S0140-6736(16)00576-6 (2016).
- 319 Ramalho, A. S. et al. Five percent of normal cystic fibrosis transmembrane conductance regulator mRNA ameliorates the severity of pulmonary disease in cystic fibrosis. *American journal of respiratory cell and molecular biology* 27, 619-627, doi:10.1165/rcmb.2001-0004OC (2002).
- 320 Kerem, E. Pharmacologic therapy for stop mutations: how much CFTR activity is enough? *Curr Opin Pulm Med* 10, 547-552 (2004).
- 321 Karin, M. & Greten, F. R. NF-kappaB: linking inflammation and immunity to cancer development and progression. *Nat Rev Immunol* 5, 749-759, doi:10.1038/nri1703 (2005).
- 322 Whibley, C., Pharoah, P. D. & Hollstein, M. p53 polymorphisms: cancer implications. *Nat Rev Cancer* 9, 95-107, doi:10.1038/nrc2584 (2009).
- 323 Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nature medicine* 10, 789-799, doi:10.1038/nm1087 (2004).
- 324 Biegging, K. T., Mello, S. S. & Attardi, L. D. Unravelling mechanisms of p53-mediated tumour suppression. *Nat Rev Cancer* 14, 359-370, doi:10.1038/nrc3711 (2014).
- 325 Zuckerman, V., Wolynec, K., Sionov, R. V., Haupt, S. & Haupt, Y. Tumour suppression by p53: the importance of apoptosis and cellular senescence. *J Pathol* 219, 3-15, doi:10.1002/path.2584 (2009).
- 326 Soussi, T. & Wiman, K. G. TP53: an oncogene in disguise. *Cell Death Differ* 22, 1239-1249, doi:10.1038/cdd.2015.53 (2015).

- 327 Floquet, C., Deforges, J., Rousset, J. P. & Bidou, L. Rescue of non-sense mutated p53 tumor suppressor gene by aminoglycosides. *Nucleic acids research* 39, 3350-3362, doi:10.1093/nar/gkq1277 (2011).
- 328 Riley, T., Sontag, E., Chen, P. & Levine, A. Transcriptional control of human p53-regulated genes. *Nature reviews. Molecular cell biology* 9, 402-412, doi:10.1038/nrm2395 (2008).
- 329 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* 144, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
- 330 Weinstein, I. B. Disorders in cell circuitry during multistage carcinogenesis: the role of homeostasis. *Carcinogenesis* 21, 857-864 (2000).
- 331 Weinstein, I. B. Cancer. Addiction to oncogenes--the Achilles heel of cancer. *Science* 297, 63-64, doi:10.1126/science.1073096 (2002).
- 332 Nagel-Wolfrum, K., Moller, F., Penner, I., Baasov, T. & Wolfrum, U. Targeting Nonsense Mutations in Diseases with Translational Read-Through-Inducing Drugs (TRIDs). *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy* 30, 49-74, doi:10.1007/s40259-016-0157-6 (2016).
- 333 Tan, Y. C., Blumenfeld, J. & Rennert, H. Autosomal dominant polycystic kidney disease: genetics, mutations and microRNAs. *Biochimica et biophysica acta* 1812, 1202-1212, doi:10.1016/j.bbadis.2011.03.002 (2011).
- 334 Du, L. et al. Nonaminoglycoside compounds induce readthrough of nonsense mutations. *J Exp Med* 206, 2285-2297, doi:10.1084/jem.20081940 (2009).
- 335 Goljan, E. F. *Rapid Review Pathology*. (Mosby/Elsevier, 2010).
- 336 Brendel, C., Klahold, E., Gartner, J. & Huppke, P. Suppression of nonsense mutations in Rett syndrome by aminoglycoside antibiotics. *Pediatr Res* 65, 520-523, doi:10.1203/PDR.0b013e31819d9ebc (2009).
- 337 Goldmann, T., Overlack, N., Wolfrum, U. & Nagel-Wolfrum, K. PTC124-mediated translational readthrough of a nonsense mutation causing Usher syndrome type 1C. *Human Gene Therapy* 22, 537-547, doi:10.1089/hum.2010.067 (2011).
- 338 Perez, B., Rodriguez-Pombo, P., Ugarte, M. & Desviat, L. R. Readthrough strategies for therapeutic suppression of nonsense mutations in inherited metabolic disease. *Mol Syndromol* 3, 230-236, doi:10.1159/000343086 msy-0003-0230 (2012).
- 339 Stenson, P. D. et al. Human Gene Mutation Database: towards a comprehensive central mutation database. *Journal of medical genetics* 45, 124-126, doi:10.1136/jmg.2007.055210 (2008).
- 340 James, P. D. et al. Aminoglycoside suppression of nonsense mutations in severe hemophilia. *Blood* 106, 3043-3048, doi:10.1182/blood-2005-03-1307 (2005).
- 341 Worgan, L. C. et al. Spectrum of mutations in mut methylmalonic acidemia and identification of a common Hispanic mutation and haplotype. *Human mutation* 27, 31-43, doi:10.1002/humu.20258 (2006).
- 342 Chamberlain, J. R. & Chamberlain, J. S. Muscling in: Gene therapies for muscular dystrophy target RNA. *Nature medicine* 16, 170-171, doi:10.1038/nm0210-170 (2010).
- 343 Zamecnik, P. C. & Stephenson, M. L. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proceedings of the National Academy of Sciences of the United States of America* 75, 280-284 (1978).

- 344 Riviere, C., Danos, O. & Douar, A. M. Long-term expression and repeated administration of AAV type 1, 2 and 5 vectors in skeletal muscle of immunocompetent adult mice. *Gene therapy* 13, 1300-1308, doi:10.1038/sj.gt.3302766 (2006).
- 345 Pramono, Z. A. et al. Induction of exon skipping of the dystrophin transcript in lymphoblastoid cells by transfecting an antisense oligodeoxynucleotide complementary to an exon recognition sequence. *Biochemical and biophysical research communications* 226, 445-449 (1996).
- 346 Dunckley, M. G., Manoharan, M., Villiet, P., Eperon, I. C. & Dickson, G. Modification of splicing in the dystrophin gene in cultured Mdx muscle cells by antisense oligoribonucleotides. *Human molecular genetics* 7, 1083-1090 (1998).
- 347 Malerba, A. et al. Chronic systemic therapy with low-dose morpholino oligomers ameliorates the pathology and normalizes locomotor behavior in mdx mice. *Molecular therapy : the journal of the American Society of Gene Therapy* 19, 345-354, doi:10.1038/mt.2010.261 (2011).
- 348 Tortora, G. J., Funke, B. R. & Case, C. L. *Microbiology: An Introduction, 9th Edition*. 912 (Benjamin Cummings, 2006).
- 349 Aartsma-Rus, A. et al. Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. *Human mutation* 30, 293-299, doi:10.1002/humu.20918 (2009).
- 350 Benchaouir, R., Robin, V. & Goyenvalle, A. Gene and splicing therapies for neuromuscular diseases. *Frontiers in bioscience* 20, 1190-1233 (2015).
- 351 Temple, G. F., Dozy, A. M., Roy, K. L. & Kan, Y. W. Construction of a functional human suppressor tRNA gene: an approach to gene therapy for beta-thalassaemia. *Nature* 296, 537-540 (1982).
- 352 Karijolich, J., Yi, C. & Yu, Y. T. Transcriptome-wide dynamics of RNA pseudouridylation. *Nature reviews. Molecular cell biology* 16, 581-585, doi:10.1038/nrm4040 (2015).
- 353 Panchal, R. G., Wang, S., McDermott, J. & Link, C. J., Jr. Partial functional correction of xeroderma pigmentosum group A cells by suppressor tRNA. *Hum Gene Ther* 10, 2209-2219, doi:10.1089/10430349950017194 (1999).
- 354 Buvoli, M., Buvoli, A. & Leinwand, L. A. Suppression of nonsense mutations in cell culture and mice by multimerized suppressor tRNA genes. *Molecular and cellular biology* 20, 3116-3124 (2000).
- 355 Kiselev, A. V. et al. [Suppression of nonsense mutations in the Dystrophin gene by a suppressor tRNA gene]. *Molekuliarnaia biologiiia* 36, 43-47 (2002).
- 356 Giege, R., Sissler, M. & Florentz, C. Universal rules and idiosyncratic features in tRNA identity. *Nucleic acids research* 26, 5017-5035 (1998).
- 357 Laski, F. A., Ganguly, S., Sharp, P. A., RajBhandary, U. L. & Rubin, G. M. Construction, stable transformation, and function of an amber suppressor tRNA gene in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* 86, 6696-6698 (1989).
- 358 Keeling, K. M., Xue, X., Gunn, G. & Bedwell, D. M. Therapeutics based on stop codon readthrough. *Annu Rev Genomics Hum Genet* 15, 371-394, doi:10.1146/annurev-genom-091212-153527 (2014).

- 359 Huang, C., Wu, G. & Yu, Y. T. Inducing nonsense suppression by targeted pseudouridylation. *Nature protocols* 7, 789-800, doi:10.1038/nprot.2012.029 (2012).
- 360 Linde, L. et al. Nonsense-mediated mRNA decay affects nonsense transcript levels and governs response of cystic fibrosis patients to gentamicin. *The Journal of clinical investigation* 117, 683-692, doi:10.1172/JCI28523 (2007).
- 361 Viegas, M. H., Gehring, N. H., Breit, S., Hentze, M. W. & Kulozik, A. E. The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the Nonsense Mediated Decay pathway. *Nucleic acids research* 35, 4542-4551, doi:10.1093/nar/gkm461 (2007).
- 362 Khajavi, M., Inoue, K. & Lupski, J. R. Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease. *European journal of human genetics : EJHG* 14, 1074-1081, doi:10.1038/sj.ejhg.5201649 (2006).
- 363 Kerem, E. et al. Effectiveness of PTC124 treatment of cystic fibrosis caused by nonsense mutations: a prospective phase II trial. *Lancet* 372, 719-727, doi:S0140-6736(08)61168-X 10.1016/S0140-6736(08)61168-X (2008).
- 364 Wilschanski, M. et al. Gentamicin-induced correction of CFTR function in patients with cystic fibrosis and CFTR stop mutations. *The New England journal of medicine* 349, 1433-1441, doi:10.1056/NEJMoa022170 (2003).
- 365 Usuki, F. et al. Inhibition of nonsense-mediated mRNA decay rescues the phenotype in Ullrich's disease. *Annals of neurology* 55, 740-744, doi:10.1002/ana.20107 (2004).
- 366 Durand, S. et al. Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. *The Journal of cell biology* 178, 1145-1160, doi:10.1083/jcb.200611086 (2007).
- 367 Martin, L. et al. Identification and characterization of small molecules that inhibit nonsense-mediated RNA decay and suppress nonsense p53 mutations. *Cancer research* 74, 3104-3113, doi:10.1158/0008-5472.CAN-13-2235 (2014).
- 368 Gonzalez-Hilarion, S. et al. Rescue of nonsense mutations by amlexanox in human cells. *Orphanet J Rare Dis* 7, 58, doi:10.1186/1750-1172-7-58 (2012).
- 369 Keeling, K. M. et al. Attenuation of nonsense-mediated mRNA decay enhances in vivo nonsense suppression. *PloS one* 8, e60478, doi:10.1371/journal.pone.0060478 (2013).
- 370 Popp, M. W. & Maquat, L. E. Attenuation of nonsense-mediated mRNA decay facilitates the response to chemotherapeutics. *Nature communications* 6, 6632, doi:10.1038/ncomms7632 (2015).
- 371 Nomakuchi, T. T., Rigo, F., Aznarez, I. & Krainer, A. R. Antisense oligonucleotide-directed inhibition of nonsense-mediated mRNA decay. *Nature biotechnology* 34, 164-166, doi:10.1038/nbt.3427 (2016).
- 372 McIlwain, D. R. et al. Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. *Proceedings of the National Academy of Sciences of the United States of America* 107, 12186-12191, doi:10.1073/pnas.1007336107 (2010).
- 373 Weischenfeldt, J. et al. NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. *Genes & development* 22, 1381-1396, doi:10.1101/gad.468808 (2008).

- 374 Mutyam, V. et al. Discovery of Clinically Approved Agents That Promote Suppression of CFTR Nonsense Mutations. *American journal of respiratory and critical care medicine*, doi:10.1164/rccm.201601-0154OC (2016).
- 375 Barton-Davis, E. R., Cordier, L., Shoturma, D. I., Leland, S. E. & Sweeney, H. L. Aminoglycoside antibiotics restore dystrophin function to skeletal muscles of mdx mice. *The Journal of clinical investigation* 104, 375-381, doi:10.1172/JCI7866 (1999).
- 376 Dunant, P., Walter, M. C., Karpati, G. & Lochmuller, H. Gentamicin fails to increase dystrophin expression in dystrophin-deficient muscle. *Muscle Nerve* 27, 624-627, doi:10.1002/mus.10341 (2003).
- 377 Keeling, K. M., Du, M. & Bedwell, D. M. *Therapies of nonsens-associated diseases*. Chapter 10 (2006).
- 378 Bidou, L. et al. Premature stop codons involved in muscular dystrophies show a broad spectrum of readthrough efficiencies in response to gentamicin treatment. *Gene therapy* 11, 619-627, doi:10.1038/sj.gt.3302211 (2004).
- 379 Howard, M. T. et al. Readthrough of dystrophin stop codon mutations induced by aminoglycosides. *Annals of neurology* 55, 422-426, doi:10.1002/ana.20052 (2004).
- 380 Matalonga, L. et al. Effect of Readthrough Treatment in Fibroblasts of Patients Affected by Lysosomal Diseases Caused by Premature Termination Codons. *Neurotherapeutics* 12, 874-886, doi:10.1007/s13311-015-0368-4 (2015).
- 381 Allamand, V. et al. Drug-induced readthrough of premature stop codons leads to the stabilization of laminin alpha2 chain mRNA in CMD myotubes. *The journal of gene medicine* 10, 217-224, doi:10.1002/jgm.1140 (2008).
- 382 Shalev, M. & Baasov, T. When Proteins Start to Make Sense: Fine-tuning Aminoglycosides for PTC Suppression Therapy. *Medchemcomm* 5, 1092-1105, doi:10.1039/C4MD00081A (2014).
- 383 Wagner, K. R. et al. Gentamicin treatment of Duchenne and Becker muscular dystrophy due to nonsense mutations. *Annals of neurology* 49, 706-711 (2001).
- 384 Politano, L. et al. Gentamicin administration in Duchenne patients with premature stop codon. Preliminary results. *Acta myologica : myopathies and cardiomyopathies : official journal of the Mediterranean Society of Myology / edited by the Gaetano Conte Academy for the study of striated muscle diseases* 22, 15-21 (2003).
- 385 Malik, V. et al. Gentamicin-induced readthrough of stop codons in Duchenne muscular dystrophy. *Annals of neurology* 67, 771-780, doi:10.1002/ana.22024 (2010).
- 386 Bedwell, D. M. et al. Suppression of a CFTR premature stop mutation in a bronchial epithelial cell line. *Nature medicine* 3, 1280-1284 (1997).
- 387 Du, M. et al. Aminoglycoside suppression of a premature stop mutation in a Cftr<sup>-/-</sup> mouse carrying a human CFTR-G542X transgene. *Journal of molecular medicine* 80, 595-604, doi:10.1007/s00109-002-0363-1 (2002).
- 388 Wilschanski, M. et al. A pilot study of the effect of gentamicin on nasal potential difference measurements in cystic fibrosis patients carrying stop mutations. *American journal of respiratory and critical care medicine* 161, 860-865, doi:10.1164/ajrccm.161.3.9904116 (2000).
- 389 Sermet-Gaudelus, I. et al. In vitro prediction of stop-codon suppression by intravenous gentamicin in patients with cystic fibrosis: a pilot study. *BMC medicine* 5, 5, doi:10.1186/1741-7015-5-5 (2007).

- 390 Clancy, J. P. et al. Evidence that systemic gentamicin suppresses premature stop mutations in patients with cystic fibrosis. *American journal of respiratory and critical care medicine* 163, 1683-1692, doi:10.1164/ajrccm.163.7.2004001 (2001).
- 391 Keeling, K. M. & Bedwell, D. M. Clinically relevant aminoglycosides can suppress disease-associated premature stop mutations in the IDUA and P53 cDNAs in a mammalian translation system. *Journal of molecular medicine* 80, 367-376, doi:10.1007/s00109-001-0317-z (2002).
- 392 Zilberberg, A., Lahav, L. & Rosin-Arbesfeld, R. Restoration of APC gene function in colorectal cancer cells by aminoglycoside- and macrolide-induced read-through of premature termination codons. *Gut* 59, 496-507, doi:10.1136/gut.2008.169805 (2010).
- 393 Lubamba, B., Dhooghe, B., Noel, S. & Leal, T. Cystic fibrosis: insight into CFTR pathophysiology and pharmacotherapy. *Clin Biochem* 45, 1132-1144, doi:10.1016/j.clinbiochem.2012.05.034 (2012).
- 394 Lopez-Novoa, J. M., Quiros, Y., Vicente, L., Morales, A. I. & Lopez-Hernandez, F. J. New insights into the mechanism of aminoglycoside nephrotoxicity: an integrative point of view. *Kidney Int* 79, 33-45, doi:10.1038/ki.2010.337 (2011).
- 395 Xie, J., Talaska, A. E. & Schacht, J. New developments in aminoglycoside therapy and ototoxicity. *Hear Res* 281, 28-37, doi:10.1016/j.heares.2011.05.008 (2011).
- 396 Goldmann, T. et al. A comparative evaluation of NB30, NB54 and PTC124 in translational read-through efficacy for treatment of an USH1C nonsense mutation. *EMBO molecular medicine* 4, 1186-1199, doi:10.1002/emmm.201201438 (2012).
- 397 Shulman, E. et al. Designer aminoglycosides that selectively inhibit cytoplasmic rather than mitochondrial ribosomes show decreased ototoxicity: a strategy for the treatment of genetic diseases. *The Journal of biological chemistry* 289, 2318-2330, doi:10.1074/jbc.M113.533588 (2014).
- 398 Mattis, V. B. et al. Novel aminoglycosides increase SMN levels in spinal muscular atrophy fibroblasts. *Hum Genet* 120, 589-601, doi:10.1007/s00439-006-0245-7 (2006).
- 399 Le, T. T. et al. SMNDelta7, the major product of the centromeric survival motor neuron (SMN2) gene, extends survival in mice with spinal muscular atrophy and associates with full-length SMN. *Human molecular genetics* 14, 845-857, doi:10.1093/hmg/ddi078 (2005).
- 400 Heier, C. R. & DiDonato, C. J. Translational readthrough by the aminoglycoside geneticin (G418) modulates SMN stability in vitro and improves motor function in SMA mice in vivo. *Human molecular genetics* 18, 1310-1322, doi:10.1093/hmg/ddp030 (2009).
- 401 Mattis, V. B., Ebert, A. D., Fosso, M. Y., Chang, C. W. & Lorson, C. L. Delivery of a read-through inducing compound, TC007, lessens the severity of a spinal muscular atrophy animal model. *Human molecular genetics* 18, 3906-3913, doi:10.1093/hmg/ddp333 (2009).
- 402 Nudelman, I. et al. Redesign of aminoglycosides for treatment of human genetic diseases caused by premature stop mutations. *Bioorganic & medicinal chemistry letters* 16, 6310-6315, doi:10.1016/j.bmcl.2006.09.013 (2006).

- 403 Brendel, C. et al. Readthrough of nonsense mutations in Rett syndrome: evaluation of novel aminoglycosides and generation of a new mouse model. *Journal of molecular medicine* 89, 389-398, doi:10.1007/s00109-010-0704-4 (2011).
- 404 Nudelman, I. et al. Development of novel aminoglycoside (NB54) with reduced toxicity and enhanced suppression of disease-causing premature stop mutations. *Journal of medicinal chemistry* 52, 2836-2845, doi:10.1021/jm801640k (2009).
- 405 Kondo, S. & Hotta, K. Semisynthetic aminoglycoside antibiotics: Development and enzymatic modifications. *J Infect Chemother* 5, 1-9, doi:10.1007/s101569900000 (1999).
- 406 Russell, R. J., Murray, J. B., Lentzen, G., Haddad, J. & Mobashery, S. The complex of a designer antibiotic with a model aminoacyl site of the 30S ribosomal subunit revealed by X-ray crystallography. *Journal of the American Chemical Society* 125, 3410-3411, doi:10.1021/ja029736h (2003).
- 407 Lee, H. L., Chen, C. C., Baasov, T., Ron, Y. & Dougherty, J. P. Post-transcriptionally regulated expression system in human xenogeneic transplantation models. *Molecular therapy : the journal of the American Society of Gene Therapy* 19, 1645-1655, doi:10.1038/mt.2011.90 (2011).
- 408 Kandasamy, J. et al. Increased selectivity toward cytoplasmic versus mitochondrial ribosome confers improved efficiency of synthetic aminoglycosides in fixing damaged genes: a strategy for treatment of genetic diseases caused by nonsense mutations. *Journal of medicinal chemistry* 55, 10630-10643, doi:10.1021/jm3012992 (2012).
- 409 Wang, D. et al. The designer aminoglycoside NB84 significantly reduces glycosaminoglycan accumulation associated with MPS I-H in the Idua-W392X mouse. *Molecular genetics and metabolism* 105, 116-125, doi:10.1016/j.ymgme.2011.10.005 (2012).
- 410 Sabbavarapu, N. M. et al. Design of Novel Aminoglycoside Derivatives with Enhanced Suppression of Diseases-Causing Nonsense Mutations. *ACS medicinal chemistry letters* 7, 418-423, doi:10.1021/acsmedchemlett.6b00006 (2016).
- 411 Nudelman, I. et al. Repairing faulty genes by aminoglycosides: development of new derivatives of geneticin (G418) with enhanced suppression of diseases-causing nonsense mutations. *Bioorganic & medicinal chemistry* 18, 3735-3746, doi:10.1016/j.bmc.2010.03.060 (2010).
- 412 Gunn, G. et al. Long-term nonsense suppression therapy moderates MPS I-H disease progression. *Molecular genetics and metabolism* 111, 374-381, doi:10.1016/j.ymgme.2013.12.007 (2014).
- 413 Xue, X. et al. Synthetic aminoglycosides efficiently suppress cystic fibrosis transmembrane conductance regulator nonsense mutations and are enhanced by ivacaftor. *American journal of respiratory cell and molecular biology* 50, 805-816, doi:10.1165/rcmb.2013-0282OC (2014).
- 414 Taguchi, A. et al. Negamycin analogue with readthrough-promoting activity as a potential drug candidate for duchenne muscular dystrophy. *ACS medicinal chemistry letters* 3, 118-122, doi:10.1021/ml200245t (2012).
- 415 Guo, J. et al. Pharmacokinetics of the natural antibiotic negamycin. *Xenobiotica* 45, 625-633, doi:10.3109/00498254.2015.1006301 (2015).

- 416 Welch, E. M. et al. PTC124 targets genetic disorders caused by nonsense mutations. *Nature* 447, 87-91, doi:10.1038/nature05756 (2007).
- 417 Gatti, R. A. SMRT compounds correct nonsense mutations in primary immunodeficiency and other genetic models. *Annals of the New York Academy of Sciences* 1250, 33-40, doi:10.1111/j.1749-6632.2012.06467.x (2012).
- 418 Kayali, R. et al. Read-through compound 13 restores dystrophin expression and improves muscle function in the mdx mouse model for Duchenne muscular dystrophy. *Human molecular genetics* 21, 4007-4020, doi:10.1093/hmg/dds223 (2012).
- 419 Lee, P. et al. SMRT compounds abrogate cellular phenotypes of ataxia telangiectasia in neural derivatives of patient-specific hiPSCs. *Nature communications* 4, 1824, doi:10.1038/ncomms2824 (2013).
- 420 Gomez-Grau, M. et al. Evaluation of Aminoglycoside and Non-Aminoglycoside Compounds for Stop-Codon Readthrough Therapy in Four Lysosomal Storage Diseases. *PLoS one* 10, e0135873, doi:10.1371/journal.pone.0135873 (2015).
- 421 Du, L. et al. A New Series of Small Molecular Weight Compounds Induce Read Through of All Three Types of Nonsense Mutations in the ATM Gene. *Molecular Therapy*, doi:10.1038/mt.2013.150 mt2013150 (2013).
- 422 Peltz, S. W., Morsy, M., Welch, E. M. & Jacobson, A. Ataluren as an agent for therapeutic nonsense suppression. *Annu Rev Med* 64, 407-425, doi:10.1146/annurev-med-120611-144851 (2013).
- 423 Auld, D. S., Thorne, N., Maguire, W. F. & Inglese, J. Mechanism of PTC124 activity in cell-based luciferase assays of nonsense codon suppression. *Proceedings of the National Academy of Sciences of the United States of America* 106, 3585-3590, doi:0813345106 10.1073/pnas.0813345106 (2009).
- 424 Peltz, S. W. et al. Nonsense suppression activity of PTC124 (ataluren). *Proceedings of the National Academy of Sciences of the United States of America* 106, E64; author reply E65, doi:10.1073/pnas.0901936106 (2009).
- 425 Inglese, J., Thorne, N. & Auld, D. S. Reply to Peltz et al: Post-translational stabilization of the firefly luciferase reporter by PTC124 (Ataluren). *Proceedings of the National Academy of Sciences* 106, E65, doi:10.1073/pnas.0905457106 (2009).
- 426 Dranchak, P. K. et al. Nonsense suppressor therapies rescue peroxisome lipid metabolism and assembly in cells from patients with specific PEX gene mutations. *J Cell Biochem* 112, 1250-1258, doi:10.1002/jcb.22979 (2011).
- 427 McElroy, S. P. et al. A lack of premature termination codon read-through efficacy of PTC124 (Ataluren) in a diverse array of reporter assays. *PLoS biology* 11, e1001593, doi:10.1371/journal.pbio.1001593 (2013).
- 428 Cogan, J. et al. Aminoglycosides restore full-length type VII collagen by overcoming premature termination codons: therapeutic implications for dystrophic epidermolysis bullosa. *Molecular therapy : the journal of the American Society of Gene Therapy* 22, 1741-1752, doi:10.1038/mt.2014.140 (2014).
- 429 Du, M. et al. PTC124 is an orally bioavailable compound that promotes suppression of the human CFTR-G542X nonsense allele in a CF mouse model. *Proceedings of the National Academy of Sciences of the United States of America* 105, 2064-2069, doi:0711795105 10.1073/pnas.0711795105 (2008).

- 430 Tan, L., Narayan, S. B., Chen, J., Meyers, G. D. & Bennett, M. J. PTC124 improves readthrough and increases enzymatic activity of the CPT1A R160X nonsense mutation. *Journal of inherited metabolic disease* 34, 443-447, doi:10.1007/s10545-010-9265-5 (2011).
- 431 Yu, H. et al. Comparison of read-through effects of aminoglycosides and PTC124 on rescuing nonsense mutations of HERG gene associated with long QT syndrome. *International journal of molecular medicine* 33, 729-735, doi:10.3892/ijmm.2013.1601 (2014).
- 432 Kerem, E. et al. Ataluren for the treatment of nonsense-mutation cystic fibrosis: a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet Respir Med* 2, 539-547, doi:10.1016/S2213-2600(14)70100-6 (2014).
- 433 Hirawat, S. et al. Safety, tolerability, and pharmacokinetics of PTC124, a nonaminoglycoside nonsense mutation suppressor, following single- and multiple-dose administration to healthy male and female adult volunteers. *J Clin Pharmacol* 47, 430-444, doi:Doi 10.1177/0091270006297140 (2007).
- 434 Hirawat, S. et al. Safety, tolerability, and pharmacokinetics of PTC124, a nonaminoglycoside nonsense mutation suppressor, following single- and multiple-dose administration to healthy male and female adult volunteers. *J Clin Pharmacol* 47, 430-444, doi:10.1177/0091270006297140 (2007).
- 435 Sermet-Gaudelus, I. et al. Ataluren (PTC124) induces cystic fibrosis transmembrane conductance regulator protein expression and activity in children with nonsense mutation cystic fibrosis. *American journal of respiratory and critical care medicine* 182, 1262-1272, doi:10.1164/rccm.201001-0137OC (2010).
- 436 Wilschanski, M. et al. Chronic ataluren (PTC124) treatment of nonsense mutation cystic fibrosis. *Eur Respir J* 38, 59-69, doi:10.1183/09031936.00120910 (2011).
- 437 Nguyen, L. S., Wilkinson, M. F. & Gecz, J. Nonsense-mediated mRNA decay: inter-individual variability and human disease. *Neurosci Biobehav Rev* 46 Pt 2, 175-186, doi:10.1016/j.neubiorev.2013.10.016 (2014).
- 438 Finkel, R. S. et al. Phase 2a study of ataluren-mediated dystrophin production in patients with nonsense mutation duchenne muscular dystrophy. *PloS one* 8, e81302, doi:10.1371/journal.pone.0081302 (2013).
- 439 Finkel, R. S. Read-through strategies for suppression of nonsense mutations in Duchenne/ Becker muscular dystrophy: aminoglycosides and ataluren (PTC124). *Journal of child neurology* 25, 1158-1164, doi:10.1177/0883073810371129 (2010).
- 440 Sirtori, C. R. Aescin: pharmacology, pharmacokinetics and therapeutic profile. *Pharmacol Res* 44, 183-193, doi:10.1006/phrs.2001.0847 (2001).
- 441 Cruz-Vera, L. R., Sachs, M. S., Squires, C. L. & Yanofsky, C. Nascent polypeptide sequences that influence ribosome function. *Current opinion in microbiology* 14, 160-166, doi:10.1016/j.mib.2011.01.011 (2011).
- 442 Fernandez, I. S., Bai, X. C., Murshudov, G., Scheres, S. H. & Ramakrishnan, V. Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* 157, 823-831, doi:10.1016/j.cell.2014.04.015 (2014).
- 443 Fischer, N. et al. Structure of the E. coli ribosome-EF-Tu complex at <3 Å resolution by C-corrected cryo-EM. *Nature*, doi:10.1038/nature14275 (2015).

- 444 Markova, S. V., Golz, S., Frank, L. A., Kalthof, B. & Vysotski, E. S. Cloning and expression of cDNA for a luciferase from the marine copepod *Metridia longa*. A novel secreted bioluminescent reporter enzyme. *The Journal of biological chemistry* 279, 3212-3217, doi:10.1074/jbc.M309639200 (2004).
- 445 Stepanyuk, G. A. et al. Expression, purification and characterization of the secreted luciferase of the copepod *Metridia longa* from Sf9 insect cells. *Protein Expr Purif* 61, 142-148, doi:10.1016/j.pep.2008.05.013 (2008).
- 446 Matt, T. et al. Dissociation of antibacterial activity and aminoglycoside ototoxicity in the 4-monosubstituted 2-deoxystreptomycin apramycin. *Proceedings of the National Academy of Sciences of the United States of America* 109, 10984-10989, doi:10.1073/pnas.1204073109 (2012).
- 447 Zhang, X. D. Novel analytic criteria and effective plate designs for quality control in genome-scale RNAi screens. *J Biomol Screen* 13, 363-377, doi:10.1177/1087057108317062 (2008).
- 448 Zhang, X. D. A pair of new statistical parameters for quality control in RNA interference high-throughput screening assays. *Genomics* 89, 552-561, doi:10.1016/j.ygeno.2006.12.014 (2007).
- 449 Lupold, S. E., Johnson, T., Chowdhury, W. H. & Rodriguez, R. A real time *Metridia* luciferase based non-invasive reporter assay of mammalian cell viability and cytotoxicity via the beta-actin promoter and enhancer. *PloS one* 7, e36535, doi:10.1371/journal.pone.0036535 (2012).
- 450 Markova, S. V., Burakova, L. P. & Vysotski, E. S. High-active truncated luciferase of copepod *Metridia longa*. *Biochemical and biophysical research communications* 417, 98-103, doi:10.1016/j.bbrc.2011.11.063 (2012).
- 451 Ohmiya, Y. & Wu, C. Stabilizing composition and stabilizing method of coelenterazine solution for high-throughput measurement of luciferase activity (Google Patents, 2010).
- 452 Goktug, A. N., Chai, S. C. C. & Chen, T. *Data analysis approaches in high throughput screening*. (2013).
- 453 Zhang, X. D. Illustration of SSMD, z score, SSMD\*, z\* score, and t statistic for hit selection in RNAi high-throughput screens. *J Biomol Screen* 16, 775-785, doi:10.1177/1087057111405851 (2011).
- 454 Wang, C. S. et al. Establishment and characterization of a new cell line derived from a human primary breast carcinoma. *Cancer genetics and cytogenetics* 120, 58-72 (2000).
- 455 Floquet, C., Rousset, J. P. & Bidou, L. [Rescue of nonsense mutated p53 by readthrough leads to apoptosis in cancers cells]. *Medecine sciences : M/S* 27, 585-586, doi:10.1051/medsci/2011276008 (2011).
- 456 Duttler, S., Pechmann, S. & Frydman, J. Principles of cotranslational ubiquitination and quality control at the ribosome. *Molecular cell* 50, 379-393, doi:10.1016/j.molcel.2013.03.010 (2013).
- 457 Heck, J. W., Cheung, S. K. & Hampton, R. Y. Cytoplasmic protein quality control degradation mediated by parallel actions of the E3 ubiquitin ligases Ubr1 and San1. *Proceedings of the National Academy of Sciences of the United States of America* 107, 1106-1111, doi:10.1073/pnas.0910591107 (2010).

- 458 Petrov, A. & Puglisi, J. D. Site-specific labeling of *Saccharomyces cerevisiae* ribosomes for single-molecule manipulations. *Nucleic acids research* 38, e143, doi:10.1093/nar/gkq390 (2010).
- 459 Fizman, N. et al. in *OptDiag 2012 Diagnostic et Imagerie Optique en Médecine et Biologie, Biophotonique*. 1 (<https://hal-iogs.archives-ouvertes.fr/hal-00708294/file/Fizman.pdf>).
- 460 Murray, J. et al. Structural characterization of ribosome recruitment and translocation by type IV IRES. *eLife* 5, doi:10.7554/eLife.13567 (2016).
- 461 Petrov, A., Grosely, R., Chen, J., O'Leary, S. E. & Puglisi, J. D. Multiple Parallel Pathways of Translation Initiation on the CrPV IRES. *Molecular cell* 62, 92-103, doi:10.1016/j.molcel.2016.03.020 (2016).
- 462 Zhang, H., Ng, M. Y., Chen, Y. & Cooperman, B. S. Kinetics of initiating polypeptide elongation in an IRES-dependent system. *eLife* 5, doi:10.7554/eLife.13429 (2016).
- 463 Lander, E. S. et al. Initial sequencing and analysis of the human genome. *Nature* 409, 860-921, doi:10.1038/35057062 (2001).
- 464 Scott, M. P., Storti, R. V., Pardue, M. L. & Rich, A. Cell-free protein synthesis in lysates of *Drosophila melanogaster* cells. *Biochemistry* 18, 1588-1594 (1979).
- 465 Boehlke, K. W. & Friesen, J. D. Cellular content of ribonucleic acid and protein in *Saccharomyces cerevisiae* as a function of exponential growth rate: calculation of the apparent peptide chain elongation rate. *Journal of bacteriology* 121, 429-433 (1975).
- 466 Hershey, J. W. Translational control in mammalian cells. *Annual review of biochemistry* 60, 717-755, doi:10.1146/annurev.bi.60.070191.003441 (1991).
- 467 David, A. et al. RNA binding targets aminoacyl-tRNA synthetases to translating ribosomes. *The Journal of biological chemistry* 286, 20688-20700, doi:10.1074/jbc.M110.209452 (2011).
- 468 Negrutskii, B. S., Stapulionis, R. & Deutscher, M. P. Supramolecular organization of the mammalian translation system. *Proceedings of the National Academy of Sciences of the United States of America* 91, 964-968 (1994).
- 469 Baradaran-Heravi, A. et al. Novel small molecules potentiate premature termination codon readthrough by aminoglycosides. *Nucleic acids research*, doi:10.1093/nar/gkw638 (2016).
- 470 Chang, J., Kim, Y. & Kwon, H. J. Advances in identification and validation of protein targets of natural products without chemical modification. *Nat Prod Rep* 33, 719-730, doi:10.1039/c5np00107b (2016).
- 471 Hoja, U., Wellein, C., Greiner, E. & Schweizer, E. Pleiotropic phenotype of acetyl-CoA-carboxylase-defective yeast cells--viability of a BPL1-amber mutation depending on its readthrough by normal tRNA(Gln)(CAG). *European journal of biochemistry / FEBS* 254, 520-526 (1998).

# ***Matériels et Méthodes***

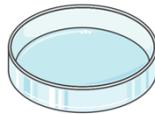
***Pour doubler votre taux de succès,  
doublez votre taux d'échec !!  
(Thomas Watson)***

### Culture cellulaire 3T3pML213



**Boîte culture traité cellules 60 cm<sup>2</sup>**  
Environ 3,8M de cellules/boîte  
3T3pML213 (C14)  
Milieu DMEM 10% (10mL)

### Détachement des cellules



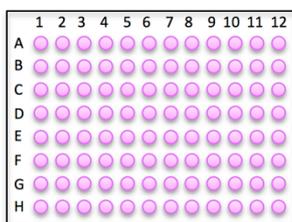
**Traitement à la trypsine (2mL)**  
5' 37°C, 5% CO<sub>2</sub>

### Comptage des cellules



**Scepter avec sondes 60µm**  
Suspension cellulaire :  
1boîte/13mL milieu

### Plaque cellules



**Plaques 96 blanche fond blanc traité cellules**  
2500 cellules/puits  
24H 37°C, 5% CO<sub>2</sub>

### Dilution et Ensemencement en plaque 96



25000 cellules/mL

100µL/puits

**Figure 1 : Etape d'ensemencement du crible à grande échelle**

Les cellules 3T3pML213 sont mises en culture puis détachées par traitement à la trypsine et remises en suspension dans du milieu de culture. Les cellules sont alors comptées et diluées à 25000 cellules/mL. Enfin, les cellules sont ensemencées en plaque 96 puits de façon semi-automatisée puis incubées durant 24h.

## Matériel et Méthodes

### I. Le crible à grande échelle

#### 1. Protocole de réalisation du crible

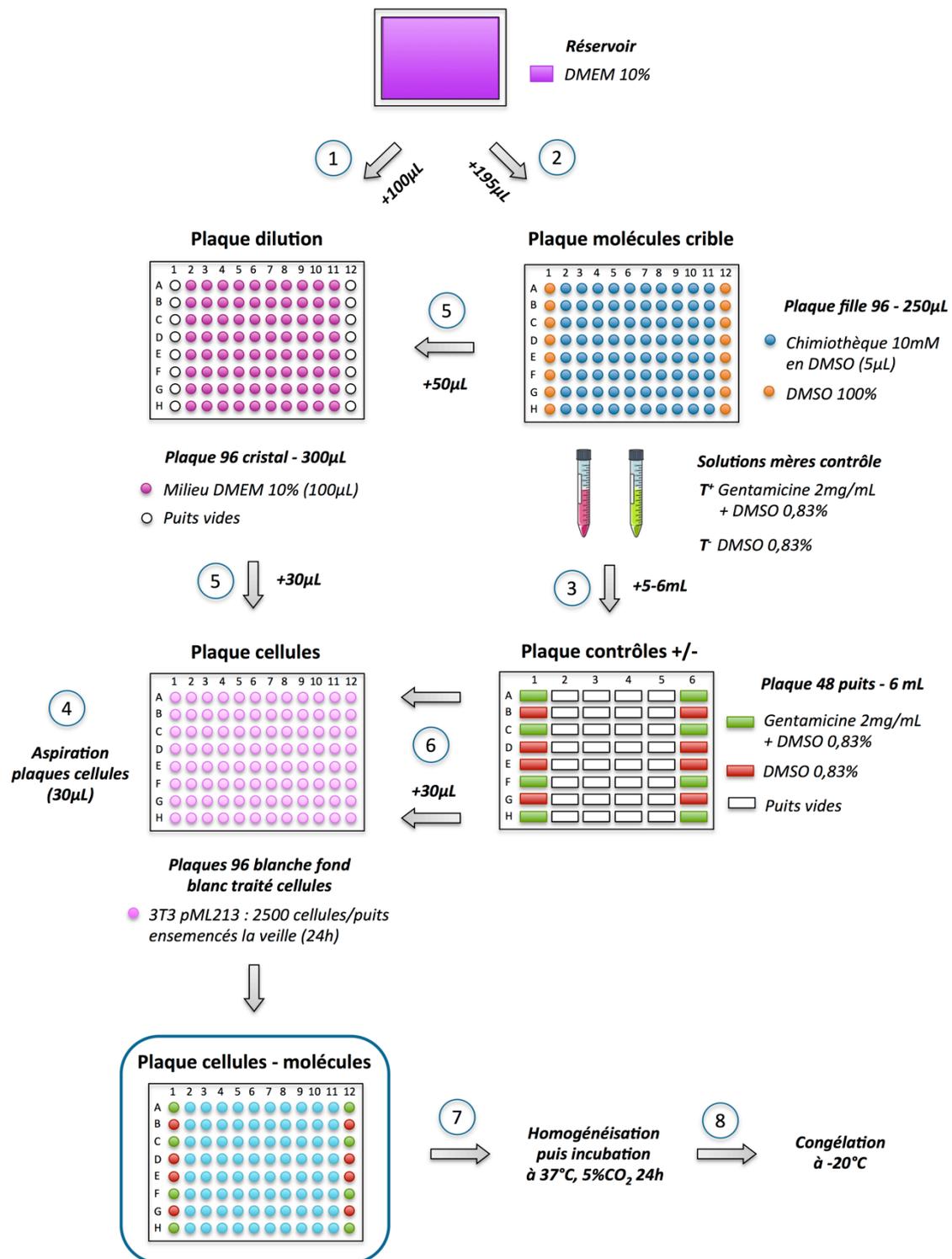
##### ★ Chimiothèque

- Deux banques de molécules sont testées : Prestwick (14 plaques) et Chembridge (206 plaques 96 puits). Toutefois, je ne décrirai le protocole du crible soit un total de 220 plaques.
- Des plaques contrôles négatifs/positifs sont également réalisées : 2 toute les 25 plaques 96 puits de chimiothèques soit 18 plaques.
- Des plaques contrôles culture sont prévues pour vérifier au microscope l'ensemencement : 1 toute les 25 plaques 96 puits de chimiothèques soit 5 plaques soit 9 plaques.
- Un total de 250 plaques 96 puits est donc nécessaire pour réaliser le criblage.

##### ★ Ensemencement

*Cette étape est schématisée dans la figure 1.*

- Les cellules 3T3pML213 (C14) sont mises en culture dans des boîtes de culture traitées cellules de 60 cm<sup>2</sup> dans du milieu DMEM Glutamax (LifeTech) (10mL) supplémenté en sérum (10%). Les cellules sont incubées à 37°C et 5% de CO<sub>2</sub>. Une boîte à 80% de confluence correspond environ à 3,8 millions de cellules.
- Deux ensemencements de 125 plaques sont réalisés. Cela nécessite donc pour 125 plaques, 1,2 L de culture cellulaire à 25000 cellules/mL auquel nous ajoutons 0,2 L de marge soit un total de 1,4 L, ce qui représente environ 8 boîtes de culture à confluence.
- Les cellules sont détachées par un traitement à la trypsine (2 mL/boîte) avec une incubation de 5' à 37°C et 5% de CO<sub>2</sub>. Les cellules sont ensuite transférées dans du milieu de culture DMEM 10% dans un erlenmeyer sous agitation lente (1 boîte dans 13mL de milieu). La suspension cellulaire est comptée 5 fois avec un Scepter (Millipore) et des sondes 60 µm. La suspension est ajustée à une concentration de 25 000 cellules par mL.



**Figure 2 : Etape d'ajout des molécules lors du crible à grande échelle**

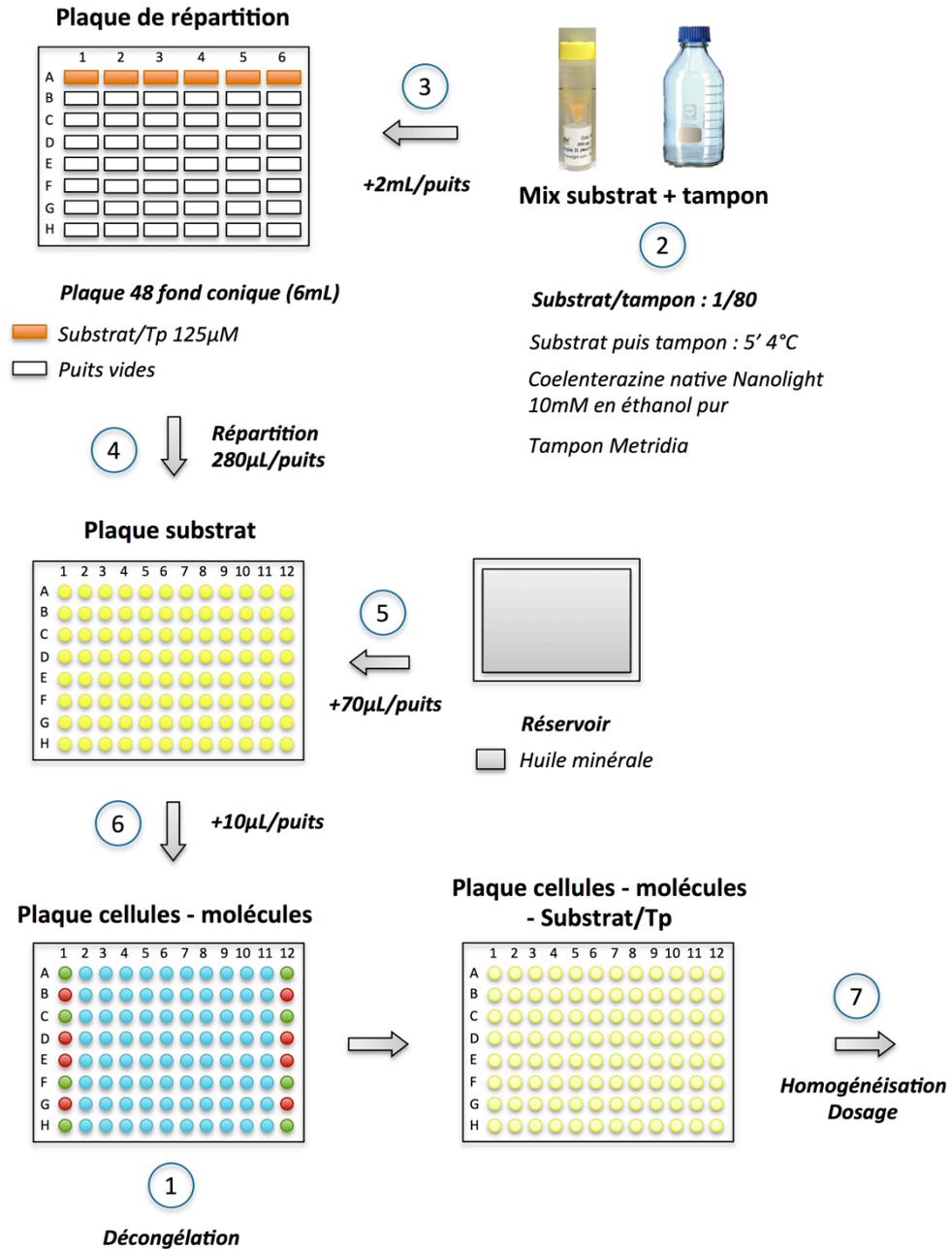
Les plaques de molécules sont diluées deux fois successivement dans du milieu de culture. Les solutions contrôles positifs (gentamicine) et négatifs sont préparées en tubes puis disposées dans une plaque. Le milieu de culture cellulaire est alors aspiré partiellement et les molécules et contrôles sont alors ajoutées sur les cellules. Les plaques de cellules sont replacées dans l'incubateur pendant 24h puis sont congelées.

- L'ensemencement est semi-automatisé en conditions stériles avec un Multidrop (Thermo). La tête d'aspiration est placée dans la suspension cellulaire et 100  $\mu$ L sont distribués dans chacun des 96 puits des 250 plaques de culture soit 2500 cellules par puits. Les plaques traitées et les contrôles positifs et négatifs sont des plaques 96 puits blanches fond blanc traitées cellules (Corning). Les plaques contrôles culture sont des plaques 96 puits blanches fond transparent traitées cellules <sup>471</sup>. Après ensemencement, ne surtout pas homogénéiser !!
- Les plaques sont ensuite incubées 24 heures à 37°C et 5% CO<sub>2</sub>.

#### ★ Ajout des molécules

*Cette étape est schématisée dans la figure 2.*

- Du milieu de culture DMEM 10% est versé dans une plaque réservoir. Ce dernier va permettre de préparer les plaques dilutions et diluer les plaques de molécules des chimiothèques.
- Dans un premier temps, 100 $\mu$ L de milieu provenant du réservoir sont déposés dans les puits des colonnes 2 à 11 des plaques de dilution à l'aide du Liquidator qui a 96 têtes d'aspiration (Rainin). Ces plaques sont transparentes et ont un couvercle pour limiter les contaminations. Leur contenance est de 300 $\mu$ L. <sup>22</sup>
- A partir de ce même réservoir, 195 $\mu$ L de milieu est transféré dans les plaques de molécules des chimiothèques. Ces plaques sont appelées plaques fille et ont une contenance de 250 $\mu$ L et contiennent 5 $\mu$ L de molécule par puits en DMSO 100% à 10mM dans les colonnes 2 à 11. Les colonnes 1 et 12 contiennent 5 $\mu$ L de DMSO 100%. A l'issue de cette dilution, les molécules sont à 0,25mM et 2,5M de DMSO. <sup>22</sup>
- Parallèlement, les solutions contrôles sont préparées dans des tubes 50mL. Le contrôle positif est une solution de gentamicine (Sigma) à 2mg/mL avec 0,83% de DMSO dans du milieu de culture. Le témoin négatif est une solution de milieu de culture cellulaire avec 0,83% de DMSO. Ces contrôles sont ajoutés dans une plaque 48 puits d'une contenance de 6mL telle que présentée sur le schéma (Figure 2) (3).
- Au dernier moment, le milieu de culture des plaques cellules ensemencées la veille, par groupe de 5, sont aspirées par une tête 96 d'aspiration. Sur les 100 $\mu$ L de milieu, il reste 30 $\mu$ L après aspiration. <sup>22</sup>
- Parallèlement, 50 $\mu$ L des plaques filles (molécules) sont transvasées dans la plaque de dilution avec le Liquidator et homogénéisées. Les molécules sont alors à 83 $\mu$ M dans 0,83% de DMSO. Immédiatement après, 30 $\mu$ L de ces molécules sont prélevés et déposés dans les plaques cellules aspirées. Les molécules sont alors à 50 $\mu$ L et 0,5% de DMSO. (5)



**Figure 3 : Etape de dosage du crible à grande échelle**

Les plaques de cultures traitées sont décongelées. Le substrat est ensuite repris dans de l'éthanol pur puis dilué dans du tampon. La solution tampon/substrat est alors répartie une première fois dans une plaque de 48 puits puis une seconde fois dans une plaque de 96 puits. Une fine pellicule d'huile minérale est déposée sur le mix tampon/substrat et la plaque est maintenue à 4°C. Le substrat est ensuite ajouté au milieu de culture d'une plaque cellules-molécules et la plaque est homogénéisée puis dosée par un luminomètre à plaques.

- Cette étape est suivie du transfert de 30 $\mu$ L des contrôles positifs et négatifs dans les colonnes 1 et 12 de la plaque cellules aspirée et contenant les molécules. Les contrôles positifs sont après dilution à 1,2 mg/mL et 0,5% de DMSO. Les contrôles négatifs sont à 0,5% de DMSO. (6)
- Les plaques cellules + molécules sont alors homogénéisées et incubées à 37°C et 5% de CO<sub>2</sub> pendant 24 heures. <sup>22</sup>
- Les étapes 4 à 7 sont répétées jusqu'à ce toutes les plaques soient traitées.
- Les plaques sont ensuite congelées à -20°C. (8)

### ★ Dosage

*Cette étape est schématisée dans la figure 3.*

- Le dosage est réalisé sur des groupes de 20 plaques.
- Le tampon Metridia (10mL) est constitué comme suit :
  - NaCl 5M  $\rightarrow$  1,1M : 2,2mL
  - Na<sub>2</sub>EDTA 0,1M  $\rightarrow$  2,2mM : 220 $\mu$ L
  - K<sub>2</sub>HPO<sub>4</sub> 1M pH5,1  $\rightarrow$  0,22M : 2,2mL
  - NaN<sub>3</sub> 1M  $\rightarrow$  0,1M : 130 $\mu$ L
  - H<sub>2</sub>O : 5,25mL
- Les plaques sont tout d'abord décongelées passivement dans la salle climatisée à 16°C et assombrie. <sup>22</sup>
- Le tube de 500 $\mu$ g de substrat, la coelenterazine native (Nanolight), est repris à 10mM en éthanol pur puis bullé à l'azote. Puis, un mix substrat / tampon est réalisé à 125 $\mu$ M finale en coelenterazine en tube 15mL pour limiter la surface d'oxydation. Ce mix est bullé à l'azote et incubé 5' à 4°C. <sup>22</sup>
- 2mL de mix tampon/substrat sont ensuite répartis dans la première ligne d'une plaque 48 puits de 6mL de contenance à la pipette monocanal. (3)
- 280 $\mu$ L du mix substrat tampon de la plaque de répartition sont ensuite transférées à l'aide d'une pipette multicanaux dans une plaque substrat à fond conique pour minimiser le volume mort. Cette étape est immédiatement suivie d'une centrifugation à 500g pendant 30 secondes. <sup>22</sup>
- Puis, 70 $\mu$ L d'huile minérale (Sigma) sont déposés à l'aide du Liquidator en surface du mix tampon/substrat de la plaque de substrat. Cette plaque est maintenue à 4°C. (5)
- A l'aide du Liquidator, 10 $\mu$ L de mix tampon/substrat de la plaque substrat sont prélevés et déposés dans la plaque cellules/molécules décongelée. (6)
- La plaque est immédiatement homogénéisée manuellement par un mouvement orbital puis placée dans le luminomètre à plaque Spectramax M5e pour être dosé.

<sup>22</sup>

- Les valeurs sont ensuite enregistrées dans un tableau Excel (Microsoft) créé pour l'analyse de nos données.
- Les étapes 1 à 5 sont réalisées pour des jeux de 20 plaques. Les étapes 6 et 7 sont réalisées pour chaque plaque.

## 2. Normalisation des résultats

### ★ Normalisation

- Les données brutes sont dans un premier temps normalisées.
- Dans un premier temps, la pente des contrôles positifs et négatifs est calculée en fonction de leur position de lecture sur la plaque 96 à l'aide de la fonction « PENTE ».
- Les contrôles positifs sont alors normalisés selon leur pente. Les contrôles négatifs et les puits traités sont normalisés par la pente des contrôles négatifs. La formule utilisée pour la normalisation est :  $V_{brute_n} - (Position_n * V_{pente})$  où n correspond à chacun des 96 puits.

## II. La molécule unique

### 1. Transcription *in vitro*

Attention, les minipreps doivent être réalisées avec des solutions sans RNAses.

#### ★ Linéarisation du plasmide et purification de l'ADN

- Mélanger : 6µg de plasmide + 30µl de tampon 10x + 6µl d'enzyme de restriction, qsp 300µL avec de l'eau ultrapure. Incuber pendant 1h30 à 37°C.
- Afin d'ôter tous les sels qui pourraient inhiber la transcription, purifier le plasmide sur une colonne NucleoSpin PCR Clean Up (Macherey-Nagel). Eluer avec 60µl d'H<sub>2</sub>O DEPC.

#### ★ Transcription

- La transcription commence au dernier G du promoteur SP6.
- Mélanger : 10µl du mix NTP à 10mM + 5µl de DTT à 0,1M + 5µl de tampon 10x SP6 RNA polymérase + 16 à 20µL de plasmide linéarisé + 2µl de polymérase SP6 (NEB), qsp 50µL avec de l'eau DEPC. Incuber à 40°C pendant 3h30. Vérifier la présence d'ARN en analysant 5µl d'échantillon par électrophorèse.

★ **Traitement à la DnaseI et purification**

- Ce traitement a pour but de dégrader la matrice ADN qui pourrait titrer les oligonucléotides lors de l'appariement. Ajouter 5µL de tampon DNaseI et 3µL de DNaseI (NEB) pour 50µL d'échantillon et incuber à 37°C pendant 1h. Ajouter 2µL de DNaseI dans chacun des tubes et incuber de nouveau 1h à 37°C. Vérifier la digestion par électrophorèse. P
- La purification de l'ARNm est réalisée à l'aide d'une colonne « Mini quick RNA spin column » (Roche).
- Pour la microscopie, mesurer la DO à 260nm et adapter la concentration pour avoir environ 1µM d'ARNm.

## 2. Traduction sur billes magnétiques avec de la méthionine S<sup>35</sup>

★ **Produits**

- Le tampon A (20mM HEPES pH 7,8 ; 150mM KAc ; 6mM MgAc ; 2mM DTT) et le tampon de fixation 5x (100mM HEPES pH 7,8 ; 500mM KAc ; 25mM MgAc ; 5mM DTT) doivent être filtrés avant utilisation. **RRL** (Promega)
- Methionine S35 (PerkinElmer - NEG009A500UC). 400µM RNOB, 1,03 µM ARNm PPS0 (digéré par Kpn1) et 600nM de sous-unités ribosomiques purifiées à partir de RRL.

★ **Avant de commencer**

- Penser à sortir la radioactivité afin qu'elle décongèle (au moins 30' avant)
- Allumer les bains-marie/à sec : 78°C, 37°C et 30°C et 30 et 70°C.
- Décongeler dans la glace l'ensemble des solutions

★ **Appariement ARNm-RNOB**

- Mélanger 16,5µl de ARNmPPS0 + 2,5µL RNOB + 1µl de tampon 10x T4 DNA ligase (Promega).
- Incuber 5min à 68°C, 5min à 37°C et garder dans la glace 10'. (Concentrations finales : 0,46µM ARNm ; 50µM RNOB)

★ **Préparation des billes magnétiques**

- Utiliser un tube de « Streptavidin MagneSphere® Paramagnetic Particles » (Promega) par échantillon.

- Laver les billes aimantées 3 fois avec 600µl de tampon A pendant les temps d'incubation d'appariement du RNOB et des ribosomes avec l'ARNm.

★ **Accroche des ribosomes sur l'ARNm/RNOB**

	<b>T-</b>	<b>Ech RRL + Ribos 5pmol</b>	<b>T+ RRL</b>
ARNm PPS0/RNOB	/	5µL	5µL
Tp fixation 5X	3µL	3µL	/
Ribosomes purifiés	/	9µL	/
Eau DEPC	14µL	/	/
Incubation 30°	15'	15'	/
+ Tampon A	50µL	50µL	50µL
Incubation à TA	15'	15'	/

★ **Fixation sur billes**

- Après avoir ôté le dernier tampon de rinçage des billes, ajouter les complexes ARNm/RNOB/±ribosomes sur les billes aimantées pour les fixer à la streptavidine.
- Laisser incuber à température ambiante pendant 15min en agitant manuellement toutes les 5'. Puis rincer les billes 2 fois avec 500µl de tampon A.

★ **Traduction avec de la Methionine S<sup>35</sup>**

- Après avoir ôté le dernier tampon de rinçage des billes, faire un mix :
  - o 216,2µL RRL + 1,61µL aa-met + 1,61µL aa-leu + 6,9µL met S35 soit 226,32µL
- Ajouter 49,5µL dans chacun des tubes puis incuber à 30°C pendant 1h dans un bain à sec.
- Conserver les surnageants et laver les billes 3 fois avec 100µl de tampon A.

★ **Migration sur gel de protéines**

- Reprendre les billes directement dans 30µl de Laemmli 1X.
- Prélever 1µL des surnageants et ajouter 14 µL de Laemmli 1X.
- Chauffer les extraits seulement 2min à 70°C.

- Analyser en faisant migrer les échantillons dans un gel Nupage 4-12% Bis-Tris avec tampon MES à 200V constants pendant au moins 35 minutes. Les bandes les plus faibles (10 et 15 kDa) du Ladder doivent sortir. La bande attendue sort entre 35-40kDa.

**★ Séchage du gel**

- Commencer par couper au scalpel les bords du gel et retirer le stacking.
- Faire adhérer le gel à un papier Whatman légèrement plus grand que le gel (sans bulles). Si le gel n'adhère pas bien, humidifier le gel.
- Retourner le tout et positionner un film étirable sur la face du dessus.
- Faire sécher sous vide sur la plaque chauffante/vide (avec son support 3M si radioéléments) pendant au moins 2 heures. Ne surtout pas casser le vide avant ou le gel est broyé. Lors de l'arrêt, détacher le tuyau d'aspiration avant de couper la pompe pour éviter un reflux puis éteindre le chauffage de la plaque.

**★ Révélation**

- Disposer le gel dans une cassette en le fixant avec du scotch et en retirant le film étirable.
- La cassette comprend deux Hyperfilms. Ajouter un film (Amersham) dans une pièce noire et fermer la cassette. Placer à -80°C ON.
- Sortir la cassette 1h à TA avant d'ouvrir. Sortir le film dans le noir et l'introduire dans l'appareil à révélation des films à autoradiographie.

# ***Annexes***

# Annexes

## Annexe 1 : Nouvelle nomenclature des protéines ribosomales

a)

Cytoplasmic ribosome New name #	Occurrence*	Bacteria Old name	Yeast Old name	Human Old name	Mammalian mitoribosome New name #	Mammalian mitoribosome Old name
bS1	B m	S1	–	–	bS1m	MRPS28
uS2	BAE m	S2	S0	SA	uS2m	MRPS2
uS3	BAE m	S3	S3	S3	uS3m	MRPS24
uS4	BAE	S4	S9	S9	–	–
uS5	BAE m	S5	S2	S2	uS5m	MRPS5
bS6	B m	S6	–	–	bS6m	MRPS6
uS7	BAE m	S7	S5	S5	uS7m	MRPS7
uS8	BAE	S8	S22	S15A	–	–
uS9	BAE m	S9	S16	S16	uS9m	MRPS9
uS10	BAE m	S10	S20	S20	uS10m	MRPS10
uS11	BAE m	S11	S14	S14	uS11m	MRPS11
uS12	BAE m	S12	S23	S23	uS12m	MRPS12
uS13	BAE	S13	S18	S18	–	–
uS14	BAE m	S14	S29	S29	uS14m	MRPS14
uS15	BAE m	S15	S13	S13	uS15m	MRPS15
bS16	B m	S16	–	–	bS16m	MRPS16
uS17	BAE m	S17	S11	S11	uS17m	MRPS17
bS18	B m	S18	–	–	bS18m	MRPS18C
uS19	BAE	S19	S15	S15	–	–
bS20	B	S20	–	–	–	–
bS21	B m	S21	–	–	bS21m	MRPS21
bTHX	B	THX	–	–	–	–
eS1	AE	–	S1	S3A	–	–
eS4	AE	–	S4	S4	–	–
eS6	AE	–	S6	S6	–	–
eS7	E	–	S7	S7	–	–
eS8	AE	–	S8	S8	–	–
eS10	E	–	S10	S10	–	–
eS12	E	–	S12	S12	–	–
eS17	AE	–	S17	S17	–	–
eS19	AE	–	S19	S19	–	–
eS21	E	–	S21	S21	–	–
eS24	AE	–	S24	S24	–	–
eS25	AE	–	S25	S25	–	–
eS26	E	–	S26	S26	–	–
eS27	AE	–	S27	S27	–	–
eS28	AE	–	S28	S28	–	–
eS30	AE	–	S30	S30	–	–
eS31	AE	–	S31	S27A	–	–
RACK1	E	–	Asc1	RACK1	–	–
m	–	–	–	–	mS22	MRPS22
m	–	–	–	–	mS23	MRPS23
m	–	–	–	–	mS25	MRPS25
m	–	–	–	–	mS26	MRPS26
m	–	–	–	–	mS27	MRPS27
m	–	–	–	–	mS29 (DAP3)	MRPS29, DAP3
m	–	–	–	–	mS31	MRPS31
m	–	–	–	–	mS33	MRPS33
m	–	–	–	–	mS34	MRPS34
m	–	–	–	–	mS35	MRPS35
m	–	–	–	–	mS37	MRPS37, CHCHD1
m	–	–	–	–	mS38	MRPS38, AURKAIP1
m	–	–	–	–	mS39	MRPS39, PTC3
m	–	–	–	–	mS40	MSRPS18B

# b: bacterial, e: eukaryotic, u: universal, m: mitochondrial  
\* B: Bacteria, A: Archaea, E: Eukaryotes, m: Mitochondria



bacterial 30S  
small subunit

■ ■  
B BAE



eukaryotic 40S  
small subunit

■ ■ ■  
E AE BAE



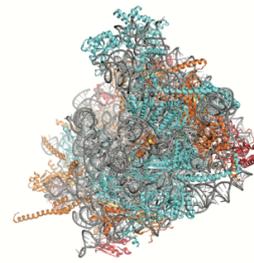
mammalian 28S  
mitochondrial small subunit

■ ■ ■  
B BAE m

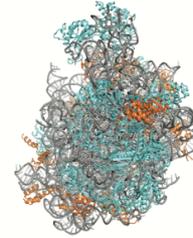
b)

Cytoplasmic ribosome					Mitochondrial ribosome		
New Name*	Occurrence*	Bacteria Old name	Yeast Old name	Human Old name	New name*	Mammalian mitoribosome Old name	Yeast mitoribosome Old name
uL1	BAE m	L1	L1	L10A	uL1m	MRPL1	MRPL1
uL2	BAE m	L2	L2	L8	uL2m	MRPL2	RML2
uL3	BAE m	L3	L3	L3	uL3m	MRPL3	MRPL9
uL4	BAE m	L4	L4	L4	uL4m	MRPL4	YML6
uL5	BAE m	L5	L11	L11	uL5m	–	MRPL7
uL6	BAE m	L6	L9	L9	uL6m	–	MRPL6
bl9	B m	L9	–	–	bl9m	MRPL9	MRPL50
uL10	BAE m	L10	P0	P0	uL10m	MRPL10	MRPL11
uL11	BAE m	L11	L12	L12	uL11m	MRPL11	MRPL19
bl12	B m	L7/L12	–	–	bl12m	MRPL12	MNP1
uL13	BAE m	L13	L16	L13A	uL13m	MRPL13	MRPL23
uL14	BAE m	L14	L23	L23	uL14m	MRPL14	MRPL38
uL15	BAE m	L15	L28	L27A	uL15m	MRPL15	MRPL10
uL16	BAE m	L16	L10	L10	uL16m	MRPL16	MRPL16
bl17	B m	L17	–	–	bl17m	MRPL17	MRPL8
uL18	BAE m	L18	L5	L5	uL18m	MRPL18	–
bl19	B m	L19	–	–	bl19m	MRPL19	IMG1
bl20	B m	L20	–	–	bl20m	MRPL20	–
bl21	B m	L21	–	–	bl21m	MRPL21	MRPL49
uL22	BAE m	L22	L17	L17	uL22m	MRPL22	MRPL22
uL23	BAE m	L23	L25	L23A	uL23m	MRPL23	MRP20 (MRPL41)
uL24	BAE m	L24	L26	L26	uL24m	MRPL24	MRPL40
bl25	B	L25	–	–	–	–	–
bl27	B m	L27	–	–	bl27m	MRPL27	MRP7
bl28	B m	L28	–	–	bl28m	MRPL28	MRPL24
uL29	BAE m	L29	L35	L35	uL29m	MRPL47	MRPL4
uL30	BAE m	L30	L7	L7	uL30m	MRPL30	MRPL33
bl31	B m	L31	–	–	bl31m	MRPL55	MRPL36
bl32	B m	L32	–	–	bl32m	MRPL32	MRPL32
bl33	B m	L33	–	–	bl33m	MRPL33	MRPL39
bl34	B m	L34	–	–	bl34m	MRPL34	MRPL34
bl35	B m	L35	–	–	bl35m	MRPL35	YNL122C
bl36	B m	L36	–	–	bl36m	MRPL36	RTC6
eL6	E	–	L6	L6	–	–	–
eL8	AE	–	L8	L7A	–	–	–
eL13	AE	–	L13	L13	–	–	–
eL14	AE	–	L14	L14	–	–	–
eL15	AE	–	L15	L15	–	–	–
eL18	AE	–	L18	L18	–	–	–
eL19	AE	–	L19	L19	–	–	–
eL20 <sup>†</sup>	AE	–	L20	L18A	–	–	–
eL21	AE	–	L21	L21	–	–	–
eL22	E	–	L22	L22	–	–	–
eL24	AE	–	L24	L24	–	–	–
eL27	E	–	L27	L27	–	–	–
eL28	E	–	–	L28	–	–	–
eL29	E	–	L29	L29	–	–	–
eL30	AE	–	L30	L30	–	–	–
eL31	AE	–	L31	L31	–	–	–
eL32	AE	–	L32	L32	–	–	–
eL33	AE	–	L33	L35A	–	–	–
eL34	AE	–	L34	L34	–	–	–
eL36	E	–	L36	L36	–	–	–
eL37	AE	–	L37	L37	–	–	–
eL38	AE	–	L38	L38	–	–	–
eL39	AE	–	L39	L39	–	–	–
eL40	AE	–	L40	L40	–	–	–
eL41	AE	–	L41	L41	–	–	–
eL42	AE	–	L42	L36A	–	–	–
eL43	AE	–	L43	L37A	–	–	–
P1/P2	AE	–	P1/P2 (AB)	P1/P2 (αβ)	–	–	–
–	m	–	–	–	mL37	MRPL37	–
–	m	–	–	–	mL38	MRPL38	MRPL35
–	m	–	–	–	mL39	MRPL39	–
–	m	–	–	–	mL40	MRPL40	MRPL28
–	m	–	–	–	mL41	MRPL41	MRPL27
–	m	–	–	–	mL42	MRPL42	–
–	m	–	–	–	mL43	MRPL43	MRPL51
–	m	–	–	–	mL44	MRPL44	MRPL3
–	m	–	–	–	mL45	MRPL45	–
–	m	–	–	–	mL46	MRPL46	MRPL17
–	m	–	–	–	mL48	MRPL48	–
–	m	–	–	–	mL49	MRPL49	IMG2
–	m	–	–	–	mL50	MRPL50	MRPL13
–	m	–	–	–	mL51	MRPL51	–
–	m	–	–	–	mL52	MRPL52	–
–	m	–	–	–	mL53	MRPL53	MRPL44
–	m	–	–	–	mL54	MRPL54	MRPL37
–	m	–	–	–	mL57	–	MRPL15
–	m	–	–	–	mL58	–	MRPL20
–	m	–	–	–	mL59	–	MRPL25
–	m	–	–	–	mL60	–	MRPL31
–	m	–	–	–	mL61	–	MRP49
–	m	–	–	–	mL62 (ICT1)	MRPL58, ICT1	–
–	m	–	–	–	mL63	MRPL57, MRP63	–
–	m	–	–	–	mL64 (CRIF1)	MRPL59, CRIF1	–
–	m	–	–	–	mL65	MRPS30, PDCC9	–
–	m	–	–	–	mL66	MRPS18A	–
–	m	–	–	–	MHR1	–	MHR1

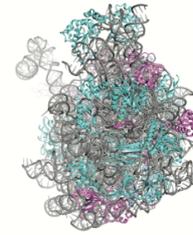
\* b: bacterial, e: eukaryotic, u: universal, m: mitochondrial  
 † B: Bacteria, A: Archaea, E: Eukaryotes, m: Mitochondria  
 ‡ protein LX in Archaea



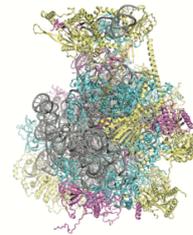
eukaryotic 60S large subunit  
 E AE BAE



archaeal 50S large subunit  
 AE BAE



bacterial 50S large subunit  
 B BAE

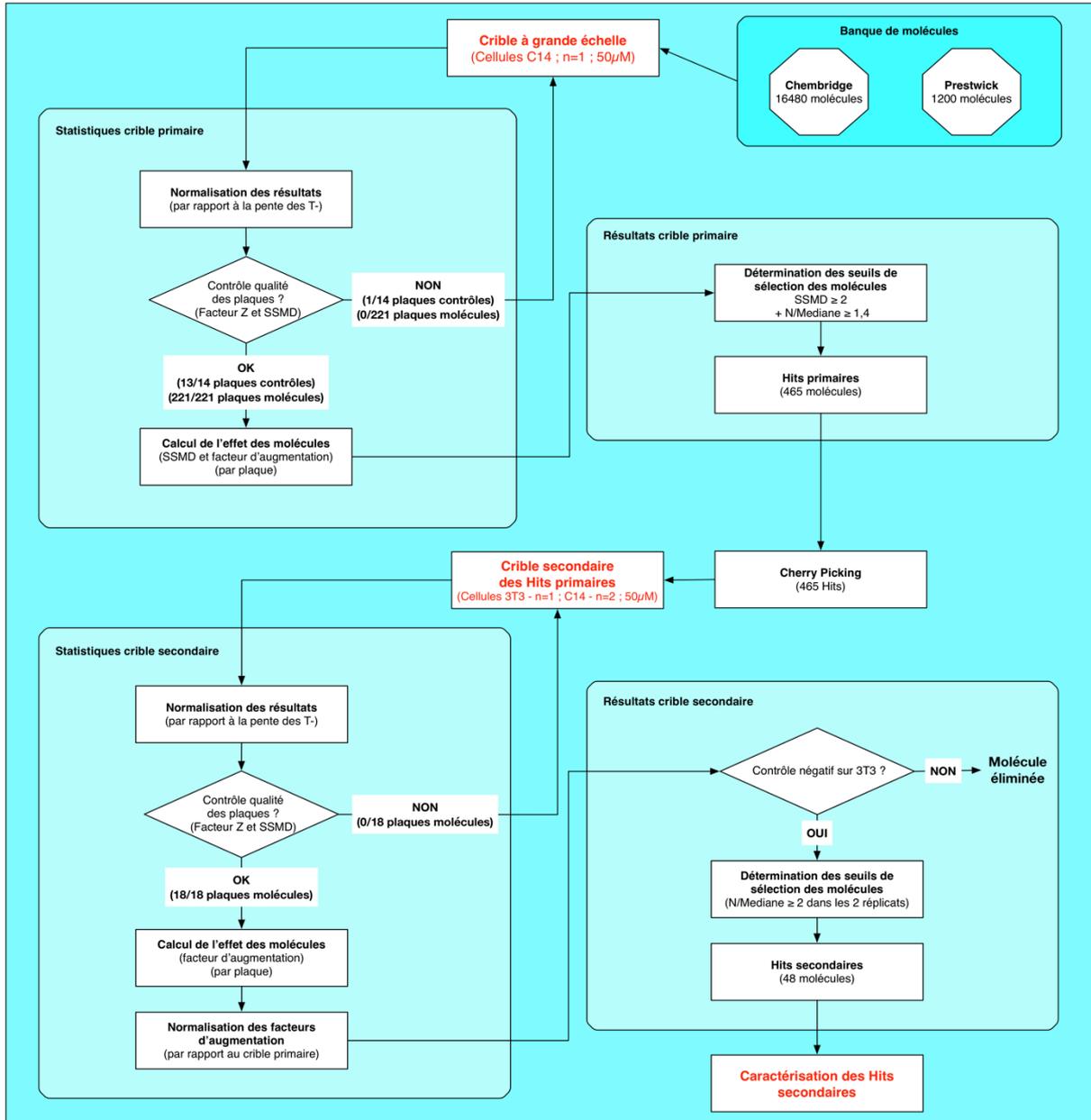


mammalian 39S mitochondrial large subunit  
 B BAE m

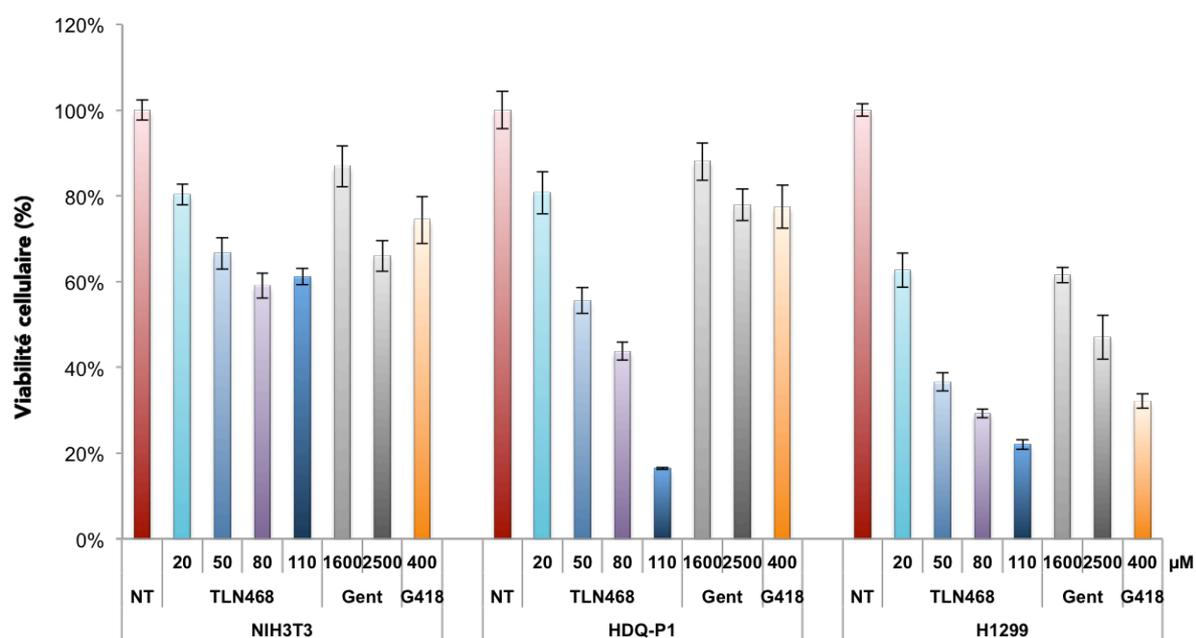


yeast 54S mitochondrial large subunit  
 B BAE m

## Annexe 2 : Arbre décisionnel de l'analyse statistique et des résultats des cribles primaires et secondaires



### Annexe 3 : Viabilité cellulaire des lignées NIH3T3, HDQ-P1 et H1299 après traitement par la molécule TLN468



Les composés G418, gentamicine et TLN468 sont ajoutés aux lignées cellulaires NIH3T3, HDQ-P1 et H1299 durant 24h (n=4). Le WST1 est alors ajouté sur les cellules et l'absorbance est quantifiée. Ce composé est réduit en formazan (détectable à 440nm), qui s'accumule de façon proportionnelle au nombre de cellules vivantes. La viabilité cellulaire est alors déterminée en faisant le rapport des mesures d'absorbance dans la condition traitée par rapport à la médiane des absorbances de la condition contrôle.

## Annexe 4 : Analyses statistiques de la contribution des ribosomes du RRL

**Expériences de contrôle sans pré-incuber les ribosomes.** Nous avons fait 3 séries d'expériences sans préincuber les ribosomes donnant des signaux exploitables en statistique suffisante. Les résultats des fits sont dans le tableau ci-dessous. Les barres d'erreur proviennent d'une procédure de 500 bootstraps.

	mars 2016	avril 2016	juillet 2016
durée entre 2 images	5s	2.5s	2.5s
ON objectif	1.45	1.45	1.49
nb de points exploitables	8000	15 à 20 000	3 à 4 000
reproductibilité	TB	TB	bonne
rapport signal à bruit	B	B	TB
départs spontanés	limités	limités	limités
fits pour R-DOWN			
A	$0.25 \pm 0.02$	$0.15 \pm 0.02$	$0.15 \pm 0.02$
$t_{el}$ (s)	0.3	0.3	0.4
$t_0$ (s)		22	22
fits pour G-UP			
A	$0.15 \pm 0.02$	$0.07 \pm 0.02$	$0.14 \pm 0.02$
$t_{el}$ (s)	NaN	0.6	NaN
$t_0$ (s)		10	25

Cette procédure visait surtout à déterminer A. Très peu de contraintes ont été mises sur  $t_{el}$ , ce qui mène à une mauvaise évaluation de ce paramètre. Mais ça n'impacte pas la suite du raisonnement.

On peut conclure de cette série d'expérience que les ribosomes libres du RRL sont capables de s'initier et de traduire 10 à 25 % d'une population d'ARNm libres (i.e. sans ribosomes pré-incubés).

**Expériences en préincubant les ribosomes.** On obtenait alors  $A = 50\%$ .

A est significativement plus élevé que si on ne préincube pas les ribosomes. Donc les ribosomes préincubés jouent un rôle important. En conséquence, on considérera que si un ribosome est préincubé sur un ARNm, un ribosome libre du RRL ne viendra pas prendre sa place.

**Conséquence sur les contributions respectives des 2 types de ribosomes sur la statistique.**

On divise les ARNm en 2 catégories :

- ceux qui ont un ribosome préincubé, notés  $n_p$ ,
- ceux qui n'en ont pas, notés  $n_l$ .

On a

$$n_p + n_l = 1 \quad (1)$$

Pour commencer, disons que 15% de  $n_l$  sera traduit, grâce à l'expérience de contrôle. On fera ensuite le même raisonnement pour les valeurs extrêmes de A, soit 10 et 25%.

Il faut aussi, pour satisfaire les résultats de l'expérience de traduction, que

$$0.15n_l + p_p n_p = 0.5 \quad (2)$$

avec  $p_p$  la probabilité qu'un ribosome préincubé traduise sont ARNm.

On a donc 3 inconnues ( $n_p$ ,  $n_l$  et  $p_p$ ) et seulement 2 équations.

On va donc tracer  $p_p$  en fonction de  $n_p$  et on en déduira la contribution des ribosomes pré-incubés (resp. libres dans le RRL) à la statistique totale. On a les équations suivantes :

pour  $p_p$ , en combinant les equations 1 et 2,

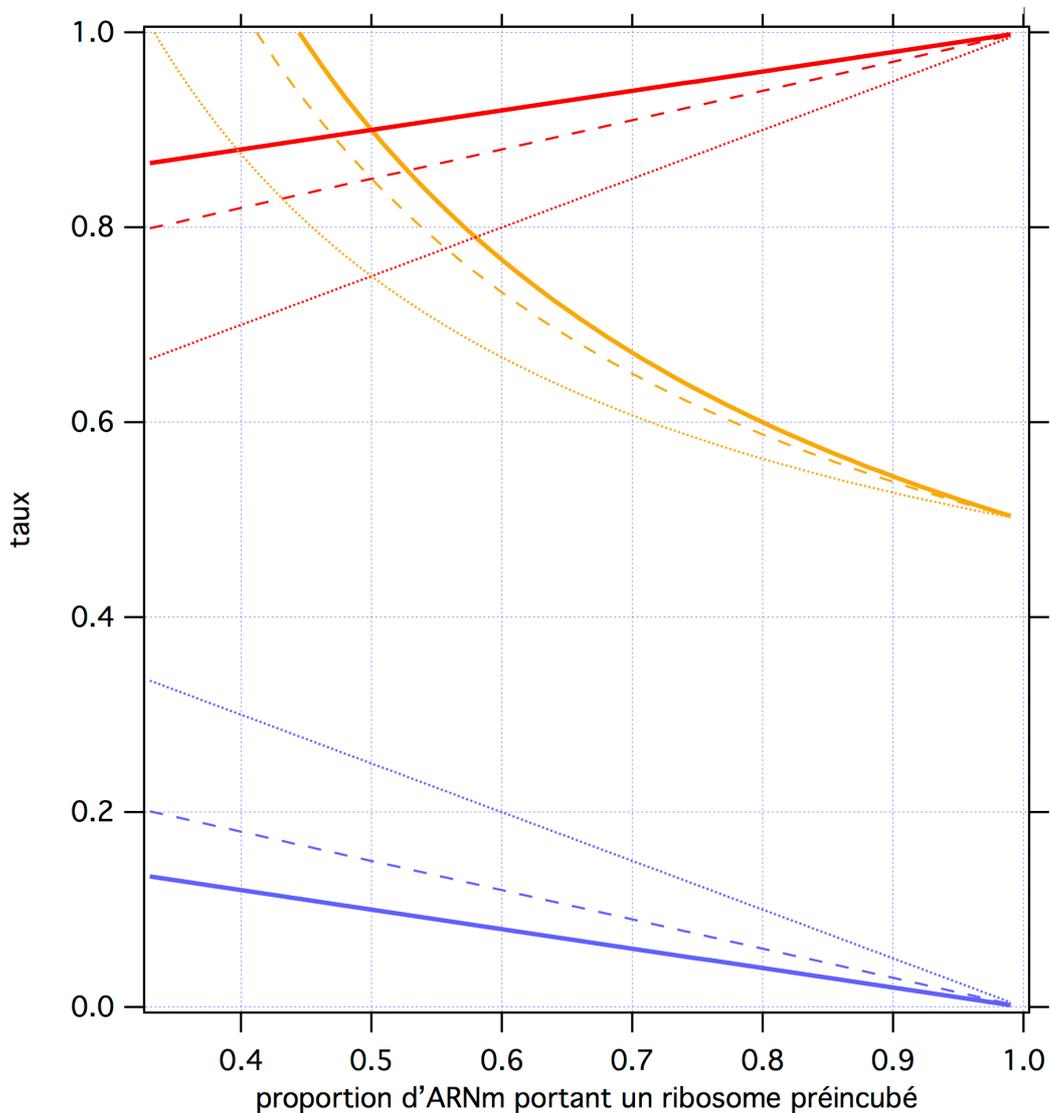
$$p_p = \frac{0.35}{n_p} + 0,15 \quad (3)$$

La contribution des ribosomes du RRL vaut  $2 \times 0.15n_l$  (on multiplie par 2 pour se ramener à 100%, alors que A ne vaut que 50% dans la manip traduction), soit  $2 \times 0.15(1 - n_p)$ .

Celle des ribosomes préincubés vaut nécessairement le complémentaire.

On obtient les courbes suivantes en fonction de  $n_p$ .

hyp : A\_RRL=10%  
 ——— probabilité qu'un ribo préincubé traduise  
 ——— contribution des ribosomes préincubés à la statistique totale  
 ——— contribution des ribosomes du RRL à la statistique totale  
 hyp : A\_RRL=15%  
 - - - probabilité qu'un ribo préincubé traduise  
 - - - contribution des ribosomes préincubés à la statistique totale  
 - - - contribution des ribosomes du RRL à la statistique totale  
 hyp : A\_RRL=25%  
 ..... probabilité qu'un ribo préincubé traduise  
 ..... contribution des ribosomes préincubés à la statistique totale  
 ..... contribution des ribosomes du RRL à la statistique totale



On a rajouté les mêmes courbes pour les valeurs extrêmes de  $A$  dans l'expérience de contrôle.  
 Note : on ne peut pas avoir  $p_p > 1$ . Donc, les points des courbes de contribution des différents ribosomes ne veulent rien dire dans cette situation.

On lit tout d'abord que, pour conserver un pourcentage d'ARNm traduits de 50% dans l'expérience de traduction, il faut nécessairement une contribution importante des ribosomes pré-incubés.

- Un cas limite est que tous les ARNm portent un ribosome préincubé. Dans ce cas, il n'y a aucune contribution des ribosomes du RRL, puisque tous les brins sont déjà occupés, et il suffit que 50% des ribosomes pré-incubés soient actifs.
- Si tous les ARNm ne portent pas un ribosome préincubé, il y aura une contribution des ribosomes du RRL, plus ou moins importante selon les cas. Dans le cas le moins favorable où  $A_{\text{contrôle}}=0.25$ , il faut tout de même au minimum 33% d'ARNm avec un ribosome pré-incubé et une probabilité de 1 que ces ribosomes soient actifs. C'est le point le moins favorable pour nous : les ribosomes préincubés ne contribuent alors qu'à 66% de la statistique.
- Dans l'hypothèse plus raisonnable où  $A_{\text{contrôle}}=0.15$  (valeur moyenne des expériences de contrôle), dans le pire des cas, 82% des ribosomes préincubés participent à la statistique.

**Conclusion** Les ribosomes du RRL n'ont pas de contribution forte à la statistique observée.

**Titre :** Suppression traductionnelle des codons stop chez les mammifères

**Mots clés :** Terminaison de la traduction, mutation non-sens / codon stop prématuré, translecture, aminoglycosides.

**Résumé :** Entre 10% et 30% des maladies génétiques humaines sont liées à l'apparition d'une mutation non-sens (PTC). La synthèse protéique est alors arrêté prématurément. Cet arrêt peut être inhibé par des molécules inductrices de translecture qui permettent l'incorporation d'un ARNt suppresseur naturel au niveau du PTC (translecture). Le ribosome peut alors franchir le PTC et restaurer l'expression de la protéine.

Au cours de ma thèse, je me suis intéressé à la suppression des codons stop en caractérisant de nouvelles molécules inductrices de translecture et en analysant les mécanismes de la fidélité de la traduction. J'ai tout d'abord mis au point un système de criblage innovant avec lequel j'ai testé plus de 17 000 molécules et identifié un hit, la molécule TLN468. La caractérisation de cette molécule a montré qu'elle est capable de supprimer une mutation non-sens retrouvée dans le gène suppresseur de tumeur *P53* et d'induire la réexpression de la protéine p53 active.

J'ai également caractérisé l'efficacité de translecture de nouveaux composés dérivés d'aminoglycosides. J'ai pu montrer que le NB124 est capable d'induire l'apoptose de cellules tumorales via la réexpression de la protéine p53 tout ayant une toxicité bien plus faible que la gentamicine.

En parallèle, j'ai développé une approche en molécule unique permettant d'étudier les erreurs traductionnelles. J'ai ainsi pu analyser la cinétique d'élongation des ribosomes eucaryotes et montré que l'initiation de la traduction sur un site d'entrée interne (IRES) ralentit le ribosome lors des premiers cycles d'élongation. Cet outil permettra d'étudier la suppression traductionnelle des codons stop et d'améliorer les molécules qui ciblent l'appareil traductionnel. L'obtention de données dynamique sera une avancée majeure dans la compréhension de la traduction eucaryote.

**Title :** Translational suppression of stop codons in mammals

**Keywords :** Translation termination, nonsense mutation / premature stop codon, readthrough, aminoglycosides

**Abstract :** Nonsense mutations, also known as premature termination codons (PTCs) are responsible for 10% of all genetic diseases. The presence of a PTC leads to premature translation termination. Translational nonsense suppression strategy aims to re-express a functional full-length protein to restore normal cellular functions by using drugs. These molecules increase the incorporation of near-cognate tRNA at the stop codon (readthrough) and allow ribosomes to continue translation in the same reading frame until the next stop codon.

During my PhD, I characterized new readthrough inducers and studied translation accuracy to better understand the molecular mechanisms involved in stop codon decoding. I first developed an innovative screening system that I used to test over 17,000 molecules. Using a rigorous selection workflow I identified one hit we named TLN468.

We have shown that this molecule is able to suppress various non-sense mutations and induce the re-expression of a full-length p53 protein. I also characterized readthrough efficiency of new compounds derived from aminoglycosides. I have shown that one of them, NB124 is able to induce apoptosis of tumor cells by re-expressing p53 protein, while having a much lower toxicity than gentamicin. In parallel, we developed a single molecule approach for studying translational errors. The analysis of the kinetic elongation of eukaryotic ribosomes showed that translation initiation at an Internal Ribosome Entry Site (IRES) slows ribosomes during the first elongation cycles. This tool will allow us to study the translational non-sense suppression and improve the molecules that target the translational machinery. Obtaining dynamic data will be a major advance in the understanding of the eukaryotic translation.

